uORFs with unusual translational start codons autoregulate expression of eukaryotic ornithine decarboxylase homologs

Iyavol P. Ivanov*,1, Gary Loughran*†, and John F. Atkins**†

*BioSciences Institute, University College Cork, Cork, Ireland; and †Department of Human Genetics, University of Utah, Salt Lake City, UT 84112-5330

In a minority of eukaryotic mRNAs, a small functional upstream ORF (uORF), often performing a regulatory role, precedes the translation start site for the main product(s). Here, conserved uORFs in numerous ornithine decarboxylase homologs are identified from yeast to mammals. Most have noncanonical evolutionarily conserved start codons, the main one being AUU, which has not been known as an initiator for eukaryotic chromosomal genes. The AUG-less uORF present in mouse antizyme inhibitor, one of the ornithine decarboxylase homologs in mammals, mediates polyamine-induced repression of the downstream main ORF. This repression is part of an autoregulatory circuit, and one of its sensors is the AUU codon, which suggests that translation initiation codon identity is likely used for regulation in eukaryotes.

ornithine decarboxylase (ODC) catalyzes the synthesis of putrescine from ornithine. This is the first and rate-limiting reaction in the biosynthesis of polyamines in cells (1). Because of the narrow concentration range of polyamines needed for their multiple roles, ODC expression is tightly controlled. Even modest reductions in mammalian ODC activity can lead to marked resistance to tumor development (1, 2). ODC is turned on rapidly, and this, in part, is because of it being targeted to the 26S proteasome without ubiquitination by the protein antizyme, its key regulator (3, 4). Antizyme itself is negatively regulated by antizyme inhibitor, a homolog of ODC with a higher affinity for antizyme that has lost the ability to decarboxylate ornithine (5). In addition to its posttranslational regulation, ODC is under transcriptional and translational layers of control.

Appreciation of the role and prevalence of upstream open reading frames (uORFs; short coding sequences 5′ of the main coding sequence) in expression of eukaryotic genes is increasing (6–10). Although in most cases the function of the uORF does not depend on the sequence of the encoded peptide, several sequence-dependent uORFs have been studied in depth (for review, see ref. 6). As would be expected from the scanning model of eukaryotic translation initiation, most uORFs have an inhibitory effect on the expression of the main ORF. After translating a uORF, when the 40S subunit of any dissociating ribosome remains on the mRNA and resumes scanning, it may have a depleted repertoire of associated initiation factors. Further, translation of some uORFs, especially those whose sequences are crucial, leads to ribosome stalling, mostly just before termination, with consequent queuing of any trailing ribosomes. Such ribosome stalling can also inhibit expression of the main ORF indirectly via nonsense-mediated mRNA decay (NMD) effects (11).

In vitro and in vivo experiments have demonstrated that in mammals all codons differing by a single nucleotide from AUG can be used as initiation codons albeit at significantly reduced levels with CUG and ACG being the most, and AGG and AAG the least, efficient (12, 13). In Saccharomyces cerevisiae, codons differing by one nucleotide from AUG can also initiate translation but even less efficiently, and the order of initiation codon efficiency is different from that in mammals. For example, AUU is more efficient in S. cerevisiae than either CUG or ACG (14). Initiation at both AUG and non-AUG codons in eukaryotes is enhanced or inhibited by the identity of neighboring nucleotides, with the optimal being the sequence GCC(A/G)CaugG (15) often referred to as the Kozak consensus.

Results

Conserved Upstream Coding Region in Homologs of ODC. Here, we report that many 5′ leaders of eukaryotic ODC homologs contain an upstream conserved coding region (uCC). A uCC is present in all 49 vertebrate antizyme inhibitor gene orthologs for which sequence is available. In these genes, it is ~50 codons long and strikingly lacks an in-frame AUG codon [Fig. 1 A and supporting information (SI) Fig. S1 A]. Unlike the antizyme coding sequence, initiation of translation of these uCCs does not appear to start in a different frame followed by a translational frameshift (SI Text and Fig. S2). Instead, there is a conserved in-frame AUU near the 5′ end of the uCCs, and the possibility that it might serve as an initiation codon for the uCCs is explored below.

The 5′ leaders of all available eukaryotic ODC sequences were examined for the presence of a uCC homologous to that in vertebrate antizyme inhibitor mRNAs. At least 70 additional uCC sequences from species belonging to 9 animal phyla were identified in this search (Fig. 1 A ii–v and Fig. S1 B–G). With rare exceptions, they lack appropriate in-frame initiation AUG codons and in some cases any AUG in their 5′ leaders (Figs. S3 and S4). The main exceptions are the orthologs of ODC in nonmammalian/avian vertebrates and orthologs of ODC in tetrapods (the latter ortholog is a homolog of ODC resulting from duplication in the tetrapod lineage), which have appropriately positioned AUGs and no AUU for initiation of the uCC. This AUG, however, is present in a poor Kozak context. Mammalian ODC orthologs lack the uCC. Like the antizyme inhibitor mRNAs above, most of the invertebrate uCCs lacking an in-frame AUG have a conserved AUU codon in a good Kozak context occupying a similar position within the uCC. Where a putative AUG initiation codon could be unambiguously identified, there is a strong bias for the nucleotides in positions +5 and +6 (where the A of the AUU codon is defined as +1) to be C and G, respectively (Fig. S5). In several invertebrates, it is unclear which initiation codon is used for the expression of the uCC. In some cases, this is undoubtedly caused by an incomplete 5′ sequence, and in others insufficient close relatives are available...

*BioSciences Institute, University College Cork, Cork, Ireland; and †Department of Human Genetics, University of Utah, Salt Lake City, UT 84112-5330

Edited by Joan A. Steitz, Yale University, New Haven, CT, and approved May 5, 2008 (received for review February 18, 2008)

© 2008 by The National Academy of Sciences of the USA

Additional author information is available in the Supporting Information online.
acid, proline has major effects on translation termination, and in uORF-M, but noteworthy given the known translational two terminal prolines, positioned slightly differently from those Bv i conserved (Fig. 1). This feature can be exploited. [An example with tandem prolines is in the human cytomegalovirus UL4 uORF (17). An example with a single proline is in the decoding of the Escherichia coli tryptophanase operon (18).]

Fig. 1. WebLogo representation of the amino acid conservation of different uCCs found in 5' leaders of eukaryotic ODC homologs: in animals (A), in Pezizomycotina (B), in Basidiomycota (C), and in Zygomycota (D). The middle, poorly conserved, portion of Pezizomycotina uCC is not shown and instead is represented by three dots. Each line represents alignment of a subset of sequences: from vertebrate antizyme inhibitors (i); from ODC homologs in invertebrate chordates, hemichordates, echinoderms and molluscs (ii); from ODC homologs in velvet worms and annelids (iii); from an ODC-like homolog in nematodes (iv); from ODC orthologs in nonmammalian, nonavian vertebrates (v); from Pezizomycotina uCCs likely initiated by AUU (vi); from Pezizomycotina uCCs likely initiated by UUG (vii) and from Pezizomycotina uCCs likely initiated by ACU (viii). The number of sequences used to compile each line is indicated in parentheses on its right. In each case, the alignment is shown starting from the putative initiation codon. Methionine is not shown as the first amino acid except for ODC orthologs in nonmammalian, nonavian vertebrates where uORF-M is likely initiated with AUG; however, all uORF-Ms are expected to start with methionine. The two adjacent prolines discussed in the text are indicated by arrows. (A complete alignment, including several subsets not shown here, is available in Fig. S1.)

For comparison. It cannot be assumed that all of them would be initiated by an AUU codon. Among the different animal uCCs, similarity is highest near the C terminus of the putative peptide, notably, in at least eight animal phyla, the sequence NAEPP-WDP, or close variants. In a subset of seven different phyla, the uCCs end with the dipeptide PS (Fig. 1A). In plant mRNAs encoding S-adenosylmethionine decarboxylase, which is another key enzyme in the biosynthesis of polyamines, a short, highly conserved, regulatory, AUG-initiated uORF also ends with the dipeptide PS (16). Data supporting the assertion that the observed conservation near the C terminus of the antizyme inhibitor uCC is at the amino acid level and not at the nucleotide level is provided in SI Text and Fig. S6. Additional amino acid positions closer to the N terminus of the animal uCC seem conserved in a lineage-specific manner. For the remainder of this work, the uCC in animals will be referred to as uORF-M.

The 5' leaders of nonmetazoan ODC homologs were analyzed separately for uCCs with different sequences because uORF-M could not be identified in them. Three additional orthologous uCC groups were found each in a different fungal phylum, Zygomycota, Basidiomycota, and the subphylum Pezizomycotina of Ascomycota. All 36 Pezizomycotina species with available ODC sequences have a uCC. Although none of them has an in-frame AUG, 17 have a conserved AUU codon, 10 have a conserved UUG, and 5 have a conserved ACG, all flanked with a Kozak context, near the 5' end of the uCC (Fig. S3B). On the basis of these initiation codon assignments, the Pezizomycotina uORFs are between 77 and 100 codons. Again, the most highly conserved region is near the C terminus of the putative peptide with six amino acid positions, of the last ~35, completely conserved (Fig. 1B vi–viii and Fig. S1 H–J). Prominent are the two terminal prolines, positioned slightly differently from those in uORF-M, but noteworthy given the known translational effects of proline. When present as the last or penultimate amino acid, proline has major effects on translation termination, and
C-terminal peptide of uORF-M, the last 10 sense codons of uORF-M were put out-of-frame by a 1-nt deletion followed by a downstream 1-nt insertion. Unlike the AUU to UUU mutant, there is no further derepression in polyamine-depleted cells, but like it, polyamine-induced repression of expression of the main coding sequence is essentially lost, and to the same extent (Fig. 2). Combining the out-of-frame mutation with the AUU to AUG mutant results in levels of repression of expression of the main coding sequence comparable with that of the AUU to AUG alone. This indicates that although the last 10 aa of uORF-M are important for mediating polyamine-induced repression, they are redundant for repression when uORF-M is initiated and translated efficiently. This result is not unexpected because translation of ORFs longer than 40 codons generally precludes ribosomal reinitiation downstream (19, 20).

Because at least some uORF-induced repression of main ORF expression can be caused by activating an mRNA NMD response (11), the steady-state mRNA levels of the wild type and three mutant constructs were investigated by Northern blot analysis. The results are shown in Fig. S8. Although AUU to AUG mutation results in lowering the steady-state levels of mRNA by half, indicative of a possible NMD response, this accounts for only a fraction of the observed drop in protein expression of the main ORF. Even more significantly, none of the observed polyamine-induced reduction of luciferase expression in the wild-type construct can be accounted for by corresponding reduction of mRNA steady-state levels. These results are consistent with the conclusion that NMD plays little or no role in mediating the uORF-M repression of the downstream ORF and strongly suggest that the observed repression is almost entirely at the translational level.

Translation initiation on the AUU codon was further investigated by fusing uORF-M, including the preceding 5′-UTR sequence of mouse antizyme inhibitor mRNA, to a firefly luciferase-encoding sequence in conjunction with an AUU to AUG and also AUU to UUU mutants in the same context. The results of these experiments (Fig. 3A) show that in both polyamine-depleted and supplemented cells there is no firefly activity above background (negative control) when AUU is mutated to UUU, whereas firefly activity in both the wild type and AUU to AUG constructs is clearly evident. These experiments, together with the anti-firefly Western blotting using the same cell lysates (Fig. 3B), are consistent with translation initiation of uORF-M starting at the AUU codon. Furthermore, these data also indicate that in polyamine-depleted cells, initiation on the AUU, at least in this context, is 18% as efficient as initiation on AUG in the same context. However, in polyamine-supplemented cells, initiation on the AUU codon is 54% as efficient as
Fig. 3. Analysis of the initiation of uORF-M at AUU by fusing it N-terminally to firefly luciferase. (A) Analysis of firefly fusions in mammalian HEK-293 cells by using dual luciferase assay. (Left) Schematic representation of the mutations analyzed, analogous to that shown in Fig. 2. (Right) Firefly to Renilla luciferase ratio either in cells depleted of, or supplemented with, spermidine as described in Fig. 2. Fold stimulation upon spermidine addition is shown above the columns. (B) Western blot of lysates from transfected cells as described in A by using an antibody against firefly luciferase. The full-length fusion uORF-M-firefly-luciferase protein product is indicated by an arrow. The bands below it represent initiation on downstream codons.

Discussion

The revelation of important uORFs, previously unsuspected because of their initiation with noncanonical start codons, especially AUU, expands our perspective of translation versatility. Until now, AUU has not been known to serve as a translation initiator for eukaryotic chromosomal gene expression, although it is used in the decoding of one animal and two plant viruses (21–23). However, three E. coli genes initiate with AUU. The best known is the gene encoding translation initiation factor 3 (24). This initiation factor discriminates against starting at noncanonical initiation codons. When its level decreases, the chance of initiation to replenish any deficit increases resulting in autoregulatory initiation.

Initiation of eukaryotic translation on codons differing at a single position from AUG is inefficient and rare (25). The nucleotides surrounding the AUU codon in antizyme inhibitor uCC comply with the standard context for efficient initiation (15) although not to the same extent as those flanking some previously identified non-AUG-initiating codons (26). However, with the parainfluenza virus type 1 P gene GUG initiator, additional nucleotides at positions +5 and +6 also play a critical role (26). For the antizyme inhibitor uCC, the triplet after the AUU is usually CCG (Fig. S5), which differs from the preference for nucleotides +5 and +6 found in the parainfluenza P gene.

In eukaryotes, no single initiation factor performs the functions of bacterial initiation factor 3. However, eukaryotic initiation factor 1 (eIF1) is pertinent. Dissociation of eIF1 from the ternary initiation complex (eIF2, GTP, and Met-tRNA<sup>Met</sup>) is a key step in initiation codon selection (27–29). Specifically, mutants in eIF1 that reduce its affinity to the ternary complex increase initiation at non-AUG codons. This phenotype is suppressed by overexpressing the mutant eIF1 protein presumably through mass action (28). eIF1 interacts with several partners associated with, or part of, the ternary complex eIF2, eIF3, and eIF5 (30). Any condition that reduces the binding of eIF1 to the ternary complex theoretically could lead to altered selectivity for the initiation codon. Mutations in eIF2 and eIF5 that lead to increased hydrolysis of GTP in the ternary complex increase initiation at non-AUG codons (31). Therefore, anything that alters the ratio of eIF2-GDP-Pi and eIF2-GTP in the ternary complex could also affect the fidelity of initiation. How polyamines might affect the activity of any of these factors leading to more efficient initiation on the AUU codon of uORF-M or whether there are other non-polyamine-related conditions that produce the same effect is unknown. Although eIF5A has the essential polyamine-derived modification, hypusine, it is not established that it is involved in the initiation step of protein synthesis (32).

uORFs with canonical start codons are important in mRNAs whose main products are involved in controlling cell growth such as receptors, growth factors, and other protooncogenes. Multiple factors, including arginine and polyamines, influence their expression, and they in turn can have positive or negative effects on downstream ORF translation (6).

With vertebrate S-adenosylmethionine decarboxylase mRNAs, elevated polyamines stabilize ribosomes paused, because of a special nascent peptide sequence, at the termination codon of their uORF (33). The comparable situation in plant S-adenosylmethionine decarboxylase 5′ leaders, which have evolved independently from vertebrates, may be even more relevant because initiation of the first of two highly conserved uORFs may be influenced by polyamine-dependent leaky scanning (16). Yet another example of translational autoregulation in genes influencing polyamine synthesis is the programmed ribosomal frameshifting required for antizyme synthesis acting as a sensor of polyamine levels (34, 35).

Studies on the 5′ leader of ODC mRNA in Neurospora crassa (36), showed that a region that we now recognize contains the uCC in Pezizomycotina (Fig. 1Bv), causes close to 30-fold repression of expression of ODC protein. Approximately 3-fold of this could be directly attributed to translational repression of the main coding sequence, and this translational repression is partially relieved by polyamine depletion. This strongly suggests that the uCC in Pezizomycotina might have a role analogous to that of uORF-M in antizyme inhibitor.

Even if all eukaryotes share a similar translation initiation mechanism, it appears that different branches of the eukaryotic tree, not only at the kingdom level but perhaps also at the level of phylum, use uORFs differently. For example, uORFs are apparently more common in the 5′-UTRs of mammals than they are in fungi. One consequence is that to date very few, if any, uORFs have been identified that are clearly conserved between mammals and fungi and in fact between vertebrates and invertebrates (37).

Computer and manual searches for uORFs almost invariably assume initiation at AUG codons. Our findings highlight the limitations of such assumptions. Although non-AUG-initiated uORFs have been reported before (38), there has not yet been a systematic attempt to investigate how widespread they might be.

The present discovery of polyamine-induced repression in homologs of ODC genes illustrates how yet another translational control device is used for the critical control of cellular polyamine levels.
Table 1. Mutations in the 5′-UTR of phRL-WT introduced by either standard one-step or two-step PCR

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>phRL-M1*</td>
<td>AZIUTR/S2 + AZIM1/A</td>
<td>AZIM1S5 + AZIUTR/A</td>
</tr>
<tr>
<td>phRL-M2*</td>
<td>AZIUTR/S2 + AZIM2/A</td>
<td>NA†</td>
</tr>
<tr>
<td>phRL-M3*</td>
<td>AZIUTR/S2 + AZIM3/A</td>
<td>AZIM3S5 + AZIUTR/A</td>
</tr>
<tr>
<td>phRL-M4*</td>
<td>AZIUTR/S2 + AZIM4/A</td>
<td>AZIM4S5 + AZIUTR/A</td>
</tr>
<tr>
<td>phRL-M5*</td>
<td>AZIUTR/S2 + AZIM5/A</td>
<td>AZIM5S5 + AZIUTR/A</td>
</tr>
<tr>
<td>phRL-M6*</td>
<td>AZIUTR/S2 + AZIM6/A</td>
<td>AZIM6S5 + AZIUTR/A</td>
</tr>
<tr>
<td>phRL-M4M5†</td>
<td>AZIUTR/S2 + AZIM4/A</td>
<td>AZIM4S5 + AZIUTR/A</td>
</tr>
<tr>
<td>phRL-M3M5†</td>
<td>AZIUTR/S2 + AZIM3/A</td>
<td>AZIM3S5 + AZIUTR/A</td>
</tr>
</tbody>
</table>

*phRL-WT used as template.
†NA, not applicable; one-step PCR.
‡phRL-M6 used as template.
§hRL-M5 used as template.

Methods
Plasmid Construction. For primer sequences, see SI Methods. The 5′-UTR of mouse antizyme inhibitor was fused to Renilla luciferase coding sequence (phRL-WT) by standard two-step PCR (see below and Table 1). Because of the large size (4,286 nt) of the endogenous intron in antizyme inhibitor UORF-M, we replaced it with the 133-nt intron from the 5′-UTR of Renilla luciferase in phRL-CMV (Promega) while still maintaining the same intron/exon boundaries. First, the 5′ and 3′ sections of antizyme inhibitor 5′-UTR were amplified with primers AZIUTR/S2 and INTRA2A (5′ end) and INTRA/S2 and AZIUTR/A (3′ end) by using a mouse antizyme inhibitor cDNA (NM.018745) as template. Next, the 133-nt intron from the 5′-UTR of Renilla luciferase was amplified by using primers INTRA/S and INTRA/A with phRL-CMV as template. Equimolar amounts of each of the three first-step PCR products were mixed and used as template for the second step of the PCR with primers AZIUTR/S2 and AZIUTR/A to generate WT 5′-UTR AZI. WT 5′-UTR AZ1 was purified by ethanol precipitation and digested with HindIII and AvaI and cloned into HindIII/HindIII sections of antizyme inhibitor 5′-UTR were amplified by standard two-step PCR with primers AZIUTR/S2 and AZIM1/A BamiHIl using the templates indicated: phRL-WT for p2Luc-AUG-uORF-M, phRL-M3 for p2Luc-UU-uORF-M, phRL-M4 for p2Luc-AUG-uORF-M, and phRL-M4 for p2Luc-NC-uORF-M. All amplicons were digested with HindIII and BamiHi and cloned into HindIII/BamHI-digested p2Luc (39). All clones were verified by sequencing with primer LucCDSR. NC indicates negative control and was generated by accident when making p2Luc-AUG-uORF-M. A PCR-induced frameshift mutation places an in-frame stop codon at the 3′ end of uORF-M but before the firefly coding sequence.

Cell Culture and Transfections. HEK-293 cells were obtained from American Type Culture Collection and maintained in DMEM supplemented with 10% FBS, 1 mM l-glutamine, and antibiotics. Twenty-four hours before transfections, cells were passaged with medium supplemented with 2.5 mM D- difluoromethylornithine (DFMO; a kind gift from P. Woster via Dr. Michael Howard, University of Utah). Cells were transfected by using Lipofectamine 2000 reagent (Invitrogen), with the one-day protocol in which suspension cells are added directly to the DNA complexes in 96-well plates. In all, 25 ng of DNA and 0.2 μl of Lipofectamine 2000 per well in 25 μl of Opti-MEM/DFMO (Invitrogen) were incubated and plated in opaque 96-well half-area plates (Costar). Cells were trypsin-treated, washed, and added at a concentration of 4 × 10⁴ cells per well in 50 μl of Opti-MEM/DFMO. Transfected cells were incubated overnight at 37°C in 5% CO₂, then 75 μl of DMEM/10% FBS with 1 mM final concentration aminoguanosine (Sigma), 2.5 mM DFMO ± 1 mM final concentration spermidine (Sigma) as indicated were added to each well, and the plates were incubated for an additional 48 h. For the effect of DFMO treatment and polyamine supplementation on the intracellular concentrations of polyanymes in HEK-293 cell see Fig. 6 in ref. 40.

Dual Luciferase Assay. Luciferase activities were determined by using the Dual Luciferase Stop and Glo Reporter assay system (Promega). Relative light units were measured on a Verita microplate luminometer fitted with two injectors (Turner BioSystems). Transfected cells were washed once with 1 × PBS and then lysed in 12.5 μl of passive lysis buffer (PLB; Promega), and light emission was recorded after injection of 25 μl of either either Renilla or firefly luciferase substrate. For phRL-WT and its mutants, Renilla luciferase activity was calculated relative to the activity of an internal control plasmid (cotransfected at 1/10 the concentration of the test plasmid) expressing firefly luciferase driven by the weak ubiquitin promoter (pUB-Luc; a kind gift from Mark Tagney, Cork Cancer Research Centre, Ireland). For each construct, all data points were averaged and the standard deviation calculated. Data shown represent the means ± SD from three independent experiments each done in triplicate. For siORF-M firefly fusions, firefly luciferase activity was calculated relative to the activity of an internal control plasmid (phRL-CMV; cotransfected at 1/300 the concentration of the test plasmid) expressing Renilla luciferase. Data points and the corresponding standard deviations represent results from a single experiment done in triplicate; however, the results are representative of three additional independent experiments (data not shown).

Western Analysis. Transfected cells (HEK-293 treated with DFMO and spermidine as above) were washed once with 1 × PBS and then lysed in 20 μl of PLB; 7.5 μl was removed for Western blotting, and for the remaining 12.5 μl light emission was measured as described above as a control. Cell debris was removed from cell lysates by centrifugation at 15,000 × g at 4°C for 15 min, and samples were then denatured by boiling in 5 × SDS/PAGE sample buffer for 5 min. Proteins were resolved by 12% SDS/PAGE and transferred to nitrocellulose membranes (Protran), which were incubated at 4°C overnight with a 1:1,000 dilution of goat anti-firefly luciferase (Promega). The next day, membranes were incubated at room temperature for 1 h with 1:10,000 IRDye 800CW-conjugated goat secondary antibodies, and immunoreactive bands were detected on the membranes by using a Li-COR Odyssey Infrared Imaging Scanner (LI-COR Biosciences).

Acknowledgments. We thank Drs. R. B. Weiss, C. W. Anderson, S. Matsu-fujii, and R. F. Gesteland for their reading the manuscript. This work was supported by an award from Science Foundation Ireland. J.F.A. was supported by National Institutes of Health Grant R01 GM79523.


Supporting Information

Ivanov et al. 10.1073/pnas.0801590105

Methods

Translation would be expected to start on the downstream ACG consensus in Zygomycota is different from other eukaryotes, lacks an appropriate Kozak consensus; and unless the Kozak main candidates being adjacent AUC and ACG triplets. The first putative initiation codon is harder to identify, with the two candidates being adjacent AUC and ACG triplets. The first putative Kozak consensus is the result of a sequencing error, the second shortest, present in 16 different genes, is only 13 nucleotides long; and once again, using the rules gathered from previously identified examples, no putative -1 frameshift site can be identified. Furthermore, in Macaca mulatta AZI, a 2-nucleotide insertion in the uORF1 means that it is in the +1 not -1 frame relative to uCC. Requirements for +1 and -1 ribosomal frameshifting are distinct, and it seems most improbable that even if uORF1 contained a frameshift site it would be able to facilitate both.

Analysis of the synonymous vs. nonsynonymous substitutions in the highly conserved 3' end of antizyme inhibitor uORF-M is illustrated in Fig. S6. In all, 101 substitutions are observed in the selected region. Seventy-one of them are synonymous, and only 30 are nonsynonymous, strongly supporting the suggestion that the conservation in that region is predominantly, if not exclusively, on the amino acid level.

In Basidiomycota uCC, the two conserved prolines are preceded by a negatively charged residue (glutamate) unlike the positively charged residue (arginine or lysine) preceding the two prolines in Pezizomycotina.

Analysis of the synonymous vs. nonsynonymous substitutions in the highly conserved 3' end of antizyme inhibitor uORF-M is illustrated in Fig. S6. In all, 101 substitutions are observed in the selected region. Seventy-one of them are synonymous, and only 30 are nonsynonymous, strongly supporting the suggestion that the conservation in that region is predominantly, if not exclusively, on the amino acid level.

In Basidiomycota uCC, the two conserved prolines are preceded by a negatively charged residue (glutamate) unlike the positively charged residue (arginine or lysine) preceding the two prolines in Pezizomycotina.

An uCC in Zygomycota belongs to three different species (two of the homologs belong to the same species, Rhizopus oryzae). The putative initiation codon is harder to identify, with the two main candidates being adjacent AUC and ACG triplets. The first lacks an appropriate Kozak consensus; and unless the Kozak consensus in Zygomycota is different from other eukaryotes, translation would be expected to start on the downstream ACG codon, which is the one we assign as the putative start codon.

Methods

Sequence Assembly and Analysis. All sequences were obtained from GenBank using the BLAST algorithm. In most cases, the sequences are of expressed sequence tags (ESTs). In a minority of cases, the sequences also come from Whole Genome Shotgun projects. In the latter case, intro/exon junctions were determined manually, and only the (likely) spliced sequence was further analyzed. Clustering was used to determine whether a particular sequence belonged to the antizyme inhibitor, ornithine decarboxylase or the ODCp (also known as antizyme inhibitor 2) orthologous groups. Clustering was performed with the ClustalX algorithm.

Primers

AZIUTr/S2, CCAAGAGCTCTCTCGGCCCGGT-GTTTCCG.

INTR/A2, CCTGTCTTGAACCTTGTACCTTAGTAGGT-GAGATTGCGCCCAAGAG.

AZIUTR/A, CACCTCTGCTTTCATCTCAGGCTATTTCCACAAAGGC.

INRT/S, CTCTTGGGGCCGCTTATCTCAGGTAAGTACCAAGTTACAGGAC.

INTR/A, TACTCAGGCTAGAAGGTTGCGCCCTGTGAGAGAAAGGAG.

AZIM1/S, CCCCGCAGCTCTAACCCGAGGGCCGC.

AZIM1/A, AGCGCGCTTGGGGTTTAACTGCGGGGC.

AZIM2/A, CACCTCGGTTTTCATCTCAGGCGTTTCCAAAGGCGG.

AZIM3/A, CTTTTTCGCGGCAAATTTTTTAAAGGG.

AZIM4/A, CTTTTTCGCGGCGATAAAAAAG.

AZIM4/A, CTTTTTCGCGGCGATTTTAAAGGG.

AZIM5/S, GAGCGAGTCGATCAGTACGAGCCAC-CTTTGAACCTCTTGATGTAGGG.

AZIM5/A, CCCCTATCATCAAGCTAGGTTCCAAAGGTTGAGTCGAACTGCTC.

AZIUTR/S2, CACCTCTGCTTTCATCTCAGGCTATTTCCACAAAGGC.

AZIUTR/A, CACCTCTGCTTTCATCTCAGGCTATTTCCACAAAGGC.

INRT/S, CTCTTGGGGCCGCTTATCTCAGGTAAGTACCAAGTTACAGGAC.

Northern Blotting. For transfection of cells for RNA extraction, 1.5 µg of each purified BamHI linearized phRL-clone as indicated and 1.5 ng of BamHI linearized pCMV-Luc (Plasmid Factory) were cotransfected by using the calcium phosphate method into HEK-293 cells that had been plated at a density of 3 × 10^6 cells in 5 ml of DMEM, 10% FBS supplemented with 2.5 mM α-difluoromethionel (DFMO) 24 h earlier. Transfected cells were incubated overnight at 37°C in 5% CO2, then 5 ml of DMEM, 10% FBS with 2 mM aminoguanosine (Sigma; 1 mM working concentration), 2.5 mM DFMO ± 2 mM spermidine (1 mM working concentration) as indicated were added to each plate and incubated for an additional 48 h. After 48 h, cells were lysed by trypsin treatment and resuspended in PBS. At this stage, 90% of the cells were removed, pelleted, and frozen for RNA extraction later, and the remaining 10% were pelleted and then lysed in PLB for DLA and Western blotting.

RNA was extracted by using TRIzol reagent (Invitrogen) according to the manufacturer’s instructions, and 10 µg was separated by denaturing formaldehyde gel electrophoresis and transferred to a nylon membrane. Prehybridization and hybridization were carried out at 42°C in 50% formamide, 5 × SSC, 4 × Denhardt’s solution, 0.1% SDS, and salmon sperm DNA (100 µg/ml; Sigma) for 2 and 15 h, respectively. [32P]dCTP-labeled Renilla and GAPDH probes (>1 × 10^6 cpm/ml) were prepared by the random primer method (NEBlot; New England Biolabs). Membranes were washed twice at 42°C in 2 × SSC, 0.1% SDS for 5 min and then twice at 42°C in 0.1 × SSC, 0.1% SDS for 15 min, and exposed to PhosphorImager screens for empirically determined times. ImageQuant (Molecular Dynamics) was used for densitometry analysis.

A

B

Ivanov et al. www.pnas.org/cgi/content/short/0801590105

Fig. S1. Amino acid alignments of all uCCs identified in this study. Most are also shown as WebLogo alignments in Fig. 1 of the main text. The alignments were made using the ClustalX program followed by manual optimization. The highly conserved C-terminal amino acid consensus of animal uCCs is highlighted in gray (and light blue for negatively charged amino acids similar to the consensus). The two highly conserved adjacent prolines are in red letters. All in-frame AUGs in the uCCs are in blue letters. The putative non-AUG codon initiated position, where identified, is highlighted in magenta. The two highly conserved adjacent prolines are in red letters. All in-frame AUGs in the uCCs are in blue letters.
Fig. 51. (Continued)
### H

**vi**  
Hypocrea jecorina  
Trichoderma reesei  
Trichoderma atroviride  
Trichoderma virens  
Cordyceps bassiana  
Gibberella moniliformis  
Fusarium oxysporum  
Gibberella zeae  
Verticillium dahliae  
Corynesus heterothallicus  
Chaetomium globosum  
Neurospora crassa  
Magnaporthe grisea  
Botrytis fuckeliana  
Oculimacula yallundae  
Sclerotinia sclerotiorum  
Ivanov et al.  
www.pnas.org/cgi/content/short/0801590105

---

### I

**vii**  
Trichophyton rubrum  
Aspergillus fumigatus  
Neosartorya fischeri  
Aspergillus clavatus  
Aspergillus terreus  
Aspergillus flavus  
Aspergillus niger  
Aspergillus nidulans  
Talaromyces stipitatus  
Penicillium marneffei  
Ivanov et al.  
www.pnas.org/cgi/content/short/0801590105

---

### J

**viii**  
Faraccidioideae brasiliensis  
Ajellomyces capsulatus  
Coccidioideae posadassii  
Uncinocarpus reessii  
Ivanov et al.  
www.pnas.org/cgi/content/short/0801590105

---

### K

Rhizopus oryzae_1  
Rhizopus oryzae_2  
Nucor circinellidis  
Phycomyces blakesleeanus  
Ivanov et al.  
www.pnas.org/cgi/content/short/0801590105

---

### L

Coprinopsis cinerea  
Panhornechae chrysophorum  
Beleoma cylindrosorum  
Schizopyllum commune  
Antrodia cinnamomea  
Ivanov et al.  
www.pnas.org/cgi/content/short/0801590105

---

Fig. 51. (Continued)
<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
<th>ORF Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petromyzon marinus Azi</td>
<td>CCGGCGG</td>
<td>CC</td>
</tr>
<tr>
<td>Osmerus mordax Azi2</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Squalus acanthias Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Leucoraja erinacea Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Ovis aries Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Sus scrofa Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Spermophilus lateralis Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Meleagris gallopavo Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Gallus gallus Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Tetraodon nigroviridis Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Oryzias latipes Azi1</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Gasterosteus aculeatus Azi1</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Pimephales promelas Azi1</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Gasterosteus aculeatus Azi2</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Leucoraja erinacea Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Ovis aries Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Sus scrofa Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Spermophilus lateralis Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Meleagris gallopavo Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Gallus gallus Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Tetraodon nigroviridis Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Oryzias latipes Azi1</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Gasterosteus aculeatus Azi1</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Leucoraja erinacea Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Ovis aries Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Sus scrofa Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Spermophilus lateralis Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Meleagris gallopavo Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Gallus gallus Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Tetraodon nigroviridis Azi</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Oryzias latipes Azi1</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
<tr>
<td>Gasterosteus aculeatus Azi1</td>
<td>CGGACG</td>
<td>CC</td>
</tr>
</tbody>
</table>

**Fig. S2.** Nucleotide sequence alignment of the region 5' of the main ORF of vertebrate antizyme inhibitor orthologs. The uCC ORFs bracketed by the conserved in-frame AUU (shaded in green) and its downstream in-frame stop codon (shaded in red) are shown shaded in light blue. The AUG codons of the three main conventional uORFs are also highlighted in green and all but those initiating the main ORF are also underlined. The downstream in-frame stop codons for each uORF are highlighted in red. The upstream in-frame stop codons of uCC closest to the conserved AUU are highlighted in magenta. The alignments were made using the ClustalX program followed by manual optimization.

Ivanov et al. www.pnas.org/cgi/content/short/0801590105 5 of 19
Fig. S3.  Nucleotide sequence alignment of the regions corresponding to the uC in invertebrate metazoans and fungi. The alignments were made using the ClustalX program followed by manual optimization. The uC ORFs are bracketed by in-frame stop codons (where available) shaded in red. The likely initiation codons are highlighted in green. Surrounding nucleotides that constitute a Kozak consensus are highlighted in magenta. Where there are in-frame AUG codons available, they are shown in red letters. (A) ODC homologs from invertebrate animals. In Schistosoma mansoni and Crassostrea gigas, there appears to be conserved AUU codons with consensus surrounding context that precede the upstream in-frame stop codon. In both cases the sequence is derived from a single EST, and the likelihood of sequence error is high. In each case, the likely problem nucleotides are highlighted separately in light blue. (B) ODC homologs from Pezizomycotina. (C) ODC homologs from Basidiomycota. (D) ODC homologs from Zygomycota. Phyla species coloration scheme for animal ODC homologs is same as shown in Fig. S1.
"Nucleotide (cDNA) sequences of the region 5 of the main ORF of nonvertebrate ODC homologs with identified uCCs. The uCC ORFs bracketed by TGTGTGCCTGTGTCCTGCTCGTAGCCCAGAGGAAGTGCAGTTTGTTGGTGACTCATCACAA..."

**Apis mellifera**

>**Bombyx mori** ODC (Arthropoda)

>**Nasonia vitripennis**

>**Myzus persicae** Arthropoda

>**UNKNOWN APHID** Arthropoda

**Anopheles gambiae** Arthropoda (mosquito)

**Artemia salina** Arthropoda

**Drosophila melanogaster** Arthropoda (insect)

**Bacillus subtilis** Bacteria

**Picea sitchensis** actually contamination from moth or butterfly

Fig. S4. Nucleotide (cDNA) sequences of the region 5 of the main ORF of nonvertebrate ODC homologs with identified uCCs. The uCC ORFs bracketed by TGTGTGCCTGTGTCCTGCTCGTAGCCCAGAGGAAGTGCAGTTTGTTGGTGACTCATCACAA...
> Tribolium castaneum Arthropoda

> Leptinotarsa decemlineata Arthropoda (Colorado potato beetle)

> Daphnia magna Arthropoda

> Homarus americanus Arthropoda (lobster)

> Carcinus maenas Arthropoda (green crab) isoform #1

> Tribolium castaneum Arthropoda

> Platynereis dumerilii Annelida

> Rhipicephalus appendiculatus Arthropoda (Brown ear tick)

> Ixodes scapularis Arthropoda

> Leptinotarsa decemlineata Arthropoda (Colorado potato beetle)

> Amblyomma americanum Arthropoda ('lone star' tick)

> Amblyomma variegatum (tick)

> Ornithodoros porcinus porcinus Arthropoda

> Amblyomma americanum Arthropoda

> Platynereis dumerilii Annelida

> > Epiperipatus sp. Tb 2001 Onychophora (velvet worms)

> Fig. S4. (Continued)
Ivanov et al. [www.pnas.org/cgi/content/short/0801590105]

> Saccoglossus kowalevskii (Hemichordata) Acorn worm

GACGTCGCGTTTGGCTGTCCCTACTTGTACAACAAAGATT

> Strongylocentrotus purpuratus Echinodermata

CTCAGCAGCTTGTGACATTTTCTTAAATCCCTGGTCAAGTTCCCCTAGTGTTTTGGTTACTTTTATG

> Paracentrotus lividus Echinodermata (common urchin)

TATTCCTCCTTCTGACTAGCACTTGTTAATTGTATCCCTTTGAAATTAGCCAAAATT

> Chiona intestinalis ODC Chordata

CTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Branchiostoma floridae (Florida lancelet) Chordata

TTTACTACCATG

> Crassostrea gigas Mollusca (Pacific oyster)

ACGTCCATACAAAATAAAGAAGCTCAGTCTGAGGCTTCTGTTGAGGGCAGTAGCTCGACTGTCTTCTTCAGCCAAAGTCCAGAGGCCGCTTGTGGAATAAGGAGGA

> Mytilus californianus Mollusca (mussel)

CTCCGGAATCTTCTTAGACGACGGCCATATCCTAAATTTTCTTACTCAGTTGTCGACAAACTACCAAAATT

> Strongylocentrotus purpuratus Echinodermata

CCGTCGTTTAAGAATACCCGTGTTAGTTCGACTGAACTGTCAGTCCAAGGTGGACATTGTAACAAGTACTGGAGCACTGAGCCAGGTGGAGCTCTCTACCGAGACCAGCTAGACGTT

> Patiria pectinifera (Echinodermata) (starfish)

TCCAAGTCCACTTTGTCTGGGAAAGCAACTTTTTCTTCTCTGGCCAGCAAGTACATTTCTCGACAGGAGGGTCTTTGCCTTTACAAACCGCAGCTCGACTTCAACTTCGATCCACCTT

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Angiostrongylus cantonensis Nematoda

GACGTCGCGTTTGGCTGTCCCTACTTGTACAACAAAGATT

> Ciona savignyi Chordate

AACTTTTCCAAACATTTTCACCCAAACAAAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Halocynthia roretzi Chordata

CGTGTTTTACGTTTCTAACAACAAACAAGTGCAATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata

CTTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Patiria pectinifera (Echinodermata) (starfish)

TCCAAGTCCACTTTGTCTGGGAAAGCAACTTTTTCTTCTCTGGCCAGCAAGTACATTTCTCGACAGGAGGGTCTTTGCCTTTACAAACCGCAGCTCGACTTCAACTTCGATCCACCTT

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common urchin)

TTCTTCAACGACAGCAAGCCCGAAAAGAGGACGGAGAGAACCGGCGTTCCCGAGCAGAATTCATG

> Ciona intestinalis ODC Chordata

GCTGGTGTCCCAGATTTTCTTCTTTGCTGTGTTGTTGCGAATATCATAGTCTCCTGTGTCCTAGCTCCTATCCTAGTCTATATTAACCTTATCAAGAATG

> Strongylocentrotus purpuratus Echinodermata (common  Fig. S4. (Continued)
Ivanov et al. >Bursaphelenchus mucronatus Nematoda
Schistosoma mansoni Platyhelminthes
GGACAACAGACTGTAGATTGGCTATCTGTGTTTGTCGCTCA
CGGTTTAAATATTGCCGGTGCTTTTTATTATTTAAAATTCCCTCTTTAAAGAAAACTACACCTCAAGTGCGGTCTTCCCCAGAGCTTGTAAAAGTTTTCATTGCTGATATTAGCTCCT
TGGACAATTGCAATTAAATAATAACGCAGAGCCTCCTTGGGATCCTGATCCAAAATATTTATAAATCTACTTTTCTGAGTTATTAAACCATTTCAGAAAGTCAAAAACATTTAGGCTC
TTGCAATCATTAAATAAGATTTTGACTCTAAAAGTTTAATATTATTTTTTTTAATTATTATTAAGTCTAATACGCTATACGGGATAGTGCGTATACGGACTTGTTGACAGATATCTGTCCCTAAAAAGA
Strongyloides stercoralis Nematoda
ACGAATTTCTAACAACCGACACATCCCACCATATTTTCCAGAAGTTCAAACTATTTTCCATTAAAGTAAAAAAATCCAAATAACTAATG
TTACCCCCTTCCCATTATTACTGTAGCTGACAGACGTCTCACCCTCCATCTAAGTCTTCACTACAACTGTCTTCTTTCTTCCCAGGAACGAGCATCGAGTTGGTACAAATTCCGCCTC
ACTGTCCCCAATTCGTCCAAGATT
TGAATCTGTGCTTATCAATTGCACGGATTTGACCTCCAATCCTTGTGTGTAAACACTCATTCAATACTTCCTGTTCCGTGCACTCCAAGGGAGAGAGTGTGCATG
Caenorhabditis briggsae Nematoda
CGTGTTTCCAACAACCGTCACATCCCGCCATAACTTGGCACATTCAATTTCCCCATTAAACAACAAATCGATATAAAATAACTGATG
TCCTGGTCTTCCGTCAATTATCAAAATTTCGACGGTGACCTCATCTCATTCCTCGTCTCCAGCCCGTCTCTGTCCTCCGGACGACGTCTCGTGGATCTCGCGTGAAGAGCTCGCATCG
CTGCCGACTGACCGACTCTCTCTCTTCACATCTATAACCTTTTCTTCATTTTCTATTCAATCGCCCAGTAGCGTGCGTACGAGCCGGTAATCAATAGACCGCCTCGTTTCGCCCGTTT
GTGCATG
GTTTAATTACCCAAGTTTGAGAATCCGTGCCCCGTCGTGCTCGGNATTTGAACTCGATACCTTGCGTGTGAACTACTAACACAAATCGTCTTGTTCCGTGCACTCCAAGGGAGAGAGT
GTTCAACCAAACACTTCAAAGCTTCATTCGCGAGTTGTCTTCCTGTTTATCGTATTCTTCGAGTCCGTAATCGCTTCCGGGATAGGGCGTTTATG
Globodera rostochiensis Nematoda
CCCTATCTAGCCTTTCTTCTACCTTTTACA
GAGCCGCAATTAAGAGCTAGCCAGTTACAGCTCGACAAGAACGCCGAGCCTCCGTGGGATCCAGACCCAAACGTCTGCTGAGGTCGACTCGTCCGCAGCCGTCGTCTTCGATCTCGTT
CTTGACTGTCCCTACATAAACGTCGAAGATT
GAAGCCTCGAAACACTCTCAACTCTTCAGTGAATCTTTTGTTGAACTTACGTTCAACGACCCACACCGTTATTGTCCGTCTTGCACGCTCCGAGGGACAGAGCGGTATG
Haemonchus contortus Nematoda
CGACCGCCCTAATTTGTTTGTCCTTTCATTCATTGGCATTGGCAACTCGGCATTATTTGAATTCCACCAACTACAACGATCAGCGCTCTTTTCTTCGGCAGAACATG
ACTCCGTCTTTCGTCCTCTTCCTTCGTCACTGTGCGTCCTCATCATTTGAAGGTCAACCAGGAGCGTCTCTCCCACTTTCGGTTCCGGAAGTTGGCCGCCTTCGTCCGTCGTCTGAAG
TTGTTAGTGTCCCATCCATTGTCTCGTCCCTTTGTTTCTGCCGACTTCGTCTTTGAAGCGGCGTTCTTATTCAACACGCCAAAATTCGTCTCTTTTCCAACCCCAAAGGCCTTTTCCT
GGCCCCGAACCCGAGCCTACAACACTTGGACGTGAACGCGGAGCCTCCTTGGGACCCTGATCCGAACGCAATTACCTGACATTCGTTCCCGTACATTGGTGAAATTGTCCTCCATTA
CTCAAAACCAATCCAAAATT
GTTCAACCAAACACTTCAAAGCTTCATTCGCGAGTTGTCTTCCTGTTTATCGTATTCTTCGAGTCCGTAATCGCTTCCGGGATAGGGCGTTTATG
>Fig. S4. (Continued)
**Dugesia japonica** Platyhelminthes (planaria)

Fig. S4. (Continued)
Fig. S5. Alignment of 15 nucleotides surrounding and including the putative non-AUG initiation codon of uCCs in various groups of organisms. Only sequences from genes where the initiation codon is less ambiguous are shown. In each case the putative initiation codon is highlighted in green. A summary of each block of alignments is shown in a WebLogo format below. The species name coloration scheme for invertebrate animals is the same as in Fig. S1. (A) Antizyme inhibitor orthologs from vertebrates. (B) ODC homologs from invertebrate animals. (C) ODC homologs from Pezizomycotina with AUU as initiation codon for the uCC. (D) ODC homologs from Pezizomycotina with UUG as initiation codon for the uCC. (E) ODC homologs from Pezizomycotina with ACG as initiation codon for the uCC. (F) ODC homologs from Basidiomycota. (G) ODC homologs from Zygomyota. See Fig. S3 for additional details.
**B**

- *Saccoglossus kowalevskii*: GCCAAAUUCCACCG
- *Paracentrotus lividus*: GCCAAAUUCCGUGC
- *Strongylocentrotus purpuratus*: GCCAACAUUCCGUGC
- *Patiria pectinifera*: GCCAAAUUCCUCGU
- *Branchiostoma floridae*: UACACCAUUACCCCG
- *Ciona intestinalis ODC*: ACAAUAUUCCGUGC
- *Ciona savignyi*: ACCAAAUUCCGUGC
- *Crassostrea gigas*: GGCAAAAUUCCGACC
- *Crassostrea virginica*: GGCAAGAUUCCGACC
- *Lymnaea stagnalis*: UUAAAAAUUACGUCC
- *Platynereis dumerili*: GCAACAAUUCCGUGC
- *Capitella sp. #1*: UCCACAUUCCGCCU
- *Heleobdella robusta*: UGGAAAATUCGCCC

**C)**

- *Corynascus heterothallicus*: ACGAAGAUUUCAGCC
- *Chaetomium globosum*: ACGAAGAUUUCAGCC
- *Neurospora crassa*: ACCAAGAUUUCAGCC
- *Magnaporthe grisea*: ACCAAGAUUUCAGCC
- *Hypocreopsis jecoria*: ACCAAGAUUUCAGCC
- *Trichoderma reesei*: ACCAAGAUUUCAGCC

*Fig. S5. (Continued)*
Trichoderma atroviride  
Trichoderma virens  
Gibberella moniliformis  
Gibberella zeae  
Fusarium oxysporum  
Cordyceps bassiana  
Verticillium dahliae  
Cryphonectria parasitica  
Botryotinia fuckeliana  
Sclerotinia sclerotiorum  
Oculimacula yallundae

D
Aspergillus fumigatus  
Neosartorya fischeri  
Aspergillus clavatus  
Aspergillus terreus  
Aspergillus niger  
Aspergillus flavus  
Talaromyces stipitatus  
Penicillium marneffei  
Aspergillus nidulans  
Trichophyton rubrum

E
Paracoccidioides brasiliensis  
Ajellomyces capsulatus  
Coccidioides posadasii  
Coccidioides immitis  
Uncinocarpus reesii

F
Basidiomycota  
Coprinopsis cinerea  
Phanerochaete chrysosporium  
Hebeloma cylindrosporum  
Schizophyllum commune  
Antrodia cinnamomea

Fig. S5.  (Continued)
G
Zygomycota
Rhizopus oryzae 1                  CCCAUCA CGGCUACA
Rhizopus oryzae 2                  CCCAUCA CGGCCACA
Mucor circinelloides               CCCAUCA CGGCCUCA
Phycomyces blakesleeanus           CCCAUCA CGUCCUCA

Fig. S5. (Continued)
Fig. S6. Alignment of the last 12 sense codons of 49 antizyme inhibitor uORF-Ms also showing nucleotide substitutions in the same region. Species names are indicated on the right. Alternating codons are shown either in red or black letters. Synonymous substitutions are highlighted in green. Nonsynonymous substitutions are highlighted in red.

- inferred ancestral sequence

Ivanov et al. www.pnas.org/cgi/content/short/0801590105
17 of 19
Fig. S7. Full-range mutational analysis of the 5′ leader of mouse antizyme inhibitor mRNA in mammalian HEK-293 cells using dual luciferase assay supplementing data shown in Fig. 2 of the main text. The six individual mutations described in Fig. 2 of the main text are here shown as M1–M6 instead. M1, AUG → AAA mutation of conventional uORF1; M2, AUG → AAA of conventional uORF3; M3, AUU → UUU of uORF-M; M4, AUU → AUG of uORF-M; M5, out-of-frame mutation at the C terminus of uORF-M; M6, AUG → AAA mutation of conventional uORF2. For specific details see Plasmid Construction and Primers above.
Fig. S8. Correlation between protein expression and mRNA steady-state levels for select constructs described in Fig. 2. Constructs are as follows: AUU the same as wild type (line 1 on Fig. 2); UUU is the AUU to UUU mutation (line 5 on Fig. 2); AUG is the AUU to AUG mutation (line 6 on Fig. 2); C-pep is the out-of-frame mutation changing the last 10 aa of uORF-M (line 7 on Fig. 2). (A) Anti-Renilla Western blot analysis. − and + indicate DFMO-treated cells ± 1 mM spermidine. (B) Corresponding absolute Renilla luciferase activity. (C) Northern blot of Renilla luciferase and GAPDH mRNAs. − and + indicate DFMO-treated cells ± 1 mM spermidine. (D) Densitometry results for the Northern blot shown in C. The Renilla luciferase readings are normalized relative to GAPDH. (E) Luciferase activity shown in B normalized to the mRNA levels as determined in D.