# Restriction in the cleavage activity of hammerhead ribozymes ensures ongoing evolution in prebiotic RNA world*

KUMUD K SINGH and SUBHENDU GHOSH*

*Department of Biophysics, University of Delhi South Campus, Benito Juarez Road, New Delhi 110 021, India*

**Corresponding author (Fax, 91-11-688-5270; Email, bp@dusc.ernet.in).*

Self-cleaving infectious RNAs found in many plant viruses and viroids can also cleave in *trans* and form hammerhead type secondary structure. It has been observed that the cleavage site must contain the triplet GUC. Also, in other cases, the sequence XUY holds good where X = A, C, G, U and Y = A, C, U but not G. The high electronegative nature of guanosine holds the key to its resistance to cleavage which does not allow hybrid formation between the ribozyme and substrate strands. Guanosine resistance to cleavage might have been the starting thrust for the evolution of a translational initiation codon from XUG. A hypothesis is proposed in this regard and its evolutionary consequences are discussed briefly.

## 1. Introduction

Ribozymes are catalytic molecules which consist of ribonucleotides having various possible secondary structures e.g., hammerhead, hairpin, loops etc. (Cech 1987). Hammerhead type ribozymes (Haseloff and Gerlach 1988) are known to occur in various plant viruses and viroids, e.g., avocado sun blotch viroid (ASBV, both + and − strands) (Hutchins *et al* 1986), encapsidated linear satellite RNA of tobacco ring spot virus (sTRSV, + strand only) (Buzayan *et al* 1986), satellite RNAs (virusoids) of lucerne transient streak virus (vLTSV, both the + and − strands) (Foster and Symons 1987a, b), velvet tobacco mottle virus (vSNMV, + strand only) and the + strand of subterranean clover mottle virus (vSCMoV). The transcripts of newt satellite II DNA also self-cleave via a hammerhead structure where only the + strand shows catalytic activity (Epstein and Gall 1987). Interestingly, all these above mentioned self-cleaving reactions are irreversible. The RNA species are considered to replicate by a rolling circle mechanism (Hutchins *et al* 1986). These cleavage reactions have to be site-specific to produce monomeric products. That these reactions involve self-cleaving, was proved by observing transcription from DNA in the absence of any protein moiety (Hutchins *et al* 1986; Buzayan *et al* 1986). The self-cleaving reactions take place by the non-hydrolytic cleavage of the internucleotide bond by the $Mg^{2+}$ catalyzed attack of the 2′-OH on the phosphate, which gives cleaved fragments with a 2′,3′-cyclic phosphate at the 3′ end and a 5′-OH at the 5′ end.

## 2. Observations

Self-cleaving domains have a consensus hammerhead structure with the 13 conserved bases (Haseloff and Gerlach 1988). The cleavage site at these self-cleaving reactions is a GUC triplet or in some cases it is AUA (Miller and Silver 1991) or GUA (Foster and Symons 1987a). It has been proved that the conserved sequences can be arranged to cleave in *trans* (Uhlenbeck 1987). This has opened up the possibility that synthetic hammerhead ribozymes can be designed to cleave any target RNA. The specificity for the target is achieved by the flanking sequences on the either side of the target site. Interestingly, the target site, XUY can have X = A, C, G, U but Y = A, C, U only. This indicates that a guanosine at the Y position does not allow the

---

*Presented at the National Symposium on Evolution of Life.

**Keywords.** Cleavage; hammerhead ribozymes; evolution; RNA world

*J. Biosci.*, **23**, No. 1, March 1998, pp 43–45. © Indian Academy of Sciences

hammerhead ribozyme to cleave i.e., the target site XUG is uncleavable.

It has been observed that there is no example in nature where XUG is used as a target site in a self-cleaving reaction. Moreover, no synthetic hammerhead ribozyme can cleave at XUG. This clearly indicates the importance of guanosine at the target site. If we look carefully at the guanosine structure, we observe that it has the maximum number of electronegative centres (five nitrogens and one oxygen). This would result in high electron density on the heterocyclic ring of guanine. Moreover, the partial negative charge due to the electronegative –OH groups of ribose sugar and phosphates groups of the oligo-ribonucleotide is likely to offer strong electronegative repulsion to other oligoribonucleotides approaching it. Although this phenomenon can occur with ribonucleotides containing A, C or U, steric hindrance must be below a critical level. Hence, for highly electronegative G, the possibility of hybrid formation, an essential condition for cleavage, seems remote.

## 3. The proposition

From an evolutionary point of view, the resistance of G to cleavage may be highly significant. If we look at the triplet nature of the target site, we observe that the codons formed with XUG can be AUG, CUG, GUG and UUG. It is likely that once these triplets were found resistant to cleavage by hammerhead ribozymes, they might have been selected for some function during the course of evolution. If we assume that the existence of hammerhead ribozymes coincided with the evolution of the translational machinery, perhaps the AUG codon was fixed as the universal starting codon for translation, this at a time when the translational machinery was in its preliminary stage of evolution and codons were being frozen for various important functions connected with translation. Although AUG codes for methionine in internal positions of polypeptides, it is the only codon for initiating the translation process in both prokaryotes and eukaryotes. Hence it is a very important codon for the survival of a living organism, and one whose evolution likely preceded the prokaryote-eukaryote divergence in evolutionary time. Conversely, we may speculate that it would well have been the end of the evolution of the translational machinery, had any codon other than AUG were chosen as the initiation codon.

**Table 1.** Various triplet genetic codons.

| | U | | C | | A | | G | | |
|---|---|---|---|---|---|---|---|---|---|
| U | UUU | Phe | UCU | Ser | UAU | Tyr | UGU | Cys | U |
| | UUC | Phe | UCC | Ser | UAC | Tyr | UGC | Cys | C |
| | UUA | Leu | UCA | Ser | UAA | End | UGA | End | A |
| | UUG | Leu | UCG | Ser | UAG | End | UGG | Trp | G |
| C | CUU | Leu | CCU | Pro | CAU | His | CGU | Arg | U |
| | CUC | Leu | CCC | Pro | CAC | His | CGC | Arg | C |
| | CUA | Leu | CCA | Pro | CAA | Gln | CGA | Arg | A |
| | CUG | Leu | CCG | Pro | CAG | Gln | CGG | Arg | G |
| A | AUU | Ile | ACU | Thr | AAU | Asn | AGU | Ser | U |
| | AUC | Ile | ACC | Thr | AAC | Asn | AGC | Ser | C |
| | AUA | Ile | ACA | Thr | AAA | Lys | AGA | Arg | A |
| | AUG | Met | ACG | Thr | AAG | Lys | AGG | Arg | G |
| G | GUU | Val | GCU | Ala | GAU | Asp | GGU | Gly | U |
| | GUC | Val | GCC | Ala | GAC | Asp | GGC | Gly | C |
| | GUA | Val | GCA | Ala | GAA | Glu | GGA | Gly | A |
| | GUG | Val | GCG | Ala | GAG | Glu | GGG | Gly | G |

The uppermost row designates the second letter of the codon, the extreme right and left columns indicate the first and last letters of the codons respectively. The amino acid corresponding to a codon is indicated in the same row next to the codon.

At this stage, one might ask why AUG and not CUG, UUG or GUG was selected as the starting codon? It is known that cytosine is spontaneously deaminated to uridine *in vivo*. This is a slow process, yet a very significant one. This CUG could have become UUG (by gradual deamination) and thereafter it woulld not be possible to differentiate between the original UUG and UUG formed by the deamination of C in CUG. This also holds good in favour of evolution of the stable thymine base in DNA replacing the uracil of RNA. However, UUG codons are known to start translations more frequently than GUG in *Bacillus subtilis* (Glaser *et al* 1993) and in cDNA for nuclear encoded DNA polymerase $\gamma$ in the mitochondria of *Drosophila*, CUG has been found to act as initiation codon. It is interesting to find such examples because it indicates that in the early life forms such as *B. subtilis* and *Drosophila* mitochondria (mitochondria have very likely evolved from endoparasitic bacteria), UUG and CUG could act as translation initiation codon along with coding for leucine; but such UUG/CUG initiated translations are rare in higher life forms. During the course of evolution of the translational apparatus, the need for a stable and distinctive initiation codon might have reduced the possibility of selection of UUG or CUG as a long term initiation codon.

GUG normally codes for valine but in certain cases, GUG has been found to initiate protein synthesis when the normal AUG start codon has been lost as a result of genetic deletion (Leder and Nirenberg 1964). It may be mentioned here that *in vitro*, $\text{tRNA}^{\text{f-met}}$ initiates protein synthesis at AUG as well as at GUG. But biochemically, GUG must have also been ruled out as a long term initiation codon because the *in vitro* translations initiated by GUG are found to be much less efficient than those initiated by AUG (Leder and Nirenberg 1964). GUG is also the starting codon for the *in vivo* synthesis of protein 'A' of RNA phage MS2 (Fiers *et al* 1975). This signifies that GUG could indeed code for the initiation of protein synthesis in special cases. It could be that all these four codons – UUG, CUG, GUG and AUG, acted as initiation codons for early translations but later on selection pressure has favoured AUG as the stable and most efficient initiation codon in the higher life forms.

## 4. Conclusion

It will be interesting to look into, whether XUG sequences are more frequent than XUY. However, RNA sequence data has not been analysed thoroughly to address this question.

In the end, we comment, that since XUG was not cleaved by hammerhead ribozyme, AUG could emerge as the dominant initiation codon for translation and UUG, CUG and GUG codons were restricted to only lower unicellular life forms. Once AUG was fixed as the starting codon, it did not change with further evolution and when proteins took over catalitic functions from RNA, there was no reason for the codons to revert back or modify the genetic code (as proteins, having diverse structural and functional moieties, were better catalysts than RNA). In other words, once the proteins took command, RNA species could not compete with them anymore for catalytic functions. At best, most of these RNA catalysts survived as ribonucleo-protein complexes such as spliceosomes, snRNPs etc. On the basis of the above discussion it may be concluded that the restriction of cleavage by guanosine at the target site, XUG, may have been an important factor in the evolution of translational machinery in the prebiotic RNA world (Orgel 1986; Orgel and Crick 1993).

## References

Buzayan J M, Gerlach W L and Bruening G 1986 Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA; *Nature (London)* **323** 349–353

Cech T R 1987 The chemistry of self-splicing RNA and RNA enzymes; *Science* **236** 1532–1539

Epstein L M and Gall J G 1987 Self-cleaving transcripts of satellite DNA from the newt; *Cell* **48** 535–543

Fiers W, Contreras R, Duerinck F, Haegmean G, Merregaert J, Jou W M, Raeymakers A, Volckaert G, Ysebaert M, Van de Kerckhove J, Nolf F and Van Montagu M 1975 A protein gene of bacteriophage MS2; *Nature (London)* **256** 273–278

Forster A C and Symons R H 1987a Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites; *Cell* **49** 211–220

Forster A C and Symons R H 1987b Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site; *Cell* **50** 9–16

Glaser P, Kunst F, Arnand M, Coudart M P, Gonzales W, Hullo M F, Ionescu M, Lubochinsky B, Marcelino L, Moszer I *et al* 1993 *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from 325 degrees to 333 degrees; *Mol. Microbiol.* **10** 371–481

Haseloff J and Gerlach W L 1988 Simple RNA enzymes with new and highly specific endoribonuclease activities; *Nature (London)* **334** 585–591

Hutchins C J, Rathjen P D, Forster A C and Symons R H 1986 Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid; *Nucleic Acids Res.* **14** 3627–3640

Leder P and Nirenberg M W 1964 RNA codewords and protein synthesis II. Nucleotide sequence of a valine RNA codeword; *Proc. Natl. Acad. Sci. USA* **52** 420–427

Miller W A and Silver S L 1991 Alternative tertiary structure attenuates self-cleavage of the ribozyme in the satellite RNA of barley yellow dwarf virus; *Nucleic Acids Res.* **19** 5313–5320

Orgel L E 1986 RNA catalysis and the origins of life; *J. Theor. Biol.* **123** 127–149

Orgel L E and Crick F H C 1993 Anticipating an RNA world. Some past speculations on the origin of life: where are they today: *Faseb J.* **7** 238–239

Uhlenbeck O C 1987 A small catalytic oligoribonucleotide; *Nature (London)* **328** 596–600

*MS received 28 April 1997; accepted 27 February 1998*

Corresponding editors: AMITABH JOSHI and RAGHAVENDRA GADAGKAR