mRNA:guanine-N7 cap methyltransferases: identification of novel members of the family, evolutionary analysis, homology modeling, and analysis of sequence-structure-function relationships

ABSTRACT

Using computational methods, the structure of Eukaryotic cap MTase was predicted to be reconstructed using evolutionary relationships. The identification of novel cap (cap) metabolites suggests candidates for cloning and biochemical characterization, while the structural model will be useful in designing new experiments to understand the molecular function of cap caps.

INTRODUCTION

Many disciplines have developed graphical representations of complex systems to communicate, archive and analyze knowledge. Although electronic circuit diagrams and architectural plans are accessible to knowledgeable individuals, there is no universal'real' describing functional systems in biology. Knowledge in specific areas, such as metabolic pathways, gene networks, signaling networks and molecular interactions, is stored and archived in various forms in print or online, but these representations are not standardized to account for cross-disciplinary systems – for example, the interactions between genes and metabolism at the inter- and intra-molecular levels. Inadequate cartoon diagrams are frequently used by biologists to represent multidisciplinary biological problems at various levels, such as intramolecular and disease phenotype, due to their ambiguity and lack of clarity. We propose that a standardized visual biological description language would offer easier and more clear communication, and, with computational resources, provide underlying conditions for distributed searchable archives of functional knowledge, as well as CAD functionality for simulating and analyzing biological systems. Our team is currently developing a prototype biological description language, BioD, to test its conceptual foundation, explore its utility, and identify critical issues surrounding its implementation.

CONCLUSION

Specifically, we describe how the mouse RPTP (mouse phage-tumor-like) cDNA is cloned into the human skeletal muscle, the way in which the mice and humans are genetically expressed through alternative splicing of their respective genes, and the presence of an 8 kb 3'UTR in human "RPTTP" (human minus 2), The largest PCRTA gene known to date, spanning over 1 megabase pairs of genomic DNA, with its considerable length, mostly due to expanded introns in this region). Encoding the protein domains in the extracellular segment consist of modules that are flanked by phase 1 introns, while the majority of intracellular segments are phase 0 and relatively small. These data indicate that the ectodomain originated from exon shuffling and duplication and eventually fused with another phosphatase domain at a later time. The MAM domain, which is the region defining type IIB phosphatases, has a genomic structure that is typical of all these domains when located at the N-terminus. Additionally, three exons encode the fourth fibronectin repeat in RPTP, an extra property present in only type I ATPaSeq. At least two spliced exons flank the transmembrane domain, which is the region of greatest variation between the four IIB phosphatases; another resembling an alternatively arranged exonet precedes the catalytic core of the first ATPase. The genomic structure of representative members of the RPTP family (types I-V) shows

that the intron/exon organization of both phosphatase domains is highly conserved. Significant variation exists in the length of their 3' UTRs; the longest known record of a regulated transcriptional region at 8 kb is characterized as 3'UTR, or "under regulation" UTM. We have achieved the first-ever characterization of the genomic structure of an RPTP type IIB gene. This knowledge will assist in future research on the regulatory factors that influence tissue specificity of gene expression.