IDENTIFICATION OF 5 NOVEL SELENOPROTEINS BASED ON RNA STRUCTURAL TAGS

A. Lescure, D. Gautheret*, P. Carbon, and A. Krol

*EP91 CNRS "Information Génétique et Structurale"
Marseille, France
UPR 9002 du CNRS, IBMC
15 rue René Descartes 67084 Strasbourg
France

Selenoproteins have been shown to play an important role in essential cellular processes and the prevention of cancer. However, the physiological role of selenium is not fully understood. In order to better address this question, we have undertaken the identification of new selenoproteins by an original approach.

Selenocysteine incorporation into selenoproteins arises from readthrough of an in frame UGA codon. In eukaryotes, this mechanism requires the presence of the SECIS element, a specific RNA structure residing in the 3'UTR of selenoprotein mRNAs. Based on experimental data, a consensus secondary structure for the SECIS element has been proposed in our laboratory (Walczak *et al*., 1996; see also communication by Dr Alain Krol). Using a program enabling detection of RNA secondary structures, computational searches of genomic or EST databases led to the discovery of 52 potential SECIS elements capable of adopting the consensus secondary structure. Among these, 33 corresponded to already characterized selenoprotein mRNAs. The remaining unknown SECIS elements were tested for their abilities to promote insertion of selenocysteine *in vivo*. Five novel functional SECIS elements were identified in this manner. The corresponding cDNAs were subsequently sequenced and, as expected, the new SECIS elements are indeed localized in the 3'UTRs of the mRNAs. The coding sequences have been identified and all contain at least one in frame UGA codon. Potential functions for these selenoproteins, based on protein domain homologies and cellular localization, will be presented.

To our knowledge, it is the first time that a strategy based on a search with a consensus RNA secondary structure has been taken to uncover the existence of new proteins. This allowed us to identify cDNAs coding for five novel selenoproteins of as yet unknown function.