Designing of a single gene encoding four functional proteins

Masayori inouye, Yojiro Ishida, Keiko Inouye

www.elsevier.com/locate/yjtbi

PII: S0022-5193(17)30051-6
DOI: http://dx.doi.org/10.1016/j.jtbi.2017.01.042
Reference: YJTBI8952

To appear in: *Journal of Theoretical Biology*

Received date: 6 December 2016
Revised date: 25 January 2017
Accepted date: 31 January 2017

Cite this article as: Masayori inouye, Yojiro Ishida and Keiko Inouye, Designing of a single gene encoding four functional proteins, *Journal of Theoretical Biology,* http://dx.doi.org/10.1016/j.jtbi.2017.01.042

Designing of a single gene encoding four functional proteins

Masayori inouye, Yojiro Ishida and Keiko Inouye

Department of Biochemistry, Rutgers-Robert Wood Johnson Medical School, Center for Advance Biotechnology and Medicine, Piscataway, New Jersey 08854 USA

Abstract

In the genomes of some organisms such as bacteriophages and bacteria, a DNA sequence is able to encode two different proteins, indicating that genetic information is compacted in DNA twice denser than in usual DNA. In theory, a DNA sequence has a maximal capacity to produce six different mRNAs, however, it is an intriguing question how many of these mRNAs are able to synthesize functional proteins. Here, we design a DNA sequence encoding four collagen-like proteins, two, (Gly-Arg-Pro)n and (Gly-Ala-Pro)n, from a sense mRNA and the other two, also (Gly-Arg-Pro)n and (Gly-Ala-Pro)n from its antisense mRNA, all of which are expected to form triple-helical structures unique to collagens. Other designs such as the combination of (Gly-Arg-Pro)n, (Gly-Val-Pro)n, (Gly-Thr-Pro)n and (Gly-Arg-Pro)n are also possible. The proposed DNA sequence is considered to contain the most compact genetic information ever created.

Keywords: Collagen; Protein Engineering

Production of two collagen-like proteins from an mRNA

The genomes of bacteriophages are known to have genes encoding two different proteins (Sanger et al., 1977). In some cases, a gene encodes two proteins in the same orientation, while in the other cases two genes are overlapping in the opposite orientation. More recently, the entire genomes of various organisms have been sequenced including the human genome, and numerous overlapping genes have been