BMC Systems Biology

Poster presentation

Computer aided multi-parameter gene design: impact of synthetic dnas on protein expression enhancement

Frank Notka*[1], Doris Leikam[1], Asli Bauer[2], David Raab[1], Marcus Graf[1] and Ralf Wagner[1]

Address: [1]GENEART AG, Regensburg, Germany and [2]University of Regensburg, Molecular Microbiology and Gene Therapy, Germany

Email: Ralf Wagner - ralf.wagner@geneart.com

* Corresponding author

from BioSysBio 2007: Systems Biology, Bioinformatics and Synthetic Biology
Manchester, UK. 11–13 January 2007

Published: 8 May 2007

BMC Systems Biology 2007, **1**(Suppl 1):P60 doi:10.1186/1752-0509-1-S1-P60

This abstract is available from: http://www.biomedcentral.com/1752-0509/1?issue=S1

Protein production in cells is dependent on various factors including the underlying nucleotide sequence. Gene optimization is dedicated to improve the expression properties of transgenes by codon adaptation to the individual host, increasing RNA production, stability and nuclear export. However, most gene optimization strategies depend on codon usage adaptation only, whereas RNA optimization relies on optimization of many different parameters such as removal of RNA secondary structures, adjustment of CG-values, avoidance of splice sites and elimination of instability elements. With the help of a multi-parameter optimizing algorithm, degeneration of the genetic code provides a powerful tool to identify, analyze as well as utilize parameters and respective motifs to increase and/or adjust expression yields or other important properties of a gene such as its safety or genetic stability.

Interestingly, in higher eukaryotes the overall CpG content was demonstrated to be crucial for the level of transgene expression. In particular, we report here the intragenic CpG-dinucleotide dependent expression pattern of differently designed synthetic genes, encoding human and murine cytokines, the green fluorescence protein and HIV-1 proteins. Expression yields were monitored on the protein level for all genetic variants and on the RNA level for GFP-recombinant cell lines. For all tested transiently transfected and stable cell lines, a clear correlation of intragenic CpG-content, levels of cytoplasmic mRNA and protein yields has been demonstrated. Using the maximum number of CpGs, expression yields could be increased by more than 100% with respect to simple codon usage adaptation and several-fold with respect to wild type sequence.