Journal of Computational Mathematics
Vol.32, No.6, 2014, 601–629.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1405-m4370

DISCONTINUOUS GALERKIN CALCULATIONS FOR A NONLINEAR PDE MODEL OF DNA TRANSCRIPTION WITH SHORT, TRANSIENT AND FREQUENT PAUSING*

Lisa Davis Tomáš Gedeon
Department of Mathematical Sciences, Montana State University,
P.O. Box 172400, Bozeman MT 59717
Email: davis@math.montana.edu gedeon@math.montana.edu
Jennifer Thorenson
Department of Mathematics, University of Oregon, Eugene, United States
Email: jthoren2@uoregon.edu

Abstract

A discontinuous Galerkin finite element method is used to approximate solutions to a classical traffic flow PDE. This PDE is used to model the biological process of transcription; the process of transferring genetic information from DNA either to mRNA or to rRNA. The transcription process is punctuated by short, frequent RNAP pauses which are incorporated into the model as traffic lights. These pauses cause a delay in the average transcription process. The DG solution of the nonlinear model is used to calculate the delay and to determine the effect of the pauses on the average transcription time. Numerical error measurements between the DG solution and the true solution (derived by the method of characteristics) are given for a simple model problem. It shows an excellent agreement in a neighborhood away from the shocks as well as $\mathcal{O}(\Delta x)$ convergence for the delay calculation. Preliminary parameter studies indicate that in a system with multiple pauses both the location and time duration of the pauses can significantly affect the average delay experienced by an RNAP.

Mathematics subject classification: 65M60, 35R05, 35Q92, 35L65.
Key words: Discontinuous Galerkin finite element method, Transcription, LWR traffic flow model.

1. Introduction

Protein synthesis is a complex biological process that requires an immense amount of cellular energy as well as coordination and regulation of hundreds of molecules. The two stages of protein synthesis are *transcription* and *translation*. These two processes form the crucial components in the transfer of genetic information from DNA to protein. Transcription involves the transfer of the genetic information from the DNA to messenger mRNA. In bacteria, RNA polymerases (RNAP) translocate along the DNA strand and transcribe the DNA's genetic code into the mRNA molecule. Translation is the process by which ribosomes translocate along a mRNA strand and produce proteins. The ribosomes provide the so-called assembly machinery which uses the genetic information contained in the mRNA to construct a polypeptide chain which subsequently folds to form a protein. As noted in [20], each of these processes is quite complicated and requires the availability of various amino acids as well as assembly of various

* Received March 12, 2013 / Revised version received May 4, 2014 / Accepted May 26, 2014 / Published online September 3, 2014 /

initiator and termination complexes. The recent paper by Zia and collaborators [20] gives an excellent overview of these processes for the reader that may be unfamiliar with the subject. Moreover, the authors point out that the systematic study of such a process can follow either of two paths. The first is the path of building a complex and detailed model that may provide quantitative information that can be compared with experimental data of a very particular setting. Typically, it is very difficult to mathematically analyze the model or to extract from it any fundamental behaviors associated with the physical phenomena. An alternative approach is to consider a simple model that includes only a few of the fundamental underlying mechanisms of the process but hopes to capture the behavior of the system in some general sense. The disadvantage of such an approach is that these types of models rarely exhibit the kind of fidelity that allows for comparison with experimental data. The focus of this paper is the latter approach where a relatively simplified mathematical model is considered as an attempt to gain insight into the limiting behavior of the protein production process under certain cellular conditions.

On a very basic level both transcription and translation consist of a motion of a complex machine along a one-dimensional strand. Each process has roughly four parts: assembly of the machinery at the start of the strand, movement initiation, elongation and termination. In this paper, we concentrate on the process of elongation within the transcription stage, that is, RNAP motion along the DNA strand. It has been observed experimentally that the motion of RNAP is not uniform. Single molecule observation using optical traps indicate that the elongating polymerase transcribes rapidly but frequently pauses, and the duration of the pauses is roughly bi-modal [17] with means 1.2 ± 0.1 sec with amplitude 60% and 6 ± 0.4 sec with amplitude 40%. There are two types of transcriptional pauses of RNAP [13]. One type is referred to as backtracking pauses, and the other is called non-backtracking pauses, described in [11]. For the purposes of this work, we are only concerned with the non-backtracking pause case.

The existence of transcriptional pauses brings up the possibility that under conditions of high transcription initiation rates, a paused polymerase can prevent the forward motion of one or more polymerases that follow it, thus creating traffic jams [12]. While the density of polymerases on most genes is probably not sufficiently large to produce traffic jams, there are special genes, where the density of polymerases is very high. As an example, transcription of the ribosomal rRNA accounts for over half of the total transcriptional activity in Escherichia coli in high growth conditions [1], even though rRNA (rrn) operons only account for 0.5% of the total genome. Thus, under favorable environmental conditions, most of the cell's metabolic capacity is devoted to making ribosomes [5]. To sustain the high cellular demand on ribosome RNA synthesis, the density of polymerases on the rrn operon is very high. In [6], three types of models are considered in order to quantify the combined effect of high polymerase density and ubiquitous pauses on the transcription rate of ribosomal RNA. One of the three models is a continuum model taking the form of a hyperbolic conservation law with a nonlinear flux component. This paper serves as a companion to that work, and it details the Discontinuous Galerkin finite element methods used for the simulation results of the PDE model used in [6].

We present a study of the behavior of numerical solutions of various formulations of the nonlinear model, and these results are used as a means to analyzing the fundamental biological questions related to the transcription process. An advantage of using this particular nonlinear PDE is that RNAP pauses at bottleneck nucleotide locations along the DNA strand can be modeled in a very simple manner, and closed form solutions can be written for several model problems of interest. This allows us to validate our simulation results prior to tackling the