2014

Simulation Supported Estimation of End-to-End Transmission Parameters in Non-Viral Gene Delivery

Beata Wysocki
University of Nebraska-Lincoln, bwysocki2@unl.edu

Timothy M. Martin
University of Nebraska-Lincoln, timothy.michael.martin@gmail.com

Tadeusz A. Wysocki
University of Nebraska-Lincoln, wysocki@uow.edu.au

Angela K. Pannier
University of Nebraska-Lincoln, apannier2@unl.edu

Simulation Supported Estimation of End-to-End Transmission Parameters in Non-Viral Gene Delivery

Beata J. Wysocki,
Department of Computer and Electronics Engineering,
University of Nebraska-Lincoln, Omaha, NE 68182 USA
beata.wysocki@unl.edu

Tadeusz A. Wysocki *Senior Member IEEE*
Department of Computer and Electronics Engineering,
University of Nebraska-Lincoln, Omaha, NE 68182 USA
twysocki2@unl.edu

Timothy M. Martin
Department of Biological Systems Engineering,
University of Nebraska-Lincoln, Lincoln, NE 68583 USA
timothy.michael.martin@gmail.com

Angela K. Pannier
Department of Biological Systems Engineering,
University of Nebraska-Lincoln, Lincoln, NE 68583 USA
apannier2@unl.edu

***Abstract* — Communications, in general, involve delivery of information from a source to a sink. At nano-scale, an example of a man-made communications involving interfacing with biological systems at intra-cellular level is non-viral gene delivery. From a telecommunications engineering perspective, important end-to-end parameters of such a system are: the end-to-end delay, system capacity, and packet loss rate. There are neither known methods to estimate those parameters theoretically nor they are ready available from standard measurements. The paper provides estimates for those parameters based on the simulation of non-viral gene delivery system based on the queuing theory. The simulator used has been validated through the series of in-vitro laboratory experiments.**

Keywords — Genetic communication, Molecular communication, Communication networks, Nanobioscience

I. INTRODUCTION

Substantial amount of research activity has been devoted in recent years towards developing the concept of nano-communications, including establishment of an IEEE study group P1906.1. Its "*goals are in summary to define nanoscale communication networks in a manner that captures the economic importance of these communication networks, is useful to industry-academic partnerships, is broad enough to include complexities arising in the medical or other industries, and is a framework upon which to build that will stimulate vision for a family of technologies in nanonetworking and multi-scale integration*" [1].

A significant body of research in the area of nano-communications focuses on specific communication activities that take place among living organisms at the microbial level, or communication pertaining to molecular propagation [2-5]. There is also a very large body of research dealing with nano-scale processes within living organisms at inter- and intra-cellular levels [6] that can provide insight into nano-scale communications. For that knowledge to be successfully used to enable efficient communication between macro- and nano-devices and among nano-devices themselves, the available results must be analyzed from a telecommunications engineering perspective, so that useful communication models can be derived.

There have been also some attempts to develop a physical end-to-end model for molecular communications [5] and to derive the theoretical limits, in terms of channel capacity, for molecular communications [7]. However, the telecommunications specific parameters of intra-cellular communication channels have not been fully studied to date. This has been the case partly due to the fact that those parameters are of no interest to majority of researchers conducting experiments in the area of cellular biology, and partly due to the fact that those parameters, like channel delay, or packet loss rate, are sometimes extremely difficult to measure at intra-cellular scale.

A prominent example of an engineered communication that happens at intra-cellular scale is the process of non-viral gene delivery, where genetic information encoded as DNA is to be transported to a cell nucleus in a form of plasmid DNA (pDNA). Those pDNA units contain information that is transmitted to the cell nucleus and result in the production of a protein, instructed by the code in the DNA, and expression of this protein alters the operation of the cell and can be considered thus as data packets. This approach has been used in [8] to model non-viral gene delivery as a telecommunications system using a queuing theory [9] and implemented using SIMULINK.

The model has been since improved [10], taking into account biological factors like mitosis and random death of

978-1-4799-2003-7/14/$31.00 ©2014 IEEE

successfully transfected cells (i.e. cells that successfully received the pDNA) due to toxicity caused by the presence of foreign proteins produced by the transfected cells. It has then been implemented in MATLAB and the simulation results showed remarkable agreement with experimental data [10].

Therefore, we conjecture here that this simulation model can be used to obtain estimates of some of the end-to-end parameters that are difficult to measure directly. In this paper we use the queuing model of nonviral gene delivery to estimate an effective end-to-end delay as time lapsed from internalization of a first complex carrying plasmids to the arrival of the first plasmid in the nucleus, an effective system capacity measured as a maximum plasmid arrival rate at nucleus multiplied by the information content of a single plasmid, and a packet loss rate measured as a rate of plasmids lost inside the cell to the total number of plasmids internalized to the cell.

The paper is organized as follows. Section II describes the system model of non-viral delivery considered as a telecommunications network. Simulation results for cell transfection are presented in Section III and compared with the results of in-vitro experiments. In Section IV, the communication channel between the cell membrane and nucleus is characterized, and Section V concludes the paper.

II. System Description

Plasmid pDNA used to transfect a cell is delivered to cells in a form of so-called complexes created by mixing DNA with cationic lipid, Lipofectamine 2000 (LF2000; Invitrogen), following manufacturer's instructions. Upon binding to a cell, nearly all complexes, containing on average 10 plasmids per complex, enter the cell through an endocytic pathway [12, 13]. A fraction of the complexes within endosomes are able to escape into the cytoplasm, where they can either undergo unpacking into individual plasmids or bind to nuclear localization sequence (NLS)-containing cytoplasmic proteins [13,14], which then shuttle the complexes into the nucleus. Individual plasmids, after unpacking from complexes, also bind to nuclear localization sequence (NLS)-containing cytoplasmic proteins. However, on their way to the cell nucleus the plasmid suffer degradation because of the exposure to enzymes located within the cytoplasm. As a result, only a fraction of the unpacked plasmids is delivered to the destination – i.e. cell nucleus. A more detailed description can be found in [10].

In [8], we modeled the non-viral gene delivery utilizing a four layer protocol stack, and the communication chain for gene delivery and the transfection process is given in Fig. 1. [8]. The model of Fig. 1 has been implemented in MATLAB as a queuing network with service rates μ_i, i = 1, …, 6, determined based on the available kinetic constants [10]. The details of the implementation in SIMULINK can be found in [8]. The model has been further refined to account for the process of random cell death due to overproduction of foreign proteins following a successful transfection. Implementation of that refined model in MATLAB is described in [10].

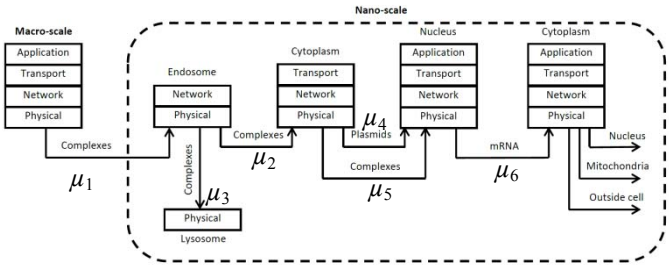

Fig. 1. Non-viral gene delivery process described using a four layer communication protocol followed by a successful transfection [8]; the service rates μ_i, i = 1, …, 6, used in simulation are listed in Table 1.

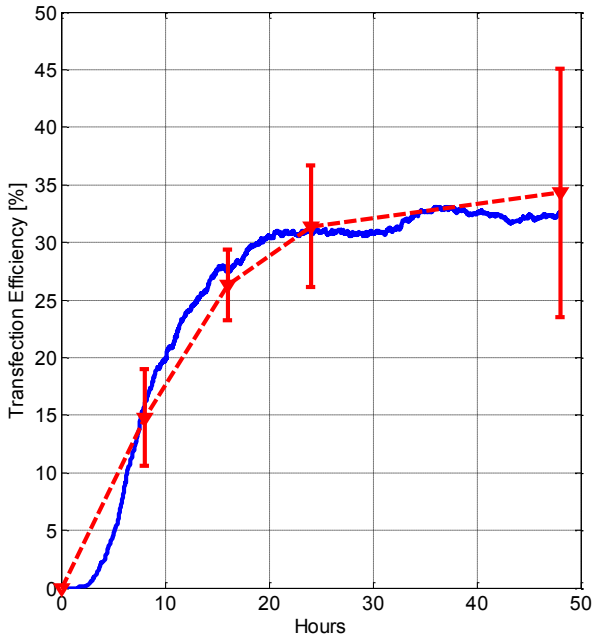

Fig. 2. Comparison between average transfection efficiency obtained from simulation – 2000 runs and the average rate obtained from the experiment [10] with experimental standard deviation shown as error bars; experimental data dashed line and simulated results solid line.

III. Simulation Results of Transfection Efficiency

To validate telecommunications model of nonviral gene delivery, simulations were performed of an experiment, where DNA encoding green fluorescent protein (GFP) was delivered to HeLa cells, using a cationic lipid, Lipofectamine 2000 as described in [10]. The simulation parameters applied were calculated based on kinetic constants published in literature.

The full list of the numerical values of kinetic constants used in is given in [10], while Table 1 shows the values of the service rates μ_i, $i = 1, \ldots, 6$, indicated in Fig. 1.

Table I. LIST OF SERVICE RATES USED FOR SIMULATION.

Service rate	Value [s^{-1}]	Source
μ_1	1.45	[15]
μ_2	1.7×10^{-4}	[14]
μ_3	3.3×10^{-4}	[15]
μ_4	1.7×10^{-3}	[15]
μ_5	3.5×10^{-3}	[16]
μ_6	5×10^{-2}	[18]

The simulation was repeated for 2000 cells with duration of the experiment set to 48 hours after delivery of DNA with 0.5 second time increment. Seeds used to draw random variables were randomized using MATLAB `rng('shuffle');` command, so different seeds were used every time a random number was drawn. The detailed description of the model is also presented in [10] together with description of an in-vitro experiment used to validate the simulation tool.

Fig. 2 shows a comparison between average transfection efficiency obtained from simulation and the rate obtained from the experiment, where transfection efficiency was measured using fluorescence activated cell sorting (FACS). The standard deviation is shown as the error bars. To validate prediction from the model, a χ^2 test was performed, with the null hypothesis that the simulated curve and the expected results from the experiment come from different distributions. The values for the test are shown in Table 2.

Table II. VALUES FOR THE χ^2 TEST.

Time instant	Predicted value	Expectation from experiment	Variance from experiment
8 h	14.82 %	14.8 %	16.81
16 h	27.52 %	26.3 %	9.61
24 h	31.24 %	31.4 %	28.09
48 h	33.82 %	34.3 %	116.64

The χ^2 estimator for the values in Table II yields:

$$\chi^2 = \frac{(14.82 - 14.8)^2}{16.81} + \frac{(27.52 - 26.3)^2}{9.61}$$

$$+ \frac{(31.24 - 31.4)^2}{28.09} + \frac{(33.82 - 34.3)^2}{116.64} = 0.1578$$

For the 3 degrees of freedom, the value for χ^2 at 0.05 significance level ($\alpha = 0.05$) is 7.815. Therefore, the null hypothesis has to be rejected and there are no reasons to consider that there is a statistical difference between the data obtained from the model and the experimental results. Hence the model can be considered as a validated one.

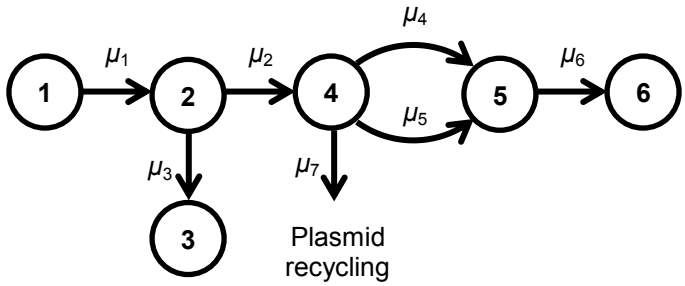

Fig. 3. Simulated queuing network for estimation of the end-to-end transmission parameters of non-viral gene delivery; 1 – outside of the cell, 2 – endosome, 3 – lysosome, 4 – cytoplasm, 5 – nucleus, 6 – cytoplasm, service rates μ_i, $i = 1, \ldots, 6$, used in simulation are listed in Table 1, μ_7 – unpacked plasmid degradation rate in cytoplasm.

IV. ESTIMATION OF END-TO-END PARAMETERS OF NON-VIRAL GENE DELIVERY PROCESS

In standard telecommunication networks it is feasible to quantify transmission parameters for any given link. However, for the communication path followed in non-viral gene delivery, it is very difficult to establish those parameters for individual links interconnecting the nodes shown in Fig. 1. This is due to the fact that some of those nodes, i.e. cytoplasm, are distributed along the link interconnecting the cytoplasm and the nucleus, while endosomes are floating inside the cytoplasm. Thus, all the nodes and the links presented in Fig. 1 are rather abstractive. Moreover, all processing inside the cell is highly parallel and the number of parallel servers and links depends on the number of packets being transmitted. Therefore, instead of establishing the parameters for the given links and processing delays for the intermediate nodes, it is much more practical to establish those parameters for the end-to-end path from a complex arriving at an endosome to the free plasmid arriving in the nucleus.

As indicated earlier, it is quite infeasible to measure exact end-to-end delay and other transmission parameters experimentally in-vitro. However, since the final results of

transfection rate from the simulation agrees almost perfectly with the measured results, the estimates from the simulation environment have been used. Because toxicity caused by overproduction of foreign proteins by the successfully transfected cell does not impact transmission parameters of the system, its influence has been omitted in estimating those parameters. In addition, it has been assumed that the parameters are estimated for cells not going through mitosis. Thus a simplified queuing network has been considered here, which is shown in Fig. 3. The direct mapping of a system model of Fig.1 onto the queuing network of Fig. 3 can be obtained by considering interconnection between the queues numbered from 1 to 6 in Fig. 3 and the service rates for the nodes of Fig. 1. To make the system complete, a plasmid degradation path has been added in Fig. 3, with the service rate $\mu_7 = 8.3 \times 10^{-5}\ s^{-1}$ [14].

Each of the queues of Fig. 3 was implemented in MATLAB as a counter, which count increases in a given time step if an object (complex/plasmid/mRNA/protein) arrives at the queue input during that time step, and decreases if a service condition is met. The service condition is assumed to be met if a number randomly drawn from the interval [0, 1] is lower than the actual service rate for that particular queue at that particular time instant.

Because the queues in the model are considered as M/M/∞ type, the actual service rates used in simulation steps are given by

$$\mu_i^* = l_i \mu_i, \quad i = 1, \ldots, 7$$

where l_i – the length of ith queue in the current simulation step. To ensure proper simulation conditions, and to avoid an effect similar to aliasing, the actual service rate must be kept below 0.5 for all the queues in the network. This has been ensured by choosing the time increment in simulation equal to 0.5 millisecond.

To estimate the end-to-end delay, the time between arrival of the first complex in an endosome in a cell and the arrival of the first free (unpacked and free from nuclear localization sequence) plasmid arrives in the cell nucleus. That end-to-end delay estimate is 555 seconds.

After arrival of the first plasmid in the nucleus, more plasmids arrive in succession to the time when all complexes internalized are processed (either unpacked or destroyed in lysosomes) and no plasmid is left in the cytoplasm. Because processing and transmission are parallel, the effective system capacity can only be measured as a maximum rate of the plasmids' arrival in nucleus. The plasmids used in the experiment were pEGFPLuc, having 6285 nucleotides. Each nucleotide can be one of four types, designated by their letter for simplicity: G, C, A, or T. Therefore a plasmid contains 6285 4-level data symbols, which is equivalent to: 12,570 bits.

Given the average rate of plasmid arrival in nucleus during the peak of their arrival between 3rd and 4th hour after the start of experiment (see Fig. 4) being 1.1983 plasmids/second, the average effective system capacity can be estimated as 15,063 bits/s.

On the way to nucleus, complexes can be destroyed in lysosomes while unpacked plasmids can be degraded in cytoplasm. Therefore, only a fraction of the internalized plasmids arrive safely in the cell nucleus. Hence, the total number of plasmids that arrive in the cell nucleus subtracted from the total number of plasmids internalized equals to the number of plasmids destroyed on the way to the nucleus. A ratio of that value to the total number of plasmids internalized provides an estimate for the packet loss rate in the system.

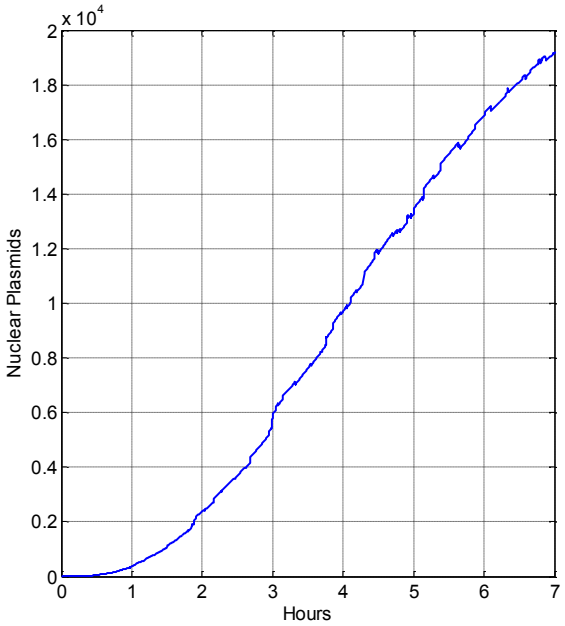

Fig. 4. Plot of the number of plasmids delivered to the cell nucleus in the first 7 hours of the non-viral gene delivery process. Kinetic constants used in simulation are listed in Table I.

Table II. SIMULATION RESULTS.

Parameter	Mean	Standard Deviation	Units
End-to-end delay	555	184.5	seconds
Effective Capacity	15,063	559	bits/second
Packet (plasmid) loss rate	82.47	0.22	%

The values of estimated parameters averaged over 1000 cells together with the standard deviations are presented in Table 2. From the simulation, it appears that the packet loss rate for the system is extremely high estimated as 82.47 %. That extremely high loss rate can explain the significant variations in the end-to-end delay estimate among the different cells. The standard deviation of the end-to-end delay estimate is equal to 184.5 second.

V. CONCLUSIONS

In the paper, we estimated some end-to-end transmission parameters for non-viral gene delivery to mammalian cells considered as a nano-communication system. The estimated parameters were: the end-to-end delay, the effective system capacity, and the packet loss rate. The used approach was based on utilizing the system simulator developed first in [8] and validated through in-vitro experiments reported in [10]. It should be noted that the estimated values of those parameters were obtained from simulation and should be confirmed in the future through series of experiments if and when the appropriate measuring techniques become readily available. In the meantime they can serve as guidelines to what can be expected from nano-communication systems developed for interfacing with biological systems at intra-cellular scale.

REFERENCES

[1] http://lifesciences.ieee.org/education/careers/132-participate-in-standards-activity-for-nanoscale-communications-networks (accessed 9/27/13).

[2] T.Suda, T.Nakano, M.Moore, A.Enomoto, and K.Fujii: "Biologically Inspired Approaches to Networks: The Bio-Networking Architecture and the Molecular Communication," *Lecture Notes in Computer Science*, Springer, vol.5151, pp.241-254, 2008.

[3] Y. Moritani, S.Hiyama, and T.Suda: "Molecular Communication for Health Care Applications", *4th IEEE PERCOM Workshop*, pp. 549-553, 2006.

[4] F.Walsh, S.Balasubramaniam, D.Botvich, T.Suda, and T.Nakano: "Hybrid DNA and Enzymatic based Computation for Address Encoding, Link Switching and Error Correction in Molecular communication", *3rd International Conf. on Nano-Networks*, 2008.

[5] M. Pierobon, I. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE Journal on Selected Areas in Communications* 28(4):602--611, May 2010.

[6] Y.Li, N.Tang, M.Aspiras, P.Lau, J.Lee, R.Ellen, and D.G.Cvitkovitch: "A Quorum-Sensing Signaling System Essential for Genetic Competence in Streptococcus mutans is Involved in Biofilm Formation," *J. Bacteriol*, 184: 2699-2708, 2002.

[7] B. Atakan, O. Akan, "On channel capacity and error compensation in molecular communication," *Transactions on Computational Systems Biology X* vol:59--80, 2008.

[8] B.J. Wysocki, T.M. Martin, T.A. Wysocki, A.K. Pannier, "Modeling nonviral gene delivery as a macro-to-nano communication system," *Nano Communication Networks*, vol. 4, Issue 1, pp. 14–22, March 2013.

[9] D.G. Kendall, "Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain," *Annals of Mathematical Statistics*, vol. 24, No. 3, pp. 338–354, Sept. 1953.

[10] T. M. Martin, B. J. Wysocki, J. P. Beyersdorf, T. A. Wysocki, A. K. Pannier, "Integrating Mitosis, Toxicity, and Transgene Expression in a Telecommunications Packet-Switched Network Model of Lipoplex-Mediated Gene Delivery" – *Biotechnology and Bioengineering* – accepted for publication (January 2014).

[11] C. M. Wiethoff, and C. R. Middaugh, "Barriers to nonviral gene delivery," *Journal of Pharmaceutical Sciences*, vol. 92, No. 2, pp. 203-17. Feb. 2003.

[12] A. Elouahabi1 and J.-M. Ruysschaert, "Formation and intracellular trafficking of lipoplexes and polyplexes," *Molecular Therapy*, vol. 11, pp. 336–347, 2005.

[13] J. Z. Gasiorowski and D. A. Dean, "Postmitotic nuclear retention of episomal plasmids is altered by DNA labeling and detection methods," *Molecular Therapy*, vol. 12, No. 3, pp. 460-467, Sept. 2005.

[14] C. M. Varga, K. Hong, and D. A. Lauffenburger, "Quantitative Analysis of Synthetic Gene Delivery Vector Design Properties," *Molecular Therapy*, vol. 4, pp. 438–446, 2001.

[15] C. M. Varga, N. C. Tedford, M. Thomas, A. M. Klibanov, L. G. Griffith, and D. A. Lauffenburger, "Quantitative comparison of polyethylenimine formulations and adenoviral vectors in terms of intracellular gene delivery processes," *Gene Therapy*, vol. 12, pp. 1023–1032, 2005.

[16] Zelphati O, Szoka FC, Jr., "Mechanism of oligonucleotide release from cationic liposomes," *Proc Natl Acad Sci U S A* 93(21):11493-8, 1996.

[17] Hakamada K, Miyake J., "Evaluation method for gene transfection by using the period of onset of gene expression and cell division," *J Biosci Bioeng*, 113(1):124-7, 2012.

[18] G. Schwake, S. Youssef, J.-T. Kuhr, S. Gude, M. P. David, E. Mendoza, E. Frey, J. O. Rädler, "Predictive modeling of non-viral gene transfer," *Biotechnology and Bioengineering*, vol. 105, No. 4, pp. 805 813, March, 2010.