Response to the Reviewers' comments

Tomáš Gedeon, Pavol Bokes

April 18, 2012

Introduction

In this document we provide responses to the comments made by the two anonymous Reviewers on our manuscript "Delayed protein synthesis reduces the correlation between mRNA and protein fluctuations", which we submitted for publication to the Biophysical Journal. Each individual comment made by the Reviewers is listed below and is immediately followed by our response to it.

We would like to thank the Reviewers for their comments, which we found helpful in improving the quality of the manuscript.

Response to the comments by Peer Reviewer 2

Overall comments by Peer Reviewer 2

*This is a report on the manuscript "Delayed protein synthesis reduces the correlation between mRNA and protein fluctuations," by Gedeon and Bokes. The authors analyze a simple stochastic model for protein production that includes a fixed time delay in transcription and translation. They use this model to explain an experimentally-observed absence of correlation between protein and mRNA concentrations. The mathematical analysis is carefully presented and the authors provide nice analytic solutions. Their result of reduced correlations is not surprising though and there are many studies on very similar systems with even more realistic delays. I am not certain whether there is existing published work precisely on the *cross* correlation between mRNA and protein, though a cursory search finds no such study. Therefore, I am inclined to recommend publication, conditioned on other strong reviews.*

However, the authors should look at (and cite) numerous publications that study very similar delayed stochastic protein production models:

- *L. F. Lafuerza and R. Toral Phys. Rev. E 84, 021128 (2011)*
- *Tobias Galla Phys. Rev. E 80, 021909 (2009)*
- *Tao Jia and R. Kulkarni PRL, 2009, 2011*
- *Barrio et al PLoS CompBio v2, 2006*

Since there are many theoretical models on the delays incurred during protein production, the authors should also maybe describe what would happen if their delays were distributed instead of a fixed constant.

Response of the authors to the overall comments by Reviewer 1

We thank the Reviewer for his encouraging comment regarding the mathematical analysis presented in our paper. The analytic solvability is, in our opinion, the single most appealing feature of the simple model that we considered. The references the Reviewer provided are indeed relevant to our work and are now mentioned in the concluding paragraph of the Introduction. A brief discussion of the additional level of complexity that would be introduced in the model if a general, stochastic, translational delay was considered is now part of the revised section "Shifting protein dynamics".

Response to the comments by Peer Reviewer 3

Overall comments by Peer Reviewer 3

In their paper Delayed protein synthesis reduces the correlation between mRNA and protein fluctuations the authors P. Bokes and T. Gideon propose a novel explanation for recent experimental results (Taniguchi et al., Science (2010)) showing that the instantaneous correlation coefficient between mRNA and their protein levels in single cell measurements in E. Coli is almost zero. The authors show that a delay time in protein translation reduces exponentially the instantaneous correlation coefficient (eq. 3), thus providing a plausible explanation for the experimental findings. This explanation based on a translational delay time seems to be new, and it is different from the one given by Taniguchi et al., where extrinsic noise in translation was offered as the most plausible explanation for the near zero correlation. The exponential reduction of the instantaneous correlation coefficient is due to the fact that the mRNA molecule, which initiated the production of a protein, can already decay while the protein is still in production.

This work is interesting and shows that the consideration of a delay time in protein translation, which is usually neglected, could be important to correctly interpret results about the correlation between mRNA and protein levels. However, I have several problems and comments with the manuscript, as stated below. In summary, we are missing a serious biological discussion about the magnitude of delay times in normal protein production. The biologically motivated delay time given in the manuscript is due to YFP maturation (around 7 min), which is a special case, adapted to the experiments performed by Taniguchi et al. We are missing here a discussion of delay times that applies to normal protein production. Would it be reasonable to assume that the low instantaneous correlation coefficient measured by Taniguchi et al. is a special case and just a consequence of YFP maturation? Would a low instantaneous correlation coefficient also be found with normal protein production? I would have expected a more elaborate discussion of the physical mechanisms that lead to a loss of the instantaneous correlation. Finally, because the idea that the instantaneous correlation is reduced due to a delay time is simple, and also the new mathematical analysis performed in the paper is simple, we suggest to revisit and shorten the manuscript in order to keep it simple and easily accessible for the readers. As written, the paper is suited for J. Th Biology or nay theoretical Journal, but not adequate for the classical readership of the BJ.

Response by the authors to the overall comments by Reviewer 3

We are grateful to the Reviewer for their extensive comments, and are glad that they found our work interesting. It seems difficult to get hold of some generic estimates, such as are available for the elongation processes, for the delay due to post-translational modification in normal proteins. Previous mathematical modelling studies that considered post-translational delay (e.g. Roussel and Zhu (2006), Phys Biol 3, 274-84) also seem to be quite unspecific about the value. We nevertheless expect that the delay can be substantial in normal protein production, having been previously implicated in driving slow circadian oscillations (Gallego and Virshup (2007), Nat. Rev. Mol. Cell Biol. 8, 139-48)

The Reviewer will find that the revised manuscript has been simplified (see also our responses to their specific comments below), and now contains a more elaborate discussion of the underlying physical mechanisms. We hope that the revised manuscript may therefore be interesting to the readership of the Biophysical Journal.

Specific comments by Reviewer 3

Comment 1 (by the Reviewer)

In the abstract (and in the the discussion) the authors claim that no correlation between mRNA and protein level put apparently severe limits on the ability of individual cells to control their protein levels. However, this does not seem to be consistent with their explanation based on a delay time. With a delay time, the lack of correlation is due to the fact that mRNA and protein fluctuations are measured at the same time, however, as the authors also point out in the abstract, a more realistic correlation coefficients can be obtained by shifting the protein abundance measurements back in time by the amount of translational delay. Hence, there is very well a correlation between mRNA and protein level in the cell, but it is hidden when considering the instantaneous correlation coefficient. This seems to be very different from situations where the lack of correlation is due to additional noise, in which case, we do not expect that the correlation increases by considering mRNA and protein fluctuations at different times. The authors should clarify these points.

Furthermore, it would be interesting if the authors could refer to experimental results, showing that the correlation coefficient increases when mRNA and protein fluctuations are considered at different times, which would strongly validate their assumption.

Response 1 (by the authors)

In the revised manuscript we have removed the claim the Reviewer refers to, partly for the reasons they detail.

We are not aware of experimental results that would confirm the predictions of the presented model. Our work suggests that such measurements would certainly be worth looking at.

Comment 2

Introduction — fourth paragraph on the second page: revisit the discussion about the consequences of a low correlation between mRNA and protein level concerning the control of a cell over its protein level (see comment 1).

Response 2

The claim about the implications of low correlation with respect to the control of protein level has been withdrawn from the revised manuscript.

Comment 3

Introduction — first paragraph on the third page: The authors should add a conclusion that leads over in a more smooth way to their work.

Response 3

We added in the Introduction a concluding paragraph providing a very brief — and certainly incomplete — review of previous approaches to modelling delayed gene expression. This leads over quite smoothly towards the model introduced right at the start of the following section.

Comment 4

Results (Eq. 1): Is it really necessary to consider a transcriptional delay time d_1? This time does not affect the loss of correlation and the results given in eq. 2 or eq. 3. The manuscript would gain in clarity without d_1 , then d_1 should be discussed.

Response 4

Indeed, it is not necessary to consider the transcriptional delay d_1. It has been eliminated and the translational delay has been renamed from d_2 to d. The role of transcriptional delay in our problem is mentioned in the revised Discussion.

Comment 5

Page 4: Why do the authors include the elongation time $\approx$ 0.5–1 min into the translational delay time d_2? We expect that the mRNA cannot be degraded during this time, and therefore the elongation time should not contribute to the loss of correlation. Why does the maturation time of the YFP protein lead to loss of correlation? This can only happen, if the protein cannot be degraded during this time, whereas the mRNA can be. The authors should explain in more detail the mechanism of YFP maturation and how this leads to a loss of correlation. What is the delay time for proteins that do not have a fluorescent part attached ?

Response 5

Since for the purposes of mRNA-protein correlation the only role of mRNA is its ability to produce protein, the mRNA joins the ranks of active mRNAs at the time when a ribosome is able to bind it. In prokaryotes this coincides with transcription ribosome binding site (RBS) on mRNA. It can well happen that mRNA is degraded (i.e. the RBS is degraded) while ribosomes that have already been bound continue to elongate amino-acid chains. Thus if we identify mRNA with its RBS, elongation contributes toward the translational delay. Other possibilities are discussed in detail in the revised Discussion.

In regards to the degradation of protein during its maturation, we simply assume that it is not possible. We think it is plausible to assume that only mature proteins can engage in biochemical reaction channels, and we treat degradation as one of such channels (this point is explicitly made in the manuscript in the revised section discussing the time shift of protein trajectory).

Comment 6

The authors should explain what it additionally brings to the manuscript to discuss the joint mRNA and protein number distribution? it is more appropriate to a more mathematical Journal. Furthermore, what do we additionally learn from considering the two conditions (Set 1 and set 2) with low and high protein abundance? The manuscript would certainly be clearer and easier to read without these parts. Furthermore, why do the authors consider for Set 1 and 2 a delay time $d_2 = 12$ *min (before they estimated* $d_2 = 7.5$ *min) ? This is based on what?*

Response 6

We believe that the analytic solvability, in terms of generating function, is one of the most appealing features of the presented simple model. The generating function also facilitates, via the numerical method based on Fourier transform, an accurate evaluation and visualisation of the probability distribution, as given in Figure 2. This figure has the additional benefit that it predicts for a very large, isogenous, population of cells the proportions of cells found in individual expression states; such a picture suitably complements the trajectories shown in Figure 1.

While we understand that the derivation of the generating function may seem technical to a part of the readership of the Journal, we would at the same time argue that similar derivations have increasingly often appeared in journals not explicitly focused on mathematical biology (see e.g. the use of exact solution in Raj et al. (2006), PLoS Biol. 4, e309). Nevertheless, we aimed to simplify Section "The joint distribution" in the revised manuscript, shifting some of the technical details into the appendices, to make that part accessible to a wider readership.

The Reviewer will find the part of the paper in which the two parameter Sets are introduced simplified too. We realised that the original version of the manuscript placed unwarranted emphasis on the parameter values used in these sets, whose sole purpose was to illustrate the behaviour we can expect from the model with large delays in comparison with no delay. We think that such a

comparison benefits the Reader by giving him a feel about the model, in particular with respect to changing the parameter d. The reason why we used two parameter sets, instead of one, is that we wanted to emphasise that our conclusions are robust with respect to gene-to-gene variation in transcription and translation rates (as opposed to degradation rates, which have to obey biological and modelling constraints to get the desirable relationships between the individual timescales operating in the model). We were also partially motivated by the distinction made in experimental studies, such as that of Taniguchi *et al.*, between low-abundance and high-abundance proteins. We hope that our intentions are clearer now that the manuscript has been revised. The value $d = 12$ min is purely illustrative, just like the rest of the parameters.

Comment 7

We propose to shorten the analysis part of the manuscript (starting from page 6). For example, the paragraph The stationary distribution for delayed protein production is somehow confusing and blurs the simple assumptions leading to eq. 2 and 3 (we further suggest to consider $d_1 = 0$*): 1) The delay time* d_2 *is constant and not stochastic. 2) The stochastic time points for protein synthesis and decay,* t_i *, in a model without translational delay corresponds to the time points* $t_i + d_2$ *in a model with delay, leading to* $N_d(t + d_2) = N_0(t)$ *(the equation after eq. 6, which should be given a label in the manuscript). The extended argumentation before eq. 6 is somehow confusing and blurs these main points. Using* $N_d(t + d_2) = N_0(t)$*, the analysis of the delayed model is reduced to the analysis of a model without delay, which has been studied previously (for example, the authors studied this model in their reference 19). The derivation of the steady state correlation coefficient with delayed times (after equation 7) is the new analysis performed in this paper.*

Response 7

We agree with the Reviewer that the original exposition given in the section they refer to was too technical. In the revised manuscript we provide a simplified version of it which follows along the lines suggested by the Reviewer in his comment. In addition, we discuss the assumptions on which the analysis relies, detailing the consequences that would follow if any of them was relaxed.

Comment 8

Page 9: We doubt that the formulation as a Theorem is suitable for the stating results in the Biophysical Journal.

Response 8

We acknowledge that and have removed the formulation from the manuscript.

Comment 9

The method to calculate the joint probability distribution based on Fourier transform was already introduced by the authors in their ref. 19. It is quite technical and not new, and we do not see what it brings to the reader to have it in the

main text. To us, the manuscript would be clearer without this part, or, if it is kept, it could be moved into an appendix.

Response 9

The exposition of the Fourier method has been moved into Appendix B. While the method is, as Reviewer correctly noticed, exactly the same as that used in Ref. 19, here we provide a slightly improved version of its derivation, from which the character of the numerical error incurred by the method (aliasing) can easily be appreciated.

Comment 10

The content of the first paragraph of the discussion has to be revisited (see our comments 1 and 2). The discussion is very short and focuses and the main conclusion of the paper. However, we miss a more elaborated biological discussion about the validity of the assumption of a delay time, and about the magnitude of the delay time. In the main text, the biological example for the delay time is due to maturation of the fluorescent protein YFP, which is a special case related to the experiment by Taniguchi et al. How does YFP maturation after elongation prevents the protein from degradation (only in this case the correlation is lost) ? What if there is no fluorescent part attached to the protein ? What are the delay times that reduce the correlation in normal protein production ? What are the biological experiments by which this can be tested?

Response 10

The assumption of no protein degradation during maturation is discussed in the section in which we determine that protein trajectories are shifted forward in time. The concluding section has been extended by a conceptual analysis of the coupling between transcriptional and translational elongation and also by a broader discussion of the mathematical techniques used in the work.

Minor comments by Reviewer 3

Minor comment 1

Change or eliminate the first part in the introduction.

Response

This suggestion has been implemented in the revised manuscript.

Minor comment 2

Explain better what modeling can do for studying stochastic gene expression and regulation: it provides a framework to quantify and analysis data....

Response

We aim to provide such an explanation in the second paragraph of the revised Introduction.

Minor comment 3

I do not see any justification for the conclusion this result seems to show that cells have very little control over this response on a single cell level. Either justify better or remove or reformulate.

Response

The statement has been removed in the revised manuscript.

Minor comment 4

Remove unnecessary comments such as "which can be thought of as a restatement of the central dogma of molecular biology in the language of elementary chemical kinetics", which are difficult to comprehend.

Response

This statement has been removed in the revised manuscript too.

Minor comment 5

The first 4 lines of the introduction should be rewritten: what is the point of these general statements without any specificity, are they in any sense related to this work ?

Response

Based on Reviewer's advice, we have removed these general statements from the revised version of the manuscript.

Minor comment 6

First sentence of the fourth paragraph on the second page: what is the meaning of "highly regarded analysis"?

Response

The phrase has been removed from the text.

Minor comment 7

Seventh sentence of the fourth paragraph on the second page: This result is disturbing on multiple levels. I suggest to reformulate this using the word surprising.

Response

The text has been amended and the adjective is no longer in use.

Minor comment 8

First paragraph on the third page: What is the central dogma of molecular biology?

Response

By central dogma we meant that the information is principally passed from DNA down to RNA and ultimately to protein, not the other way round (Crick F., Central dogma in molecular biology (1971), Nature 227, 561–563). Our original intention was to hint at a loose parallel between this concept and the two-stage model, in which DNA "catalyses" mRNA synthesis, and mRNA "catalyses" protein synthesis. In hindsight, this parallel, while worth a thought, is not very helpful, because the one-way character of the two-stage model is due to the lack of autoregulation in its simplest formulation, while the "Central dogma" is meant to be valid universally. In addition, we understand that the concept of a dogma in science — even if meant as a hyperbole — is rather controversial, so we had better do without it. The dogma is no longer mentioned in the revised manuscript.

Minor comment 9

The sentence between page 9 and 10 was interrupted.

Response

There was a typographical error in that sentence — this has been corrected.

Minor comment 10

Ref. 19 is not correctly given.

Response

The reference has been corrected in the revised manuscript.