molecular systems biology

Time-scales and bottlenecks in miRNA-dependent gene regulation.

Jean Hausser, Afzal Pasha Syed, Nathalie Selevsek, Erik van Nimwegen, Lukasz Jaskiewicz, Ruedi Aebersold and Mihaela Zavolan

Corresponding author: Mihaela Zavolan, University of Basel

Review timeline:

Submission date:	09 May 2013
Editorial Decision:	15 May 2013
Appeal:	22 May 2013
Editorial Decision:	25 July 2013
Revision received:	25 September 2013
Editorial Decision:	25 October 2013
Revision received:	28 October 2013
Accepted:	30 October 2013

Editor: Maria Polychronidou

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 15 May 2013

Thank you for having submitted a manuscript entitled "Time-scales and bottlenecks in miRNA-dependent gene regulation." for consideration for publication in Molecular Systems Biology.

Your paper has now been seen by Editors of the Journal and we have decided to return it to you without sending it for extensive peer review.

In this study, you present a mathematical model of miRNA-dependent gene regulation. We appreciate that your mathematical framework describes the regulatory dynamics in miRNA-dependent gene regulation. For example, you conclude that target protein decay rate is a limiting step in the response to miRNAs. However, we feel that these insights remain relatively modest, while they do not seem to decisively advance our understanding of miRNA-dependent gene regulation. As such, we are not convinced that your study provides the kind of decisive systems-level biological insight our audience would expect in Molecular Systems Biology and we think that your manuscript would be better suited for publication in a journal more oriented towards computational biology.

I am very sorry to have to disappoint you on this occasion, but I hope that this early decision will allow you to submit your work elsewhere without undue delay.

Appeal 22 May 2013

I appreciate your efforts in providing us with a speedy evaluation of our manuscript.

Based on your reply I must conclude that we somehow failed to clearly present the main message of our manuscript. Indeed, if the insight of our work would only be that protein decay rates influence the rate of the response to miRNAs, I would agree that this insight is rather limited.

Our manuscript provides however, much more striking and fundamental new insights into miRNA regulation. Our main finding is that, contrary to an unquestioned assumption in the field (see for instance Shimoni et al. MSB 2007; Baek et al., Nature 2008; Selbach et al., Nature 2008), miRNA-based regulation is relatively slow, being limited by the dynamics of miRNA loading in and out of the Argonaute protein. The current view in the field is summarized in the following excerpt from a review that appeared in the Science journal in 2008 (Hobert O., Gene regulation by transcription factors and miRNAs): "Speed and reversibility are other distinguishing features of miRNA-mediated gene regulation that may result in a more specialized regulatory niche of miRNAs. For transcription to be repressed, a sophisticated machinery needs to be set in place in a subcellular compartment ó the nucleus ó that is distinct from the production site of the protein product, the cytoplasm. The stability of already transcribed mRNA species sets another limit to the speed with which transcriptional repression can wipe out the expression of a target gene. In contrast, miRNAs can rapidly turn off protein production right at the site of protein production, the ribosome. Another factor that speeds up miRNA-mediated control of gene expression is that, owing to their small size and noncoding nature, miRNAs may be produced more rapidly than TFs, thereby decreasing response times to stimuli that induce gene repression." Our results show that this accepted view is incorrect, owing to the fact that it ignores the dynamics of miRNA loading and unloading into the ribonucleotide complexes that they form with Argonaute protein. Of very practical importance for experimental scientists, our results suggest that an increased miRNA turnover is necessary in situations in which miRNAs implement switches or induce oscillations in gene expression.

I believe that this aspect, on its own, would make our manuscript of great interest. Additionally, our model can support further experimental investigations into the kinetics of miRNA-dependent gene regulation. One reason is that while the kinetics of Ago-loading and miRNA degradation are difficult to manipulate experimentally, the parameters characterizing the dynamics of these processes can be inferred with the help of our model from the measured kinetics of mRNA expression. Simulations can then be used to study how changes in these parameters affect the magnitude and kinetics of miRNA-induced responses. This is highly relevant for example, in understanding the regulation of cell cycle, which is one of the processes in which miRNAs have been most clearly implicated. Our study demonstrates that accelerated miRNA turnover is necessary to achieve protein-level regulation on the time scale of the cell-cycle, and our model can support the discovery of the yet unknown underlying mechanisms. Importantly, we show that increased protein decay cannot compensate for the slow rate of miRNA loading and unloading from Argonaute. Our results also suggest a model in which the kinetics of miRNA degradation and Ago-loading differs between miRNAs, depending on their regulatory function. Evidence that the rate of degradation varies between miRNAs is accumulating [Krol et al., Cell 2010; Chatterjee et al., Dev Cell 2011] but a very recent study has also shown that Ago unloading strongly depends on the loaded small RNA and target availability [De Young et al., Mol Cell 2013]. Thus, given the current context of the field, our model could strongly support the efforts on the experimental side.

Finally, concerning your suggestion to consider a computational journal: I hope it will be clearer from my description above that the most appropriate audience for our manuscript is the systems biology community. First, we use the canonical systems biology approach, combining cycles of data analysis, modeling and experimental validation, to uncover quantitative properties of miRNA-based regulation. Second, Molecular Systems Biology has an established interest in studies characterizing the contribution of various parameters of miRNA and target gene expression to the efficacy of miRNA- and siRNA-dependent targeting, as we inferred from a few articles that were published recently in the journal (e.g. "Target mRNA abundance dilutes microRNA and siRNA activity" [MSB 6:363(2010)], "mRNA turnover rate limits siRNA and microRNA efficacy" [MSB 6:433(2010)], "Chromatin measurements reveal contributions of synthesis and decay to steady-state mRNA" [MSB 8:605(2012)]). I hope you will agree and allow our manuscript to be sent for peer

review. Of course, we would be happy to revise the manuscript so that the points I made above are more apparent. Otherwise, I would appreciate the opportunity to discuss with you by phone any aspects that I may have left unclear.

2nd Editorial Decision 25 July 2013

Thank you again for submitting your work to Molecular Systems Biology. First of all, I would like to apologize for the exceptional delay in getting back to you. Unfortunately it took us a considerably long time to find referees that were available to review your work. We have now heard back from the three referees who agreed to evaluate your manuscript. As you will see from the reports below, the reviewers acknowledge that you address a potentially interesting topic. However, they raise a series of concerns, which should be convincingly addressed in a revision of the manuscript.

We would particularly like to draw your attention to a serious concern that was raised by Reviewer #1 and refers to the log-likelihood value of 877.3 for the tri-exponential Ago-loading model. Even though this point was raised only by one reviewer, we feel that it may potentially severely undermine the main conclusions of the manuscript. As such, convincingly addressing this issue will be essential. Moreover, Reviewer #3 points out that additional kinetic bottlenecks besides loading of the Ago proteins could be involved in the described pathway.

If you feel you can satisfactorily deal with these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees. A revised manuscript will be once again subject to review and you probably understand that we can give you no guarantee at this stage that the eventual outcome will be favorable.

Reviewer #1:

The authors of this manuscript are doing important work and addressing some of the controversial questions in the field: how quickly do miRNAs degrade their targets and where is the greater impact of miRNAs (protein or mRNA level?). Answers to these questions based on mathematical modelling would make for a very interesting manuscript.

Furthermore, I think that they are taking the right approach, however I did not find their manuscript compelling. Many of the details that would help to understand the model are either in the supplement, or buried in the methods, which makes the manuscript hard to follow. A clear description of the methodology as part of the main text is important here, not only because that is part of the contribution but also its development provides the reader confidence in the accuracy of model. Furthermore, while I appreciate that it is quite difficult to track down all of the required data to fit the model but, as written, the manuscript gives the impression that values were put together in a haphazard fashion. I am sure that this is not what happened and this is likely an issue of presentation - perhaps a figure showing the source of the data for each of the parameters would help. In general, more illustrative figures would make for a more compelling story.

There was one particularly worrying sentence in the manuscript:
". With a maximum log-likelihood of 877.3, the tri-exponential Ago loading model allowed a better t the mRNA pro ling data compared to the bi-exponential model (log-likelihood of 2994.6, Fig. 2B, S2B)."

The positive log-likelihood of 877.3 is something that almost never happens unless the model is severely overfit. At first I thought this was a typo but it is repeated in figure 1B. To give you some idea of how difficult this is to get, assuming Gaussian noise, to get a positive log-likelihood of 877.3 with the 1,098 data points (i.e. ~0.8 likelihood / data point) requires a perfect fit (i.e. zero error) and a standard deviation in the noise model of 0.18. This single sentence alone is sufficient to make me suspicious of the rest of the conclusions drawn from the model.

Reviewer #2 :

The main goal of the study to use already available data in order to understand and model the dynamics of miRNA regulation is very interesting and opens a new area in the field. But this fact gives also the limitation to that the simulation is performed in 2 biological examples. The way the model will generalize in other examples remain to be seen in the future.

Minor comments

1. In order to explore the response of the KTN1 gene at both mRNA and protein abundance upon miR-199a induction, authors added to their described model represented in Fig. 1A a constant for miRNA synthesis into the Ago-accessible comportment, which is not also included in the equations in section 4.1.
2. Is a particular reason why the proposed bi/tri-exponential curve fitting model is chosen?. Could higher order exponential curves or other distributions such as weibull, gamma, lognormal also work an have a similar performance?
3 "goodness" of fit of the bi/tri exponential models has been evaluated for genes presenting miR-124 seed match in the 3' UTR. Is seed limitation for a miRNA target required in this study? Would it also work with target prediction model that take into account further binding sites next to seed or targets on the CDS as proposed in other papers from the authors themselves ?

Reviewer #3:

In the manuscript "Time scales and bottlenecks in miRNA-dependent gene regulation" Hausser et al. embark on an important aspect of small RNA-directed post-transcriptional gene silencing - the kinetics of individual steps involved in silencing. They use mathematical modeling of miRNA-mediated gene regulation by inferring the respective parameters from different experimental datasets. Based on previously published data the authors abstract a model for loading of miRNAs into Ago proteins and their functional output on mRNA and protein expression. These models were then validated using an inducible miRNA expression system in HEK293 cells, which allows the simultaneous assessment of miRNA function based on a reporter construct containing a 3¥UTR with miRNA target sites. Furthermore the authors test their models to address important issues in the filed, including the disputed effect of miRNAs on protein synthesis vs mRNA decay, and the possible biological impact of accelerated miRNA turnover on miRNA function.
While the presented data is novel and timely, the authors likely underestimate the complexity of the pathway. For example there is good evidence that association of small RNA duplexes with Ago proteins represent not the only kinetic bottleneck in the assembly of functional RISC complexes. Activation of RISC requires also expelling the passenger/miR* strand.
Nevertheless I believe the manuscript provides a starting point for dissecting the intracellular kinetics of small RNA pathways and its consequences for post-transcriptional gene silencing, an important question in the field. I therefore recommend publication of the manuscript without major changes.

Minor comments:

- page 2 last paragraph: "... for 1098 genes that had a miR-124 seed match in the 3¥UTR and were repressed on average." is incomplete.

Reviewer #1

> The authors of this manuscript are doing important work and addressing some of the controversial questions in the field: how quickly do miRNAs degrade their targets and where is the greater impact of miRNAs (protein or mRNA level?). Answers to these questions based on mathematical modelling would make for a very interesting manuscript.

> Furthermore, I think that they are taking the right approach, however I did not find their manuscript compelling. Many of the details that would help to understand the model are either in the supplement, or buried in the methods, which makes the manuscript hard to follow. A clear description of the methodology as part of the main text is important here, not only because that is part of the contribution but also its development provides the reader confidence in the accuracy of model. Furthermore, while I appreciate that it is quite difficult to track down all of the required data to fit the model but, as written, the manuscript gives the impression that values were put together in a haphazard fashion. I am sure that this is not what happened and this is likely an issue of presentation - perhaps a figure showing the source of the data for each of the parameters would help. In general, more illustrative figures would make for a more compelling story.

We thank the reviewer for his/her gracious comments concerning our choice of problem and approach to solve it. We believe that the fact that we had to reconstruct a coherent quantitative description of miRNA-dependent regulation from disparate data sets contributed to the appearance of haphazardness.

We are concerned that integrating more mathematical details in the main text would come at the expense of making the manuscript less accessible to an experimental audience (Fawcett & Higginson, PNAS 2012 109(29):11735). However, in our revision we have streamlined the description of our approach, stressing the rationale for using specific models, the parameters that we inferred at every step and that data that we used for this purpose.
We did so through changes in the main text:
- on page 2, we edited paragraphs 5 & 6
- in the methods, we rewrote the mathemical derivation of the model from the last paragraph of page 7 to the end of the section on page 8
- on page 8, we edited the first and second paragraphs of section "Fitting changes in mRNA, protein and Ribosome Protected Fragment abundance upon miRNA transfection"
- page 11, we edited the 1st and 2nd paragraphs of the "Simulations" section
We also revised main Figures 1 and 2. Finally, we added Table 1 which offers an overview on the strategy we used in our analysis.

There was one particularly worrying sentence in the manuscript: ". With a maximum log-likelihood of 877.3, the tri-exponential Ago loading model allowed a better fit the mRNA profiling data compared to the bi-exponential model (log-likelihood of −2994.6, Fig. 2B, S2B)."

> The positive log-likelihood of 877.3 is something that almost never happens unless the model is severely overfit. At first I thought this was a typo but it is repeated in figure 1B. To give you some idea of how difficult this is to get, assuming Gaussian noise, to get a positive log-likelihood of 877.3 with the 1,098 data points (i.e. ~0.8 likelihood / data point) requires a perfect fit (i.e. zero error) and a standard deviation in the noise model of 0.18. This single sentence alone is sufficient to make me suspicious of the rest of the conclusions drawn from the model.

The reviewer's remark seems to have been the result of a simple misunderstanding: we did not have 1,098 data points but rather 5490 corresponding to 5 measurements of expression of 1098 genes.

From the observed error distribution (Fig 2C of manuscript, and first panel of the attached "reviewer1.pdf" file), we can see that 80% of genes have less than 20% error. A 20% error amounts to log2(1.2) = 0.25 on log2 scale. As shown on the second panel of the PDF, under our error model (sigma = 0.25), the corresponding LL is usually positive for these prediction errors and ranges up to 0.5 per data point. A total LL of 877.3 amounts to an average LL of 0.16 per data point and is hence well compatible with such statistics.

We have edited the first paragraph of page 3 to clarify this point. If the reviewer finds it necessary, we can include these figures in the Supplementary material.

Reviewer #2:

The main goal of the study to use already available data in order to understand and model the dynamics of miRNA regulation is very interesting and opens a new area in the field. But this fact gives also the limitation to that the simulation is performed in 2 biological examples. The way the model will generalize in other examples remain to be seen in the future.

We appreciate the reviewer's positive comment about the implication of our study for the field.

> Minor comments

1. In order to explore the response of the KTN1 gene at both mRNA and protein abundance upon miR-199a induction, authors added to their described model represented in Fig. 1A a constant for miRNA synthesis into the Ago-accessible comportment, which is not also included in the equations in section 4.1.

Indeed, in section 4.1 we described a miRNA transfection experiment, in which the miRNA is provided at once. We have now added equation 14 onpage 8 together with two paragraphs that describe the full process, including the miRNA production.

> 2. Is a particular reason why the proposed bi/tri-exponential curve fitting model is chosen?. Could higher order exponential curves or other distributions such as weibull, gamma, lognormal also work an have a similar performance?

The main reason for choosing these forms is that they follow, with a small number of reasonable assumptions, from the explicit mechanistic model of miRNA loading/dissociation into/from Argonaute. It is possible that other functions give similar performance in explaining the mRNA/protein dynamics later on, but our aim was to have a relatively simple functional form that depends on easily interpretable parameters such as association, dissociation, miRNA decay rates.

> 3 "goodness" of fit of the bi/tri exponential models has been evaluated for genes presenting miR-124 seed match in the 3' UTR. Is seed limitation for a miRNA target required in this study? Would it also work with target prediction model that take into account further binding sites next to seed or targets on the CDS as proposed in other papers from the authors themselves?

The requirement of a miRNA seed match in the 3'UTR would be limitation if other types of sites would substantially contribute to the measured levels of the mRNAs, which is what we aim to explain. As we and others have shown, although non-seed sites and CDS-located sites occur with reasonable frequency, both of these types of sites have a much smaller impact mRNA stability compared to 3' UTR sites. Thus, for simplicity, we focused on mRNAs with seed matches in their 3'UTRs.

Reviewer #3:

> In the manuscript "Time scales and bottlenecks in miRNA-dependent gene regulation" Hausser et al. embark on an important aspect of small RNA-directed post-transcriptional gene silencing - the kinetics of individual steps involved in silencing. They use mathematical modeling of miRNA-mediated gene regulation by inferring the respective parameters datasets. Based on previously

published data the authors abstract a model for loading of miRNAs into Ago proteins and their functional output on mRNA and protein expression. These models were then validated using an inducible miRNA expression system in HEK293 cells, which allows the simultaneous assessment of miRNA function based on a reporter construct containing a 3´UTR with miRNA target sites. Furthermore the authors test their models to address important issues in the filed, including the disputed effect of miRNAs on protein synthesis vs mRNA decay, and the possible biological impact of accelerated miRNA turnover on miRNA function. While the presented data is novel and timely, the authors likely underestimate the complexity of the pathway. For example there is good evidence that association of small RNA duplexes with Ago proteins represent not the only kinetic bottleneck in the assembly of functional RISC complexes. Activation of RISC requires also expelling the passenger/miR strand.*from different experimental

We certainly agree with the reviewer that we took a simplified view of the assembly/disassembly of RISC complexes. Even so, as the reviewer probably appreciated, finding data sets from which we can infer the various parameters of our models was challenging. We have edited the first paragraph of the Discussion section on page 5 to clarify these points.

> Nevertheless I believe the manuscript provides a starting point for dissecting the intracellular kinetics of small RNA pathways and its consequences for post-transcriptional gene silencing, an important question in the field. I therefore recommend publication of the manuscript without major changes.

We thank the reviewer for his/her positive recommendation.

> Minor comments:
- page 2 last paragraph: "... for 1098 genes that had a miR-124 see match in the 3´UTR and were repressed on average." is incomplete.

We corrected this mistake.

3rd Editorial Decision 25 October 2013

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the referee who agreed to evaluate your manuscript. Overall, this referee feels that their main concerns have been addressed and raises only a few minor points, which we would like to ask you to carefully address in a revision of the manuscript.

As you will see below, the reviewer suggests applying the bi-exponential model to the data from the study of Karginov et al., 2007. While we would certainly welcome inclusion of such an analysis in case it is already available or can be performed within a reasonable timeframe, we do not feel that it is mandatory at this point as we would not like to substantially delay the publication of the work.

Thank you for submitting this paper to Molecular Systems Biology.

Reviewer #1

The manuscript was substantially improved by the changes and the model is now clear. As is the surprisingly high log likelihood. Nice revision, this looks like it will be a classic paper in the field.

One minor but important point is that the result of applying the bi-exponential model to other data was not discussed. It should, at least, be applied to the Karginov et al, 2007 data? The true test of a model is how much better it generalizes.

Minor details:
P2: Sentence starting with: "The fluorescence cross-correlation..." is unclear and convoluted. Try

rewriting it.
Last sentence of second paragraph of section 2.2 starting with: "Furthermore, as data...". The phrase "using instead systems" is awkward.

2nd Revision - authors' response 28 October 2013

Thank you for giving us the opportunity to revise our manuscript entitled "Time-scales and bottlenecks in miRNA-dependent gene regulation". We are again thankful for the constructive comments and happy that Reviewer #1 found our revised manuscript improved.

We have addressed the final comments of Reviewer #1 as follows:
1. We rephrased the sentence on page 2 which begins with "The fluorescence cross-correlation spectroscopy..."

2. We rephrased the sentence in section 2.2 that starts with "Furthermore, as data from other groups also...".

3. Reviewer #1 suggested we applied the bi- and tri-exponential models to the dataset of Karginov et al. (2007). This dataset however, contains only one time-point, and would therefore provide little new insight or validation. We agree with reviewer #1 that the reality test for a model is in how well it generalizes. In the manuscript we used 10 different miRNA perturbation experiments with several time-points, from different labs (data from Wang & Wang 2006, Selbach et al. 2008; Guo et al. 2010, and our own reporter assays) to validate our
model. We believe that these experiments provide substantial evidence for the generality of our model.

