Review of Bc. thesis Functional analysis of a putative RNA-editing associated protein 1 (REAP1) in procyclic *T. brucei* written by Zdenka Cicova.

After having read the proposed thesis work performed by Zdenka as a bachelor student of the Laboratory of molecular biology of protists, I am happy to state that, in my view, the work is more than acceptable to grant Zdenka a BSc. degree.

However having stated this, I still have remarks and questions regarding the content of the thesis. I would like to point out that this thesis is written by impressive English and contains only minor mistakes. My remarks and criticism is based on fact that Zdenka would like to publish this story as a short paper in an impact journal.

Introduction:

Page 2: using abbreviations without explanation ("mt" as mitochondrial and "Us" as uridines);

Page 3: I believe that guide RNAs are mostly 70 nucleotides long, not 30.

Methods:

Chapter 2.1.: Why Zdenka did not verify the cloned sequence in p2T7-177 vector by sequencing ?

Fig. 2.7.: Why is it shown that the length of the REAP-1 sequence is 397AA when the proposed sequence by Madison-Antenucci is 418AA long?

Chapter 2.6.1.: Zdenka mentioned presence of 1% glucose in the media used for growing the BL21DE3 cells and the induction. Why was glucose added to media? Moreover I think that glucose has to be washed out before the induction since it might interfere with an IPTG-induced expression of target protein.

Chapter 2.7.1 and 2.7.2.: Two different concentrations of lysozyme were used for the pilot experiment and actual purification, any reason for this?

Chapter 2.7.3: Zdenka mentioned that the concentration of REAP-1 protein for immunizations is 1000ug/ml? How was the protein concentration measured? By Bradford assay or only by comparison with known concentration of BSA protein based on a coomassie stained SDS PAGE gel? Does Zdenka think that the latter method is precise enough?

Is it really necessary to mix homogenized coomassie stained SDS PAGE gel band with Freud's adjuvant for immunization? Why?

It would be helpful to have one figure of REAP-1 sequence which will include all primers used in this study and not 3 separate figures. The reader gets confused easily by which part of the sequence was used for RNAi, which part for the over-expression, which part for the real-time PCR, and where the primers REAp-1 Rv1 and Rv2 are located within the sequence etc.

Results:

Chapter 3.1. Why the signal for dsRNA is located around 1.3kb when the gene fragment used for RNAi was only 876bp long?

Chapter 3.3

How many times did Zdenka repeat the growth curves for RNAi cell line and clones? Is the same growth phenotype reproducible for each cell line?

Chapter 3.5.: Real time PCR

Zdenka used only one gene (beta-tubulin) as a reference gene. I know that this method is very precise and also elaborate but does Zdenka think that one reference gene is enough?

Especially if Zdenka considers that cells are dying after induction of RNAi and concentration of cell tubulin mRNAs might be changed. I would recommend to use one more internal reference. Has Zdenka looked at cells after RNAi induction? Do the cells look normal or has Zdenka noticed any abnormalities?

Table 3.1 second and third (REAP-1 tet- and tet+) columns are switched. What is meant by "PCR efficiency"?

In this chapter (3.5) I am missing a lot of text. Tables are very informative, but Zdenka actually did not explain what the numbers mean and which values she used for analysis and for figure 3.7.

Are these data reproducible?

Chapter 3.6

What is Zdenka's opinion about prediction programs used for localization prediction? Are there any false-negative or false-positive? Any guess?

The difference between GeneDB and M-A (NCBI) sequences are obvious, however Zdenka has not mentioned that it might be because of different strains, which were used for analysis. Strain Treu667 used by Madison-Antenucci and strain 927 used for the genome sequencing and by GeneDB. There is one more discrepancy between sequences and it is 20AA insertion in M-A C-terminal part of protein. It might be interesting to use DNA from strain Treu667 for the same SL-RACE experiment.

Table 3.5. Why does Zdenka expect signal for RP_N in whole cell lysate IPTG minus? What does it mean "obtained RP_N signal"? Coomassie stained band or signal from western? Why does Zdenka expect no signal in "wash 60"?

References:

Beverly SM (2002), Panigrahi et al., (2006), Panigrahi et al., (2001), Pelletier et al., (2000).....etc. They are not in text.

Discussion and conclusions

I would like to point out that Zdenka dealt with a large amount of data and methods. I highly appreciated how Zdenka was able to absorb all available information and identified the discrepancies between her data and published papers. In my view, there are two possible explanations for this discrepancy. The antibody sent by prof. Hajduk was mislabeled and only by coincidence recognizes the same size protein in *T. brucei*. In this case the gene sequence for REAP-1 is right and codes mitochondrial protein possibly involved in RNA editing. However it is also possible that the gene sequence was identified wrong during library screening and in this case Zdenka is dealing with a hypothetical protein which can be localized in any cell compartments and be involved in any cell processes. There are still a few experiments which have to be done before this story is ready for publication. The new antibody against N-terminus of Reap-1 protein will be crucial for protein localization and for verification of RNAi cell lines etc. Also it will be crucial to find out what protein is recognized by Hajduk's monoclonal antibody by immunoprecipitaion followed mass spectrometry analysis. Based on these results, Zdenka will be able to make a decision how to present and publish her data.

At the end, I would like to say that I am impressed by Zdenka`s work, which should receive a grade of 1.

Alena Zikova PhD
SBRI, Seattle