NOV 09, 2022

< SHARE

WORKS FOR ME   1

🌐 Using ebFRET for Hidden Markov Modeling

This protocol is published without a DOI.

Clark Fritsch[1]

[1]University of Pennsylvania

Clark Fritsch
University of Pennsylvania

COMMENTS 0

ABSTRACT

This protocol follows from the "Formatting Data for Use With ebFRET Hidden Markov Modeling Software" protocol and is the third step towards analyzing your single-molecule FRET traces using Hidden Markov Modeling. In this protocol, we take your formatted traces and run them through ebFRET to generate idealized FRET states (Viterbi States) and to obtain the mean values of each of these states (Viterbi Means) after ebFRET has fit your data to a certain number of states.
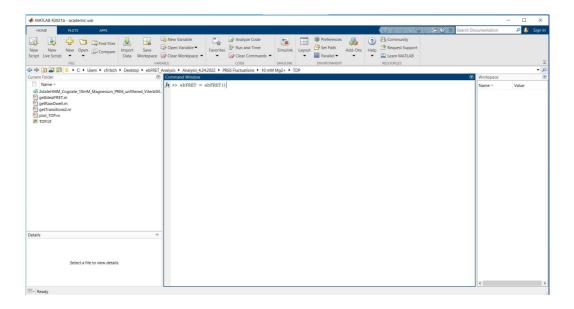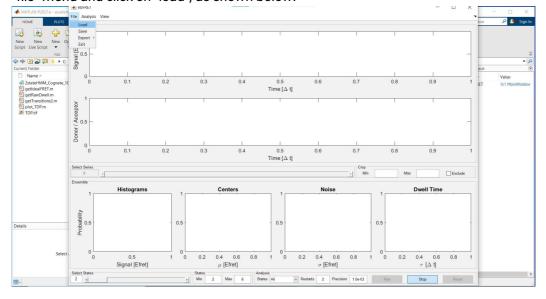
PROTOCOL CITATION

LICENSE

CREATED

Nov 05, 2022

1   After creating your ".dat" file for ebFRET, you need to open up MATLAB and open "ebFRET" by typing "ebFRET = ebFRET()" in the console, as shown below:
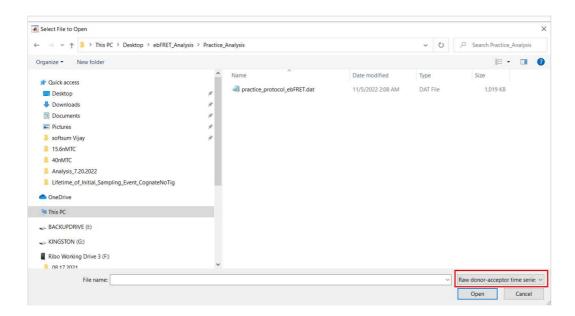

Then click "Enter" and ebFRET will open. Because of the size of the program, it may take some time for it to open depending on the computer that you are using.
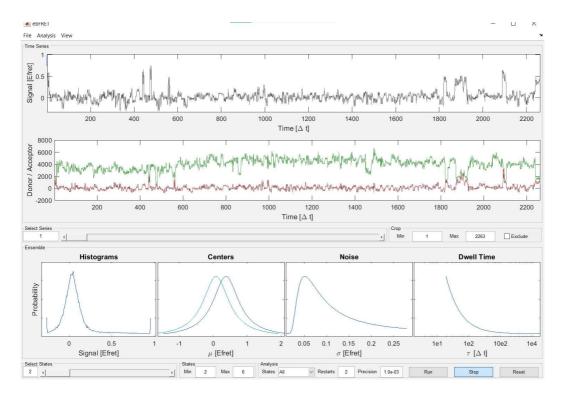
2   Once you have ebFRET open, you will need to load in your ".dat" file. To do this, simply go to the "file" menu and click on "load", as shown below:

3   This will open a navigation window and allow you to navigate to the directory that contains your ".dat" file. Note that you will need to change the type of file that ebFRET is looking for to "Raw Donor-Acceptor time series" (.dat):
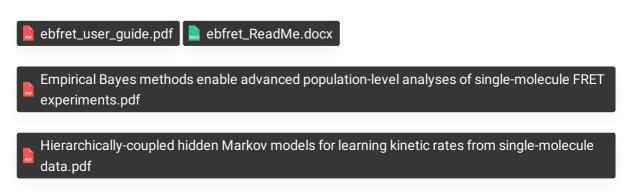

Then open your ".dat" file and you will see your traces loaded into ebFRET:

4   Once you have loaded your traces into ebFRET, you will notice a great deal of information and
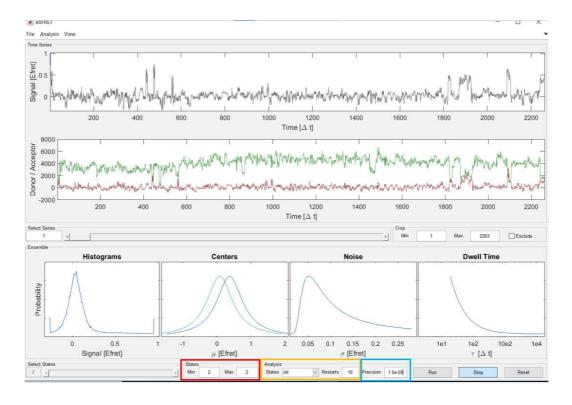
settings become available to you. Due to the number of settings, it is best to refer to the user manual and Read Me file for the ebFRET program that are provided on the ebFRET github website (http://ebfret.github.io/). I have attached several of these files and the original ebFRET papers below:

📄 **ebfret_user_guide.pdf**   📄 **ebfret_ReadMe.docx**

📄 Empirical Bayes methods enable advanced population-level analyses of single-molecule FRET experiments.pdf

📄 Hierarchically-coupled hidden Markov models for learning kinetic rates from single-molecule data.pdf
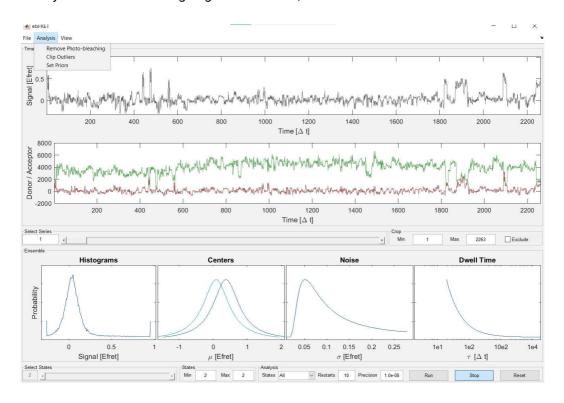
5   For the purposes of this example protocol, I will use the following settings to fit my traces:

Number of States (red box): 2
Restarts (orange box): 10
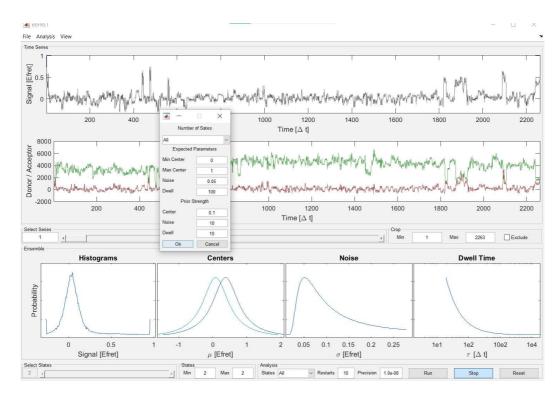Precision (blue box): 1.0e-08

Note that for the number of states, it is best to enter a few more states than what you think is reasonable. However, because I know that there are only 2 states in my data, I will restrict the number of states to 2.

Additionally, you will want to set the priors used for the fitting to default settings by first going to the "Analysis" menu and then going to "Set Priors", as shown below:

Once you click "Set Priors", a menu will appear and you can just click "OK" to set the priors to their default settings:

Once you have done this, you are ready to run the program. You can do this by simply clicking "Run"

in the bottom right hand corner of the program. ebFRET will proceed to fit to fit your traces according to the settings that you entered. Note that you more states that you attempt to fit and the more traces that you have to fit, the longer the analysis time will take. Additionally, the more restarts that you use (if you have 10 restarts, then ebFRET fits your traces 10 times), the longer the fitting will take. If you have over a 100 traces (I would recommend using 100 - 200 traces for your analysis), it may take several hours to finish the fitting.

6 You can tell when the fitting is complete when it becomes possible to click on the "Run" button again. Once your analysis is complete, you should see the following information:

7 The first important information that you can see from you analysis is the different states that were fit to each of your traces, as shown by the red box below:

For each of the traces included in your ".dat" file, ebFRET will attempt to fit up to the number of states that you entered prior to running ebFRET. Remember that for this practice protocol, I entered 2 states into the settings for ebFRET. This means that ebFRET will attempt to fit either 1 or 2 states to each trace that was included in the ".dat" file. If you had entered 4 states during the setup, ebFRET would attempt to fit up to 4 states to each trace, meaning that you would get traces that included either 1, 2, 3 or 4 states in them.

You can scroll through each of the traces that was included in your ".dat" file by using the sliding menu in the ebFRET interface (yellow box).
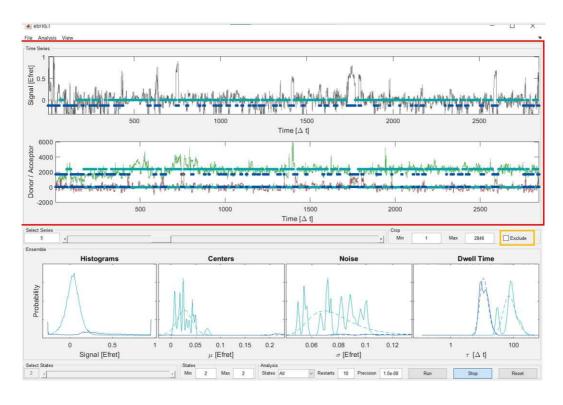
If you did enter more than 2 states for ebFRET to fit to your traces, you can change the number of states that are fit to your traces by using the sliding menu in the ebFRET interface (blue box). If you fit a total of 4 states to your traces, but adjust the sliding bar to "2 states", then you will see only up to 2 states fit to your traces. This can be useful when trying to determine whether additional states are warranted for your data. Although there are "quantitative" ways of determining the number of states that are optimal for you data, at the end of the day you have to confirm whether the states look reasonable or not by eye.

8   From the results of the ebFRET analysis, you will also see that ebFRET has generated a variety of histograms for your data, as shown in the red box below:

These histograms are a good indication for what you data looks like for each state of the Hidden Markov Model and can also give you a good indication of how much noise is present in your data. The histograms for "Dwell Time", as shown above, are clearly noisy and do not fit well to the expected two distributions. This is likely because some of the traces were not fit well by ebFRET.
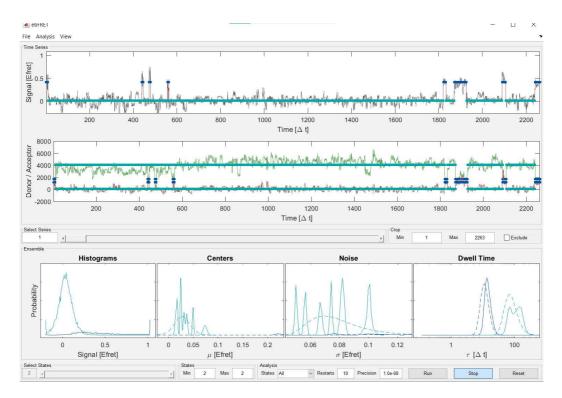
The below trace (red box) was clearly fit poorly by ebFRET and should not be used in future analysis:

If this is the case and the fit does not seem reasonable or is simply too noisy, it is best to exclude the trace. You can exclude any given trace from being exported later on by clicking the "Exclude" box that is in the orange box, as shown above.
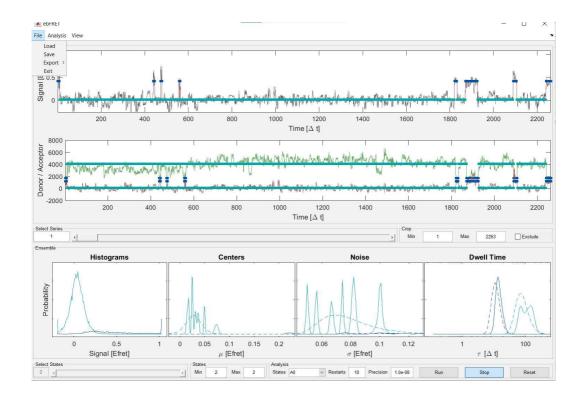
Note, you should only click this box once. If you click the box and then unclick it, for example, you may get a bug that prevents you from exporting your data later on, which would require you to redo your analysis from the beginning to fix.

As you exclude poorly fit traces, you will see the histograms update to reflect the new set of data:
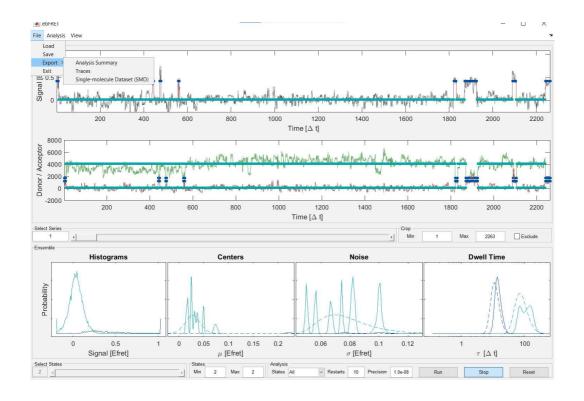

9   Once you are satisfied with the traces and their fits, you can save your ebFRET session so that you can refer to it in the future if necessary. By saving your session, it allows you to refer to your data again without having to redo the fits which, as stated previously, can take a significant amount of time to obtain.

To save you ebFRET session, first go to the "File" menu and then click on "Save" in the menu:
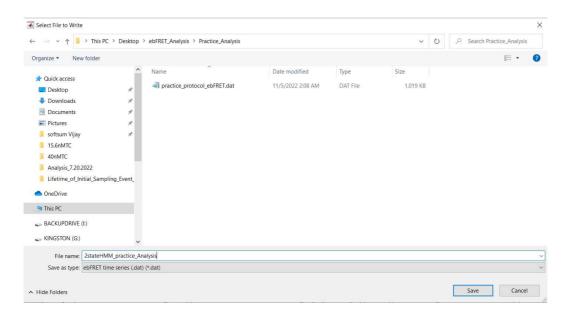
You can then save your ebFRET saved session (.mat) file at the same location that you saved your ".dat" file for convenience.

10   Once you have saved your ebFRET session, you can then begin to export your data / Hidden Markov Model for further analysis. You can do this by first going to the "File" menu, then the "Export" dropdown menu and clicking "Traces". You can also export an "Analysis Summary" that can be helpful for determining the optimal number of FRET States to fit to your data, but it is best to read about using the analysis summary from the user guide provided on the ebFRET github account.
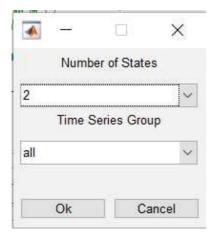
11   Once you click on "Traces", a menu will appear that allows you to save the data that you have generated to a file. I recommend clearly labeling how many states you will be exporting to the data file. If you fit your data to a maximum of 4 states, for example, you will have the option of outputting your data that has been fit to 2 states, 3 states, or 4 states. Since I have selected only 2 states for this practice protocol, I will only output the data that I have fit to 2 states. Because of this, I will name my file "2stateHMM_practice_Analysis.dat":

Once you have clicked "Save", the following window will appear and ask you what information you would like to include in the exported file. For most analyses, I recommend just including everything:
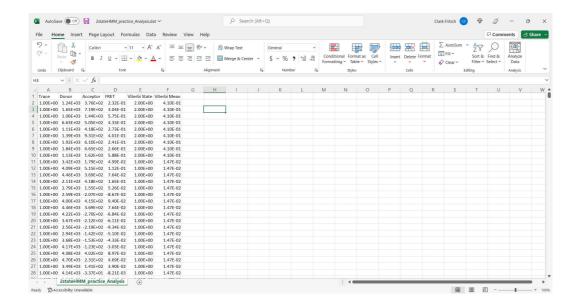
12   Once you click "Ok", another window will appear that asks you which Hidden Markov Model you would like to export to the file. Again, since I only fit a maximum of 2 states to my practice data, I only have the option to select "2 states". However, if You chose to fit your data to a maximum of 4 states, then you will have the option to select "2 states", "3 states" or "4 states" here:

If you choose "3 states" here, then your traces will be fit to a maximum of 3 states and exported. If you selected this option, it would be best if your export file was named something like "3stateHMM_practice_Analysis.dat" to prevent any confusion later on.

13   You can then open this "2stateHMM_practice_Analysis.dat" file and see the following:

Note that ordinarily, the columns don't have column headers on them. However, I have added headers here to show what each column is for convenience. The columns represent the following information:

Trace: The "Trace" column contains the unique ID for each of the traces that you analyzed using ebFRET. Note that if you excluded traces during your analysis, they will not be included in this export file.

Donor: The "Donor" column contains the raw intensity values for your donor fluorophore. This data is the same as the data that you originally included in your ".dat" file.

Acceptor: The "Acceptor" column contains the raw intensity values for your acceptor fluorophore / sensitized emission values. This data is the same as the data that you originally included in your ".dat" file.

FRET: The "FRET" column contains that FRET efficiencies calculated from the raw donor and acceptor intensity values that you provided to ebFRET for analysis.

Viterbi State: The "Viterbi State" column contains information regarding which state that ebFRET fit to a given FRET efficiency. If you fit a maximum of 2 FRET states to your data, then this value could be either "1" or "2".

Viterbi Mean: The "Viterbi Mean" is the idealized FRET efficiency that ebFRET determined based on the given state the the data was fit to.