Guidelines for Acquiring and Analyzing Stopped-Flow Data

1. **Data Collection:** When data is acquired by the stopped-flow you can collect either in **LINEAR sampling mode** or **LOGARITHM sampling mode**. Linear sampling is good for short time courses when the data is collected for less than 30 seconds. You have good distribution of the 8000 data points acquired by the instrument. Logarithm sampling mode is preferred when you plan to collect data for long time courses (over 30 seconds) and there are fast and slow phases (multiple phases). Acquiring data by logarithm sampling allows you to collect more data points for the early fast phase. This helps in data analysis by curve fitting. BioKine 32 will let you collect data over different time windows. For example, you can setup the program to collect 4000 data points from 0 to 10 sec and then 4000 data points from 10 sec to 50 sec. Therefore, plan ahead which mode to collect the data for subsequent curve fitting.

2. **Data Analysis Software:** There are several software packages for curve fitting by non-linear regression such as Prism, BioKine 32, Origin, MatLab, etc. They use the Levenberg-Marquardt method for finding the best-fit values, except BioKine 32 which uses a Simplex method. MatLab requires programming and is an overkill. Prism should be used for most curve fit analysis because it does an excellent job and provides statistical evaluation of the fitted curve. The Prism Manual and online help guide is very good to learn about curve fitting. I also recommend using BioKine 32 for getting Initial Values for fitting with Prism. However, BioKine 32 will not provide standard error or confidence interval for the fitted parameters. It is a good idea to analyze the data using both the Simplex and LM methods to avoid any potential fitting problems due to "local minima" in a program.

3. **Curve Fitting:** Import data tables into Prism for analysis.
   1.1. Each shot is not an independent experiment but should be treated as a replicate of a single experiment because they are coming from the same reaction tube. Therefore, traces from multiple shots should be averaged before analysis. Data from 5 to 6 shots that are similar can be averaged within BioKine 32 and the averaged data can be analyzed by Prism. As a precaution, it is also a good idea to analyze few traces from individual shots separately to make sure that best-fit values do not change due to averaging.
   1.2. Use the **Dead time** of the instrument to determine the first data point in the curve (usually this is about 2 to 3 millisecond depending on flow rate, etc. Note this down when you perform the stopped-flow experiment). Data points earlier than the dead time are removed from the analysis.
   1.3. **Transform** the data by dividing each Y value with k. This can be accomplished with Prism which has a menu called TRANSFORM. **Select equation Y = Y/k.** k is the Y-value for the first real data point. The transformed data will begin with Y =1 and is in a separate table in Prism. Do not change the X-value (time). Transforming the data is better than normalizing the data from 100% to 0% because the extent of the fluorescence change in the transformed data is identical to the

raw data. This is important for comparing the extent of fluorescence change in different experiments.

1.4. Selecting the cut-off point for analysis is sometimes tricky. A good rule is that the reaction will be 98% complete over six half-life. Therefore, select data points that cover at least six half-life.

Estimate the half-life **($t_{1/2}$) = 0.693/k_{obs}**
Example, if k_{obs} is 10 s^{-1}, then $t_{1/2}$ = 0.693/10 = 0.0693. Six half-life = 0.0693 x 6 = **0.4 sec**

1.5. Two phases are observed during translocation kinetics of wild type ribosomes: A **fast phase (k_{obs} = 10-20 s^{-1})** that correspond to mRNA-tRNA movement and a **slow phase (k_{obs} =0.5-2 s^{-1})**, the origin of which is unclear at the present time. Applying the six half-life rule, the fast phase will be complete in 0.4 seconds and the slow phase will be complete in 4 seconds. Generally, analyzing the translocation data up to 5 seconds gives good fits for the fast translocation phase. In some instances, analyzing the data to 2 seconds may give better fits than analyzing to 5 seconds. This needs to be empirically determined. However, make sure that wild type and mutants are analyzed identically so that relative effects on translocation rates can be compared.

1.6. Fitting translocation time course to more than 5 seconds generally give poor fits for the fast phase. There are several explanations for this: (1) The program, in trying to minimize the R^2 value, gives more weight to the later time points which contains more data points. This results in poor fits for the early phase. One way to avoid this is to collect the data in **LOGARITHM sampling mode** so there is sufficient data points for the early phase. (2) The amplitude of the early phase may become smaller than the amplitude of the late phase when data is analyzed for longer time courses. The program ignores the early phase due to its small amplitude and gives more importance to the later time points. (3) There may be additional processes going on at long time interval that cannot be correctly fit with a double exponential equation. However, we mainly care for the biologically relevant early phase so we should try to get the best fit values for this phase by analyzing shorter time courses.

1.7. For analyzing translocation kinetics use the **Biokine Double Exponential Equation:**

a*x+b + c1*exp(-k1*x) + c2*exp(-k2*x). Enter this as a custom equation in Prism. There is a slow linear slope due to instrument drift in the translocation kinetics which can be approximated by a*x+b in the equation. This is only important if the data is going to be fitted for really long times (greater than 10 seconds). You can use values obtained from BioKine 32 as starting Initial Values for fitting with Prism. The standard double exponential equation will also work for short time courses where the linear component is negligible.

1.8. **How good is the curve fit?** Inspect the fitted curves carefully by magnifying it to a large size on the computer screen. Check the **residual plots** to make sure that they are evenly distributed. Make sure the R^2 value is close to 1 (usually around 0.99 – 0.98 for good fits). Most curve fits

will report reasonable R^2 values, therefore it is NOT very useful. Instead, determine the **Standard Errors** and the **95% Confidence Intervals** with Prism. Especially important are the **95% Confidence Intervals** which will tell how precise the best-fit values are for the fitted curve. If it is very wide, the fit is poor. Finally, compare the best-fit values from **Independent experiments**.

1.9. In some cases, the mutant ribosome may translocate very slowly such that the fast phase cannot be distinguished from the slow phase. In this case it is possible to curve fit using the **Biokine Single Exponential Equation: a*x+b + c1*exp(-k1*x)**. Prism will let you compare two equations side-by-side. Use the single exponential equation if it gives as good a fit as the double exponential equation. Also verify the extent of translocation using alternative methods such as toeprinting or puromycin reaction. This will tell whether the mutant ribosomes are able to translocate over long time points and is important for the interpretation of the kinetic data.