Processing hydrolysis rate data in R

Drew Steen

May 28, 2014

Setup

To get started, you'll need to install R (cran.r-project.org) and R's Integrated Development Environment (IDE), RStudio (rstudio.com). R does all the actual calculations you need, but you will only interact with RStudio - you don't need a shortcut for R on your desktop, since you'll never open it directly.

Next, set up a folder on your hard drive for this project (say, `C:/documents/enzymeResearch`). Then, open up RStudio and make a new project: Choose `File, New Project, Existing Directory` and browse to your directory.

Next, set up the following new folders in your project directory: `R`, `plots`, `data`, and `reports`. (Be sure to make the capitalization identical to mine so that our scripts will work on each others' systems.) Save the file I've sent you, `lm_stats.R`, to your `R` directory.

Finally, install a few useful packages. In RStudio, go to the packages tab (lower left-hand window). Click "Install Packages" and type `ggplot2` in the window, and click "Install". Do the same for `reshape2`, `plyr`, and `lubridate`. You will only need to install each package once, but you will need to load them for each new R session. You can also install packages by typing the command `install.packages("ggplot2")` in RStudio's console window.

Prepare your data file

Eventually, I plan to write functions that can deal with the plate reader output files directly. For now, open the .txt or Excel file from the plate reader, copy the section containing the raw data (usually starting at row 40 or so, be sure to include the row headers) into a new spreadsheet. The new spreadsheet should have column names in row A, data in rows B and further, and the data should start in column 1. Trim out any columns that don't contain useful data.

Next, change the column names for time and temperature to `time` and `temp`. Finally, save the file as a .csv file.

1 Data processing

Run the following commands in R. The best way to