Capstone Project Proposal
James Thurmond
November 8, 2005

BioFitWeb: A Comprehensive On-Line Resource for Enzyme Kinetics Researchers

1. Introduction

BioFitWeb is to be a comprehensive, on-line set of tools and resources offered free of charge for use by researchers studying enzyme kinetics. The centerpiece of BioFitWeb will be a set of easy-to-use yet powerful and accurate mathematics tools that will analyze enzyme kinetic data and determine the parameters of greatest interest to these researchers, such as the maximum enzyme reaction velocity $\mathbf{V_{max}}$, and the Michaelis constant $\mathbf{K_m}$, as well as perform a number of tests and fittings. In addition, BioFitWeb will provide decision tools to help researchers understand the behavior of the enzymes they are working with, and decide on the best model within which to make further inferences. A collection of contextual links to scholarly resources will provide a library of information to enhance the guidance provided on the site. These will include recent papers by the developers of the site, including some as-yet-unpublished original research on the mathematics behind some of the tools. Also planned are several novel tools which will be unique to BioFitWeb, such as one to predict biochemical pathways using multi-species time-course reaction data.

2. Background

There have been a number of recent developments in the field of enzymology which create a need for a resource like BioFitWeb. Some of the methods which have been used for decades have recently been criticized for previously little-appreciated shortcomings [Ritchie & Prvan, 1996]. More robust methods are being developed by theoreticians, and the limitations within which the traditional methods can be used safely are becoming better understood [Schnell & Maini, 2003], although often not by laboratory researchers. New methods for direct fitting of time-course data have been developed [Schnell & Mendoza 2001] that provide accurate results using data from less-expensive experiments than have previously been possible. These methods have not yet been implemented in any other package.

As genomics and proteomics mature and the generation of data in these areas becomes more and more automated, some believe that demand will increase for kinetic data on enzymes [Gallagher, 2004]. A quantitative understanding of the catalytic behavior of enzymes is a necessary component for good modeling and prediction. Eventually the rate at which this data is needed may require that many of the experiments carried out painstakingly by hand today will be automated as have their predecessors. If that happens, there will be an acute need for quality, easy-to-use tools to analyze the data generated. The methods being developed for BioFitWeb are steps in the direction of these future tools.

A survey of software currently available to address these needs is revealing. The options now available are either expensive, limited in scope, difficult to use, or restricted to one or two operating systems. Many are also add-ons to other, more general applications, adding expense and decreasing specificity. These programs all require expertise and experience to use properly, as well. They are designed to be used by scientists with strong academic backgrounds. BioFitWeb will have an advantage over these tools with its integrated access to papers and other resources to help researchers with the decisions that go into choosing a model.

To summarize, what will make BioFitWeb unique is a combination of features:

- BioFitWeb is **free**.
- BioFitWeb is completely **platform-independent** because it is available to anyone with access to the internet.
- BioFitWeb will implement sophisticated algorithms and features available nowhere else, even in commercial software.
- BioFitWeb will be comprehensive, integrating a set of tools and resources more extensive than anything currently available.
- BioFitWeb will incorporate the latest academic resources and an embedded system of decision tools to assist researchers in making the right choices during analysis.

3. Methods and Goals

BioFitWeb is an internet-based, free resource consisting of documents, scripts, and software. The site consists of web pages, which guide access to the software and resource documents.

The underlying functionality is a set of software tools written in fortran. This software implements the mathematical algorithms for fitting data to various models. Already implemented tools include:

- initial reaction velocity vs. initial substrate concentration Michaelis-Menten equation fitting to determine enzyme $\mathbf{V_{max}}$ and $\mathbf{K_m}$
- reaction time-course (substrate decay or product formation vs. time) analysis to determine $\mathbf{V_{max}}$, $\mathbf{K_m}$, initial reaction rate, and initial substrate concentration, if unknown

The time-course analysis uses new and powerful techniques recently developed by Schnell and Mendoza.

Linking the web resources with the software tools is accomplished using CGI scripts written in perl. These handle the uploading and storing of data, connect the user to the appropriate tool, provide for security and availability of each user's session, and assist with the output of the fortran code.

Another unique feature of BioFitWeb will be a set of decision tools designed to help researchers to interpret data. A hypertext flowchart structure will be embedded in the site, directing users to testing tools and advising them on interpretation and further analysis until a

sound conclusion can be reached and a well-founded model for their data can be applied. The tests involved are, for the most part, well-known to enzymologists, but frequently mis-applied or not used when they should be [Ritchie & Prvan, 1996]. The structure of the BioFitWeb site will place links to the appropriate instruments and tests at the relevant points in the chain of analysis to advise users on the circumstances under which they may wish to consult them.

An on-line library of reference material, including seminal as well as very recent peer-reviewed work, is being compiled to support BioFitWeb. Context-appropriate links in the site will invite users to consult specific papers to deepen their knowledge about the theory and history behind the methods they will employ. A significant number of the papers to be included in this library will have been authored by contributors to the BioFitWeb effort, and more relevant papers are currently in development here.

The resource is currently running on a temporary server using a Linux operating system and Apache server software. The site is available now at http://156.56.92.83/BioFitWeb.htm (while still under construction).

4. Scope

BioFitWeb is a major undertaking, and will require contributions from a number of people before it is complete. My task is to lay a solid foundation for future work and to implement the most frequently-used basic tools. I will be responsible for the following tasks:

- the initial layout of the web site and its software components
- writing background and reference information to support the tools
- design of the embedded decision-aid structure
- implementing tools to estimate kinetic parameters modified to account for substrate inhibition effects
- implementing tools to perform tests such as Selwyn and discriminatory curve plots, as aids to determining the proper general enzyme kinetic model appropriate for a given data set
- implementing a modeling tool which will predict likely reaction mechanisms and pathways using time-course data for multiple reaction substrates, products, and intermediates
- implementing a system to display the bibliographic database on the web site

References:

Gallagher, R. 2004. Enzymes make the world go ‘round. (editorial) *Scientist* 18.1:6.

Ritchie, R. J., and T. Prvan. 1996. A simulation study on designing experiments to measure the K_m of the Michaelis-Menten kinetics curves. *J. Theor. Biol.* 178:239-254.

Schnell, S., and P. K. Maini. 2003. A century of enzyme kinetics: reliability of the K_m and v_{max} estimates. *Comm. Theor. Biol.* 8:169-187.

Schnell, S., and C. Mendoza. 2001. A fast method to estimate kinetic constant for enzyme inhibitors. *Acta Biotheor.* 49:109-113.