

# Quantitative Understanding in Biology

## Lab #4: Model comparisons

6 December, 2016

Consider the following data for an oxygen transport compound similar to myoglobin and hemoglobin:

| $P_{O_2}$ (mm Hg) | Percent saturation |
|-------------------|--------------------|
| 2.01              | 2.42               |
| 6.98              | 11.37              |
| 12.09             | 24.33              |
| 17.03             | 31.03              |
| 22.01             | 48.57              |
| 27.06             | 49.41              |
| 32.06             | 56.66              |
| 36.91             | 64.29              |
| 42.08             | 72.36              |
| 46.99             | 69.94              |
| 52.05             | 77.94              |
| 56.92             | 76.28              |

**Plot these data.**

**To your eye, does it appear that this system demonstrates cooperativity?**

*n.b., there is no wrong answer to this question, but explain your reasoning.*

**Fit the data to a Michelis-Menten model, prepare appropriate plots, and compute the 95% CIs for the model parameters. Do you think this is a good model for the data?**

**Fit the data to a Hill model, and again prepare appropriate plots and compute 95% CIs.**

**Using an F-test to compare the two models, demonstrate that the Hill model explains the data better than the simpler Michelis-Menten model (using an  $\alpha$**

cutoff of 0.05).

Repeat all of the above steps using a subset of the original data, as shown below. In this scenario, you should be able to show that the more complex Hill model is no longer justified.

| $P_{O_2}$ (mm Hg) | Percent saturation |
|-------------------|--------------------|
| 2.01              | 2.42               |
| 12.09             | 24.33              |
| 22.01             | 48.57              |
| 32.06             | 56.66              |
| 42.08             | 72.36              |
| 52.05             | 77.94              |

Note that you may find that R has a hard time estimating CIs for some of your models. Switching the algorithm used by `nls()` when you fit the model can sometimes help. **Why do you think it is difficult to estimate CIs in some cases, and what would your interpretation be?**

From this exercise, you should appreciate the interplay between the amount of data you collect and your ability to detect and confirm subtle behaviors of the system that you are studying. As you might imagine, the noisier your measurements, the more data you'd need to collect in order to reach statistical significance. Conversely, the stronger the signal, the less data you'd need (e.g., for hemoglobin, which has a Hill exponent between 2.5 and 3.0, six or so data points would probably be enough to “prove” cooperativity).