Molecular Mechanisms for Human Diseases: Lab 1

Bacteria’s Brain – Part A

Peter Yibo Pan

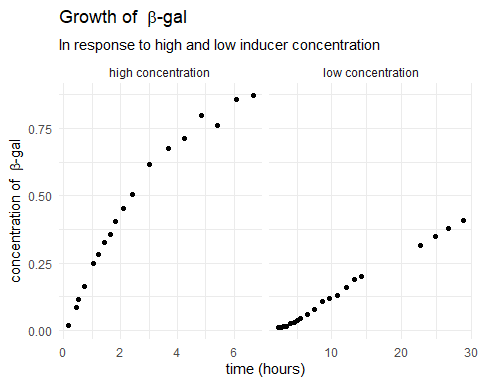
(2023-02-05)

##Results:

## # A tibble: 18 × 2  
## t concentration  
## <dbl> <dbl>  
## 1 0.170 0.019   
## 2 0.443 0.0855  
## 3 0.511 0.116   
## 4 0.716 0.164   
## 5 1.06 0.247   
## 6 1.20 0.280   
## 7 1.43 0.328   
## 8 1.65 0.356   
## 9 1.83 0.404   
## 10 2.11 0.454   
## 11 2.42 0.506   
## 12 3.00 0.615   
## 13 3.68 0.675   
## 14 4.23 0.713   
## 15 4.83 0.798   
## 16 5.41 0.760   
## 17 6.06 0.860   
## 18 6.66 0.874

## # A tibble: 20 × 2  
## t concentration  
## <dbl> <dbl>  
## 1 2.38 0.0109   
## 2 2.72 0.00934  
## 3 3.04 0.0146   
## 4 3.52 0.0145   
## 5 4.00 0.0242   
## 6 4.57 0.0286   
## 7 5.05 0.0376   
## 8 5.54 0.0435   
## 9 6.45 0.0592   
## 10 7.53 0.0764   
## 11 8.58 0.107   
## 12 9.63 0.118   
## 13 10.8 0.128   
## 14 12.0 0.16   
## 15 13.2 0.188   
## 16 14.3 0.199   
## 17 22.7 0.317   
## 18 24.9 0.349   
## 19 26.8 0.378   
## 20 28.9 0.408

Visualization of Data set A and B



## Analysis:

### Derivation of Functions used for fitting both datasets:

I will first define the different terms that will be used.

is the number of individual E.coli bacteria in the chemostat, which is maintained at a constant value.

is the flow rate both in- and outward from the chemostat, which is kept at a constant value so that can be kept constant.

is the volume of the chemostat

we define as the average rate of production by any individual bacterium at a given time. Note this is not the rate of production by individual bacterium but the mean of the population. Hence this value would change based on how many bacterium are “activated” within the population.

is defined as the concentration of -gal within the system at any moment. With these terms defined, we observe that at any moment, the growth of -gal is , which equates two terms: the rate of production by the bacteria and the rate of loss due to flow rate. Hence we get the equation:

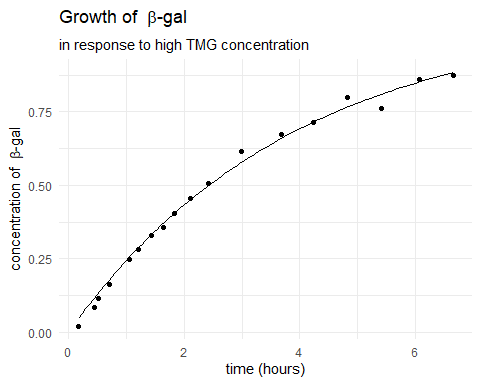
#### A: High Concentration of Inducers

For data set A, when the inducers are introduced at high concentration, we assume that is constant as all bacteria rapidly “switches on” and generates enzyme at its maximal rate. We can then express as a dimensionless variable as , hence:

#### B: Low concentration of Inducers

For data set B, when the inducers are introduced at low concentration, only a fraction of bacteria is activated, with this fraction slowly growing. Hence we assume that is a linear function of with and being the rate at which activation “spreads” through the population.

### Fitting function on Data set A

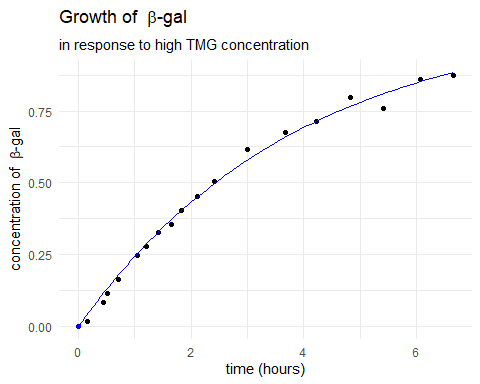


With the coefficients calculated, the nonlinear fitting of A results in the function:

With A = 1.08 and = 3.91 We observe that at , , which is valid in its physical meaning, as there would be no -gal before any inducers are introduced. Hence, the data point is included into the data and the same model is fitted again.

## ampl tau  
## . 1.07972 3.907721  
## coef\_a3 1.07972 3.907721

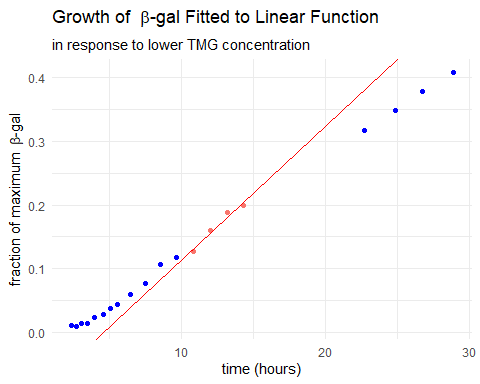
##   
## Formula: a\_q3\_df$concentration ~ growth\_a(t, ampl, tau)  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## ampl 1.07972 0.03931 27.47 1.59e-15 \*\*\*  
## tau 3.90772 0.24889 15.70 1.50e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.02143 on 17 degrees of freedom  
##   
## Number of iterations to convergence: 5   
## Achieved convergence tolerance: 1.242e-06



We see that here the two fitting parameters are identical before and after is included.

Fitting function on Data set B

## (Intercept) b\_q2$t   
## -0.09474840 0.02094596

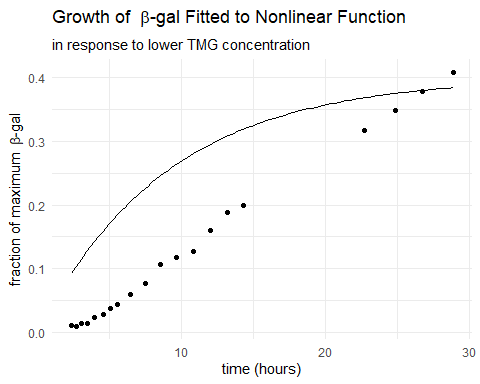


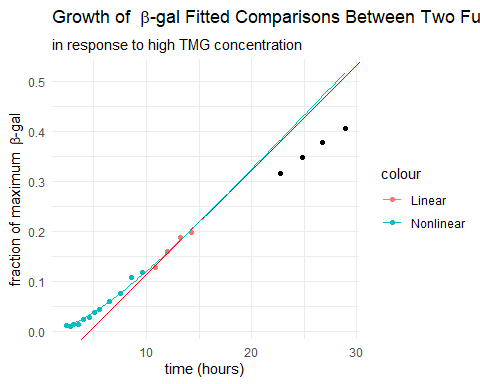
##   
## Call:  
## lm(formula = b\_q2$concentration ~ b\_q2$t)  
##   
## Residuals:  
## 1 2 3 4   
## -0.004175 0.002961 0.007005 -0.005791   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.094748 0.036016 -2.631 0.119   
## b\_q2$t 0.020946 0.002847 7.356 0.018 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.007376 on 2 degrees of freedom  
## Multiple R-squared: 0.9644, Adjusted R-squared: 0.9465   
## F-statistic: 54.11 on 1 and 2 DF, p-value: 0.01798

## Warning: 'newdata' had 20 rows but variables found have 4 rows

## # A tibble: 2 × 2  
## MSE data   
## <dbl> <chr>   
## 1 0.0000272 10<= t <= 15   
## 2 0.0176 entire dataset

As we can see, the Mean Standard Error (MSE) for the model fits much better within the range of 10<= t <= 15 with a MSE of 2.7199507^{-5}, but not well over the entire data set with a MSE of 0.0175584.





##   
## Formula: b\_q4\_df$concentration ~ growth\_b4(t, beta, tau)  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## beta 0.11755 0.02959 3.973 0.00139 \*\*   
## tau 5.34585 0.89590 5.967 3.44e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.005333 on 14 degrees of freedom  
##   
## Number of iterations to convergence: 6   
## Achieved convergence tolerance: 2.284e-06

Q5: at large t, the function becomes linear as -t/tau becomes 0 so exp(-t/tau) becomes a constant of 1