McMaster iGEM Team 2: Mathematical Modelling

Mathematical modeling is an integral component of the practice of synthetic biology: permitting the simulation of the behaviour of genetic circuits in biological systems. Mathematics could be harnessed to inform laboratory projects through predicting the characteristics and response of parts and systems *a priori* as well as to explain experimentally-derived data *post hoc.* Thus, our team developed a mathematical framework to predict the behaviour of elements of our design.

Models were developed for two main purposes:

1. Predicting the effects of changing extracellular conditions on promoter activity through a series of kinetic equations
2. Simulating the stability of residues within the acyl homo-L-serine (AHL) synthase protein structure in response to the binding of its substrate S-adenosyl-L-methionine (SAM) to elucidate the putative residues that participate in its active site

In the first model, we applied ordinary differential equations (ODEs) for transcription and translation with previously-reported parameters in the literature to predict the effect of extracellular pH on gadA promoter activity. In the second model, we utilized molecular dynamics simulations for SAM-binding determination of AHL synthase *in silico*.

Reaction Kinetics at the gadA pH-responsive Promoter

To predict the responsiveness of the pH promoter to changing extracellular pH within a heterogenous tumour microenvironment, we developed kinetic models to simulate the activity of the gadA promoter in the bacterial cell.

Method

In order to simulate the many mechanistic components of the pH-sensitive gadA promoter, we applied dissociation constants found in literature to a series of ODEs.

The behaviour of the gadA promoter was predicted using the framework of the basic model of bacterial transcription and translation at a promoter as suggested by the central dogma.

RNA Polymerase Binding:

$$DNA + RNAP \leftrightarrow DNA - RNAP$$

$$\frac{d[DNA-RNAP]}{dt} = k_{aRNAP-DNA}[DNA][RNAP] - k_{dRNAP-DNA}[DNA - RNAP]$$

RNA Polymerase Activation:

$$DNA - RNAP \rightarrow DNA - RNAP_a$$

$$\frac{d[DNA-RNAP_a]}{dt} = k_a[DNA - RNAP]$$

mRNA Transcription:

$$DNA - RNAP_a \rightarrow DNA + RNAP + mRNA$$

$$\frac{d[mRNA]}{dt} = k_{TX}[DNA - RNAP_a]$$

mRNA degradation:
$mRNA \rightarrow \emptyset$
$$\frac{d[mRNA]}{dt} = -k_{mRNAdeg}[mRNA]$$

Translation as modelled through Michaelis-Menten kinetics:
$mRNA + Ribosome \leftrightarrow mRNA + Ribosome + Protein$
$$\frac{d[Protein]}{dt} = \frac{k_{TL}[Ribosome][mRNA]}{k_M + [mRNA]}$$

The gadA pH-responsive element can be understood most simply by the following transduction pathway as described by Foster (2004):

1. Detection from the membrane-bound sensor kinase EvgS
2. Activation of the gadE transcription factor
3. gadE binding to the gadA promoter element
4. gadE-induced transcription at the gadA locus

Detection of acidic protons with the membrane-bound sensor kinase EvgS
$H^+ + EvgS \leftrightarrow EvgS_a$

Activation of the gadE transcription factor
$EvgS_a + gadE \rightarrow gadE_a$

These aforementioned processes can be modelled through least-squares curve-fitting on a Lorentzian curve with previously reported data associating pH with transcriptional activity through EvgS Histidine Kinase activation (Eguchi, Utsumi, 2014). As Eguchi and Utsumi reported their data in Miller units, the curve was transformed as a measure of gadE activation through assuming $Miller\ units \propto gadE\ activation$, with a ceiling of $gadE_{init} = 6.1\ x\ 10^{-10}\ M$.

gadE binding
$gadE_a + DNA \rightarrow DNA - gadE_a$

gadE-induced transcription at the gadA locus
$DNA - gadE_a \rightarrow mRNA + DNA + gadE$

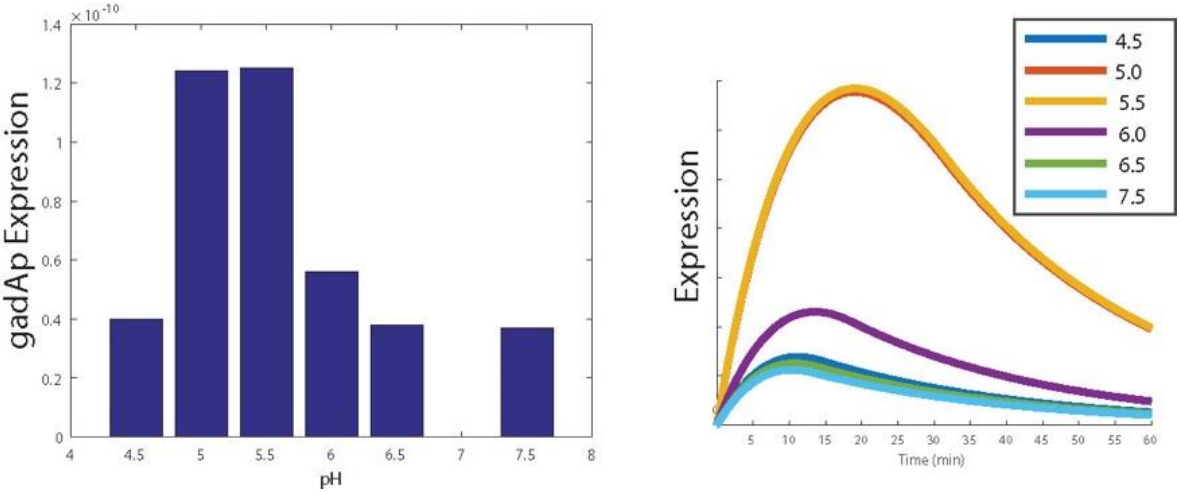
Parameter values as reported by literature:

Variable	Description	Estimated Value	Reference
$k_{aRNAP-DNA}$	RNA Polymerase Association constant with DNA	$5.7\ x\ 10^6\ M^{-1}\ s^{-1}$	Bertrand-Burggraf et al., 1987

$k_{dRNAP-DNA}$	RNA Polymerase Dissociation constant with DNA	$10\ s^{-1}$	Kierzek, Zaim & Zielenkiewicz, 2001
k_a	Closed complex isomerization	$10.5\ x\ 10^{-2}\ s^{-1}$	Bertrand-Burggraf et al., 1987
k_{TX}	Transcription kinetic constant	$1/300\ s^{-1}$	Karzbrun et al., 2011
$k_{mRNAdeg}$	mRNA degradation kinetic constant	$0.3\ s^{-1}$	Kierzek, Zaim & Zielenkiewicz, 2001
k_{TL}	Translation kinetic constant	$4/65\ s^{-1}$	Karzbrun et al., 2011
k_{dgadE}	gadE dissociation constant	6 μM	Krin, Danchin & Soutourina, 2010
$gadE_{init}$	Initial concentration of gadE transcription factor	$6.1\ x\ 10^{-10}\ M = \frac{370\ \frac{gadE}{E-coli\ Cell}}{\frac{6.02\ x\ 10^{23} gadE/mol}{\frac{1\ L}{1\ dm^3} * \frac{1\ um^3}{E-coli\ Cell} * \frac{1\ dm^3}{10^{12}\ um^3}}}$	Ishihama et al., 2014 – estimation used Dan TF as a surrogate measure (part of the LysR TF family)
DNA_{init}	Initial concentration of DNA	$26\ x\ 10^{-6}\ M = \frac{0.017\ \frac{pg\ DNA}{E-coli\ Cell}}{\frac{650\ g/mol}{\frac{1\ L}{1\ dm^3} * \frac{1\ um^3}{E-coli\ Cell} * \frac{1\ dm^3}{10^{12}\ um^3}}}$	ThermoFisher Scientific, nd
$RNAP_{init}$	Initial concentration of RNA polymerase	$2.5\ x\ 10^{-6}\ M$	Shepherd, Dennis, & Bremer, 2001
$Ribosome_{init}$	Initial concentration of cytoplasmic ribosomes	$50\ x\ 10^{-9}$ M $= \frac{3\ x\ 10^4\ \frac{Ribosomes}{E-coli\ Cell}}{\frac{6.02\ x\ 10^{23} ribosomes/mol}{\frac{1\ L}{1\ dm^3} * \frac{1\ um^3}{E-coli\ Cell} * \frac{1\ dm^3}{10^{12}\ um^3}}}$	ThermoFisher Scientific, nd

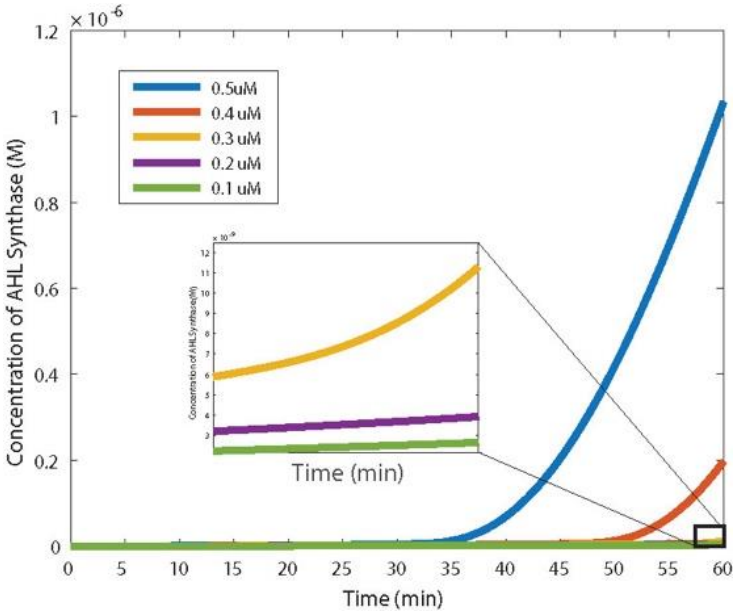
Modelling gadA Activity in Response to a pH Series

Utilizing our previously described kinetic ODEs, we modelled the differential gene expression at the gadA promoter changes in response to changing extracellular pH. The following are the transformed experimental data from Eguchi & Utsumi (2014; left) utilized for curve-fitting and the simulated results (right):



Modelling AHL Synthase Concentration in Response to Changing extracellular [H+]

As our genetic circuit placed AHL synthase downstream of the gadA promoter to initiate quorum sensing, we modelled how the intracellular concentration of AHL synthase would be altered due to altered activity of the gadA promoter to a series of changing extracellular pH (as indicated by [H+]). We thus decided to parse the behaviour of AHL synthase expression in a [H+] series. following curves are the simulated results:



Our model predicts that 0.5 uM of [H+] elicited the strongest induction of AHL synthase expression. The results of our model are in accordance with previously-determined experimental data characterizing the gadA promoter (link Dundee 2016 wiki page: http://2016.igem.org/Team:Dundee/Result). The results of this kinetic model affirm that the optimal activity of the gadA promoter lies between pH 5-5.5. This is concordant with the pH range of the tumour microenvironment as characterized by imaging studies (Chen & Pagel, 2015). For example, Castelli and colleagues (2014) determined that the *in vivo* extracellular pH of murine melanoma ranged from 5.2–6.4. This suggests that the gadA promoter can be applied to our project, which aims to develop a self-limiting tumour-killing genetic circuit responsive to an acidic tumour microenvironment.

Molecular Dynamics Simulations

Using the GROMOS 54a7 force-field with the GROMACS MD package on MacSim, a GPU-accelerated workstation in the Origins of Life Lab at McMaster University, we were able to run replicate all-atom molecular dynamics (MD) simulations. The acyl homo-L-serine synthase enzyme was taken from the PDB database (1R05) and fit in a rectangular box with ~12000 water molecules with a periodic boundary condition applied in all directions. The systems were energy minimized, and equilibrated with the NVT/NPT ensemble prior to running 20 production simulations of 5 ns each. For systems with the SAM ligand to the AHL synthase, umbrella sampling was used to find it's most preferred state in the bilayer. Then the molecule was placed, using VMD, in the water layers surrounding the protein Monte-Carlo. 20 production simulations for 10 ns were run.

The first term is a Lennard-Jones (LJ) potential, which approximates the typical interaction between pairs of neutral atoms and molecules. The second term is based on Coulomb's law and holds the electrostatics between. The last term holds the dihedrals of angle restraints. To define further the interactions between these atoms in the system, we must draw vector fields around each atom from its position in space using quantum mechanics. We utilized the CHARMM36 forcefield (Mackerell et al. 2004) to define all our atoms and residues which contained the equations defining these fields.

$$V(r_1 \ldots r_n) = \sum_{ij} \varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \sum_{ij} \frac{1}{4\pi\varepsilon_0} \frac{q_i \cdot q_j}{r_{ij}^2}$$

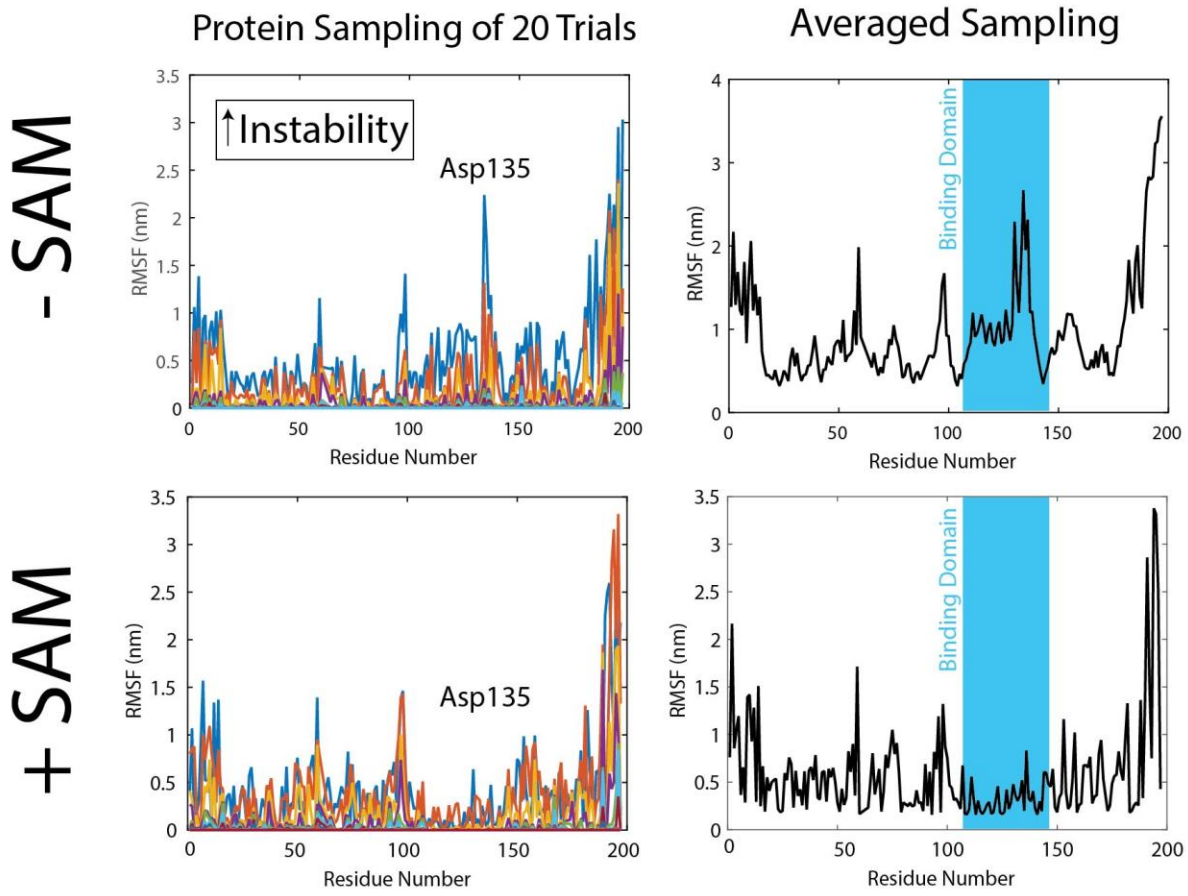
$$+ \sum_{dihedral} K_\phi [1 + cos(n\phi - \delta)] + \sum_{bonds} \frac{1}{2} K_d (d - d_0) + \sum_{angles} \frac{1}{2} K_\theta (\theta - \theta_0)$$

The root-mean square fluctuations (RMSF) of each residue was calculated as a measure for the residue instability, by the following equation:

$$RMSF = \sqrt{\frac{1}{T}\sum\left(x_i(t_j) - x_j\right)^2}$$

Where $x_i(t_j)$ is the position of a molecule at some time t, and x_j is the initial position. By calculating the sum of the squared differences, we calculate the fluctuations of these atoms.

From doing so, we sampled the protein and created an average plot of the RMSF. We find that the Aspartate at position 135 in the membrane is stable after umbrella sampling with the presence of the AHL synthase ligand SAM.



The potential of mean force (PMF) can be calculated from placing a restraining the SAM ligand with some bias, δ, and observe this stability as it is moved from position a to b. From this, we can confirm if this is the true binding site of the protein. These simulations are still being conducted.

References:

Bertrand-Burggraf E, Ling C, Schnarr M, Lefèvre JF, Pouyet J, Daune M. Fast abortive initiation of uvrA promoter in a supercoiled plasmid studied by stopped-flow techniques. FEBS letters. 1987 May 4;215(1):83-7.

ThermoFisher Scientific. Macromolecular components of E. coli and HeLa cells. n.d. Accessed from: http://www.thermofisher.com/es/en/home/references/ambion-tech-support/rna-tools-and-calculators/macromolecular-components-of-e.html. [Date Retrieved Oct. 29, 2017]

Shepherd N, Dennis P, Bremer H. Cytoplasmic RNA Polymerase inEscherichia coli. Journal of bacteriology. 2001 Apr 15;183(8):2527-34.

Kierzek AM, Zaim J, Zielenkiewicz P. The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. Journal of Biological Chemistry. 2001 Mar 16;276(11):8165-72.

Krin E, Danchin A, Soutourina O. RcsB plays a central role in H-NS-dependent regulation of motility and acid stress resistance in Escherichia coli. Research in microbiology. 2010 Jun 30;161(5):363-71.

Karzbrun E, Shin J, Bar-Ziv RH, Noireaux V. Coarse-grained dynamics of protein synthesis in a cell-free system. Physical review letters. 2011 Jan 24;106(4):048104.

Ishihama A, Kori A, Koshio E, Yamada K, Maeda H, Shimada T, Makinoshima H, Iwata A, Fujita N. Intracellular concentrations of transcription factors in Escherichia coli: 65 species with known regulatory functions. Journal of Bacteriology. 2014 May 16:JB-01579.

Eguchi Y, Utsumi R. Alkali metals in addition to acidic pH activate the EvgS histidine kinase sensor in Escherichia coli. Journal of bacteriology. 2014 Sep 1;196(17):3140-9.

Foster JW. Escherichia coli acid resistance: tales of an amateur acidophile. Nature Reviews Microbiology. 2004 Nov 1;2(11):898-907.

Chen LQ, Pagel MD. Evaluating pH in the extracellular tumor microenvironment using CEST MRI and other imaging methods. Advances in radiology. 2015 Jul 1;2015.

Delli Castelli D, Ferrauto G, Cutrin JC, Terreno E, Aime S. In vivo maps of extracellular pH in murine melanoma by CEST–MRI. Magnetic resonance in medicine. 2014 Jan 1;71(1):326-32.

Mackerell AD. Empirical force fields for biological macromolecules: overview and issues. Journal of computational chemistry. 2004 Oct 1;25(13):1584-604.