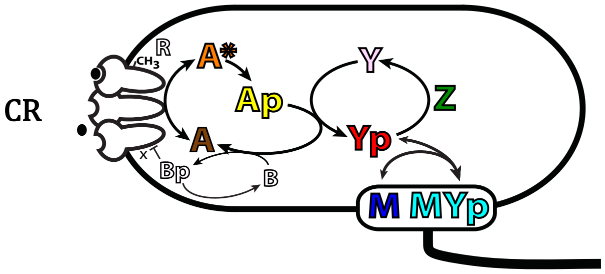
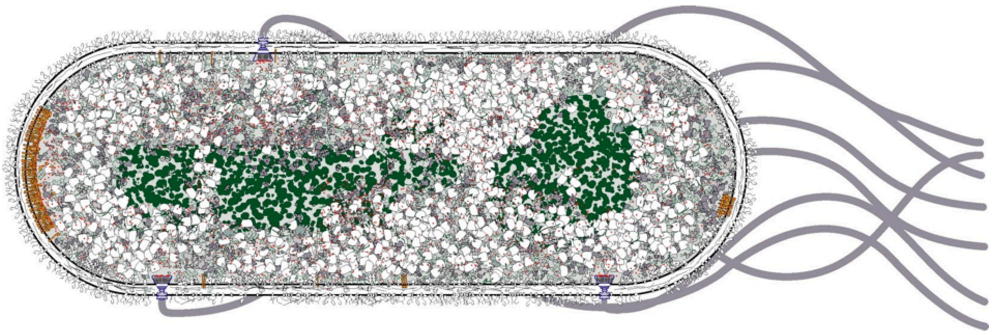
NST Part III Systems Biology
MIB2.2, 31/01/2024
Karen Lipkow, klipkow@gmail.com

Smoldyn I

In this practical, you will use the particle-based simulator Smoldyn to build a spatial, stochastic model of the *Escherichia coli* chemotaxis system.

Chemotaxis enables the bacteria to swim to a favourable environment. The core signalling pathway consists of only a couple of protein species:





The five species of transmembrane Chemoreceptors (**CR**) are arranged in a large polar cluster. They bind a variety of small chemoeffector molecules, such as amino acids and ions. The activity of the chemoreceptors is affected by their binding status, methylation status, and the activity of their neighbours in the cluster.
CheA (A): Kinase: Bound to the inside of the chemoreceptor cluster and activated by them. Active CheA (**A*, Am**) autophosphorylates to form CheA-Phosphate (**Ap**).
CheY (Y): Small, diffusing signalling molecule. Picks up phosphate from CheAp. CheY-Phosphate (**Yp**) can bind to:
FliM (M): ~34 FliM proteins form the cytoplasmic-most ring of the flagellar motors, which are distributed along the length the cell. Binding CheYp (**MYp**) increases the likelihood of the motor switching from counter-clockwise rotation and smooth running, to clockwise rotation and tumbling.
CheZ (Z): Dimeric protein, helps with the dephosphorylation of CheYp to CheY. Some of the CheZ dimers are bound to the cluster (**Z**), others are freely diffusing (**Zf**).

All required software and files are installed on the computers of the Bioinformatics Teaching Facility. If you want to run it on your own computer, you will need
- Smoldyn (download and installation instructions: smoldyn.org)
- a Terminal application
- a coding text editor, such as Visual Studio Code, Atom, SublimeText or BBEdit
- MATLAB
- basic knowledge of Unix
- Folder "Smoldyn1" from Moodle.

Exercise 0. Starting a Smoldyn simulation

Open the *Terminal* application.
Navigate into the folder:
~/Course_Materials/Smoldyn1/examples/S8_reactions/lotvolt

This can be done in two ways:

A. behind the $ prompt, type:

```
cd Smoldyn1/examples/S8_reactions/lotvolt <enter>
```

B. use the File Manager to navigate to the "S8_reactions" folder and open it. Then type behind the $ prompt: `cd <space>`
and drag and drop the "lotvolt" folder into the Terminal. Confirm with `<enter>`.

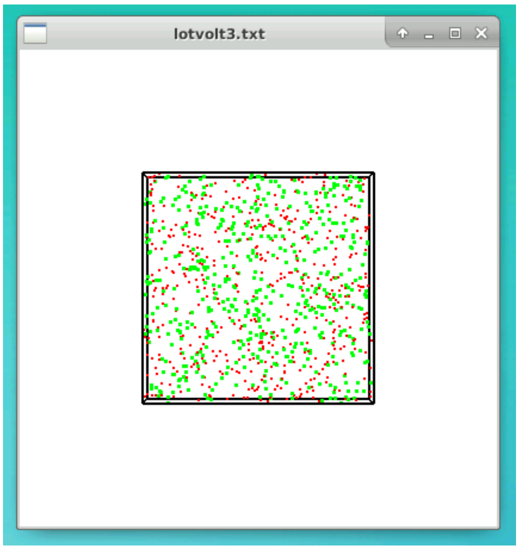
Now open the lotvolt folder to see what's in it.
Start a Smoldyn simulation:

Behind the $, type

```
smoldyn lotvolt3.txt <enter>
```

Note that the spelling is case sensitive and must be exact, and the file extension (.txt) must be included, even if it is not displayed in the File Manager.

You should see this window, which is the starting position of the Smoldyn 3D Lotka-Volterra model. Red dots represent predators (lynxes or foxes), green dots represent their prey (rabbits).



Increase the size of the window (third symbol in top right corner).

Select the graphics window and:

Unpause and pause the simulation by pressing the `<space>` bar.

Zoom out and in by pressing the minus `<->` and equal `<=>` keys.

Rotate the system using the arrow keys.

Terminate the simulation by selecting the graphics window and pressing `<shift-q>`. The first `<shift-q>` will stop a running simulation, the second `<shift-q>` will quit Smoldyn. (Only one `<shift-q>` required if the simulation has already finished.)

Note: If your simulation is creating any output, it is very important to terminate it with <shift-q>, rather than by just closing the graphics window or the terminal. Only then will the latest contents of the buffer be written down, and the stdout updated.

Exercise 1. Building a simulation system: Configuration file basics

Using the *Visual Studio Code* text editor, open the empty file Smoldyn1/MySims/MySims1.txt . Feel free to rename it. This will be your configuration file.
To start a simulation from it, you'll need to navigate the Terminal into the same directory.

Your task is now to translate my mostly-plain-English instructions into Smoldyn code. To guide you, use SmoldynQuickGuide.pdf and SmoldynUsersManual.pdf, which are in documentation/MIB_docs1.

If you get stuck, don't hesitate to ask for help. If all fails, you'll find correct configuration files in the Smoldyn1/solutions folder. (ctseb....txt)

For this practical, the time unit is milliseconds (ms) and the spatial unit is px (10 nm).

In the text file, enter:

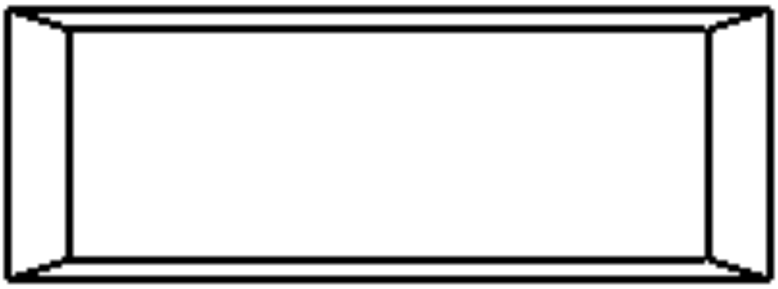
commented title line
graphics statement:
opengl_good
3 dimensions
species:
A, Am, Ap, Z, Zf, Y, Yp, M, MYp

```
time: 0-500 ms, Δt = 0.1 ms

reflective boundaries (px):
(x) -50 230
(y) -50 50
(z) -50 50

end_file
```

Solution:
You should see plain box.



(ctseb1.txt)

2. Add molecules

Add some molecule placements to the configuration file:

```
uniformly distributed:
8200 Y
780 Zf

on plane: x = -48
1276 A

on plane: x = -46
820 Z
```

Solution:
The box should now be filled with black dots.

(ctseb2.txt)

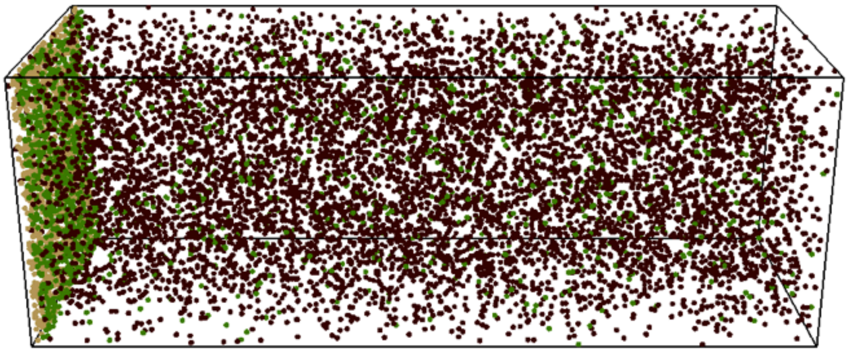
3. Define molecule colours and sizes

colours:
A 0.7 0.6 0.34
Am 1 0.5 0 or orange
Ap 1 1 0 or yellow
Z 0.25 0.5 0
Zf 0.25 0.5 0
Y 0.2 0 0
Yp 1 0 0 or red
M 0 0 0.5 or blue
MYp 0 1 1 or cyan

display sizes:
Ap and Yp: 2
all others: 1

Hints: Use American spelling, make use of wildcards (*), and pay attention to the order of statements.

Solution:
After zooming and rotation:



(ctseb3.txt)

4. Diffusion

diffusion coefficients:
Y and Yp 100
Zf 60
all others: 0 (default)

Solution:
Snapshot as before, but the cytoplasmic molecules should be moving.
(ctseb4.txt)

5. Reactions

type below the difc statements:

```
reaction a1 Am -> Ap          3.4e-2
reaction a2 Y + Ap -> Yp + Am 1.66e2
reaction y1 Y -> Yp           5e-8
reaction y2 Yp -> Y           8.5e-5
reaction m1 Yp + M -> MYp     8.3e0
reaction m2 MYp -> M + Yp     2e-2
product_placement m2 pgemmax  0.2  #optional
reaction z1 Yp + Z -> Y + Z   2.67
reaction z2 Yp + Zf -> Y + Zf 2.67
```

Solution:
As before, all molecules still only green, beige and brown, i.e. inactive and unphosphorylated.

(ctseb5.txt)

Task:
Look at the configuration file and try to understand why there are no phosphorylated molecules. Only look at the solution when you think you have found it.

Solution:
The activation cascade starts with Am = activated CheA. No Am is initiated or produced so far.

6. Commands

In real life, CheA is activated by the transmembrane chemoreceptors. As their activation and interaction is quite complex, they are not part of this model. Instead, we can directly set the activation profile of CheA. To do this at specified times during the simulation, we use **commands.**

A very useful command for these purposes is `equilmol`. When it is called, it will stoachstically equilibrate two molecule species: Please check the User's Manual for the detailed definintions of `cmd` and `equilmol`.

```
cmd @ 100 equilmol A Am 0.13
```

means that at time 100, each A and each Am molecule will have their identity reassigned: each has a 13% chance of becoming an Am, and an 87% chance of becoming an A.

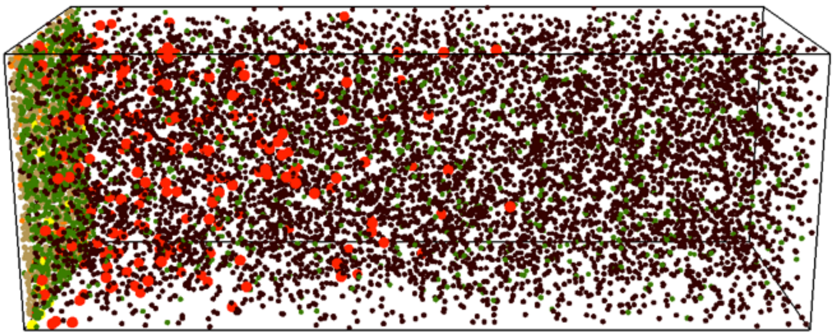
Exercise:
Use commands to give instructions:

before simulation start: pause

before simulation start and at t = 100 ms:
equilibrate A and Am, to have 13% Am

at t = 200 ms:
equilibrate A and Am, to have 40% Am

Solution:



(ctseb6.txt)

7. Readfiles

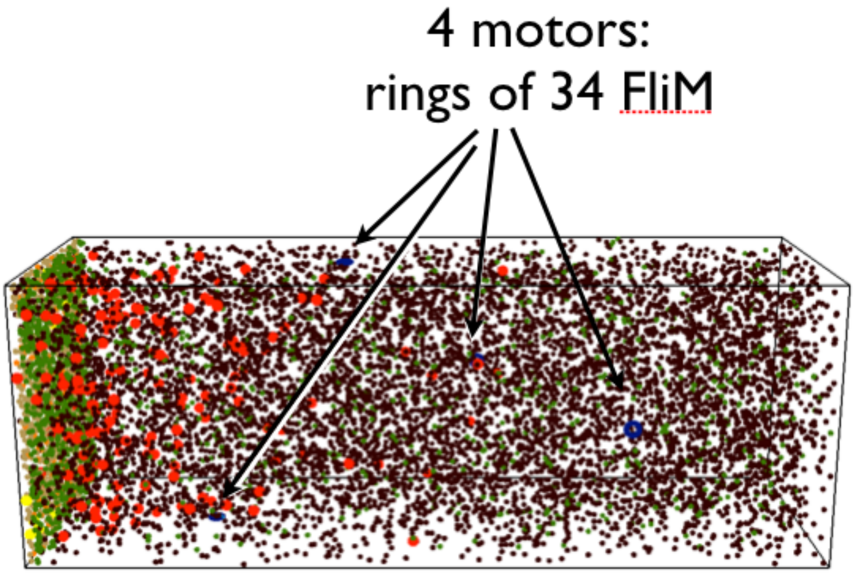
It can be useful to separate the configuration file into several files, e.g. to keep the main file short and easy to read. One example: Long lists of molecule positions.

To position the FliM molecules into four rings at the cell's lateral sides, I have used C to write a long list of `mol` statements. (Any scripting language will do.) It is saved in the file posE_M_49, which can be found in Smoldyn1/readfiles.

Copy this file into your current working directory, and add this line below the other `mol` statements:

```
read_file posE_M_49
```

Solution:

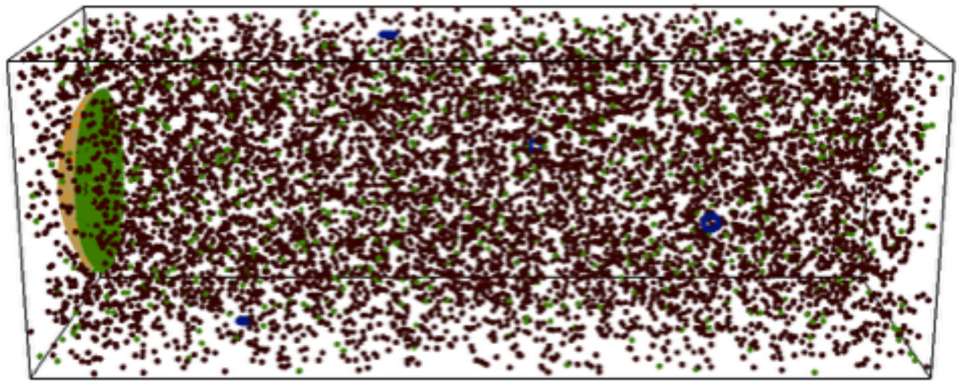


(ctseb7.txt)

8. Preparation for a more realistic *E. coli* shape

Replace the random CheA and CheZ clusters with ordered, curved arrays. Call readfiles posE_A_48_1.5_1276 and posE_Z_46_2_820.

Solution:



(ctseb8.txt)

Hint: The previous random arrays are confined to one plane in x. Remember to comment them out.

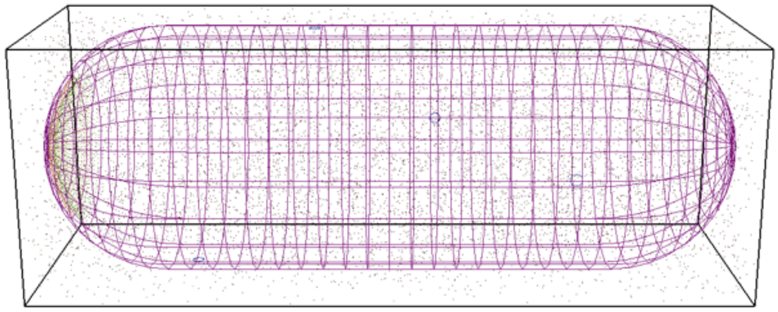
9. Surfaces

Add cell membrane as *E. coli*-shaped surface, filling the boundaries.

It helps to:

- initially choose graphics opengl (instead of opengl_good) to make molecules appear smaller
- look at the surfaces before starting the simulation (don't hit the space bar)

Solution:



Hint 1: SmoldynManual p. 127ff.

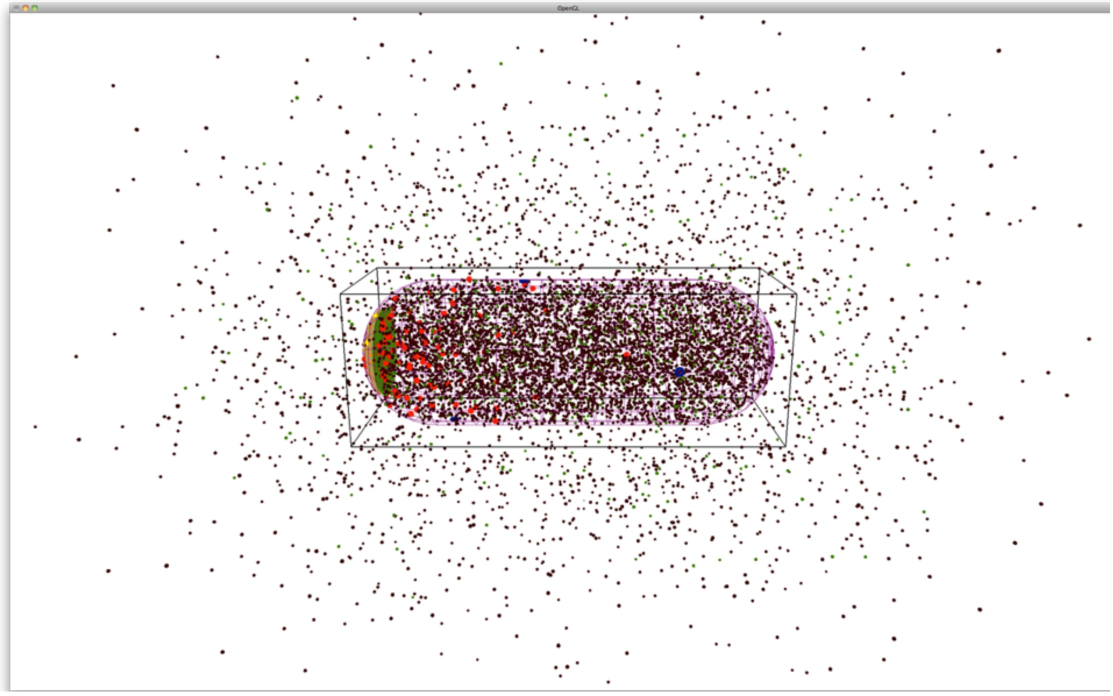
Hint 2: (on next page)

Hint 2: Two polar hemispheres, one cylinder; reflective.

(ctseb9.txt)

10. Understanding surfaces

Return to good graphics and start simulation...



(ctseb10.txt)

Task: Think about these questions before looking up the answers

Question 1: Why are the molecules escaping?

Question 2: Which molecules are escaping?

Question 3: Were they supposed to be there in the first place?

Question 4: What is a smart solution to the problem?

Answer 1: If a model includes surfaces, boundaries transmit.

Answer 2: The molecules that were initialised outside the surface (cell membrane).

Answer 3: No, all cytoplasmic molecules are supposed to be in the cytoplasm, within the cell membrane.

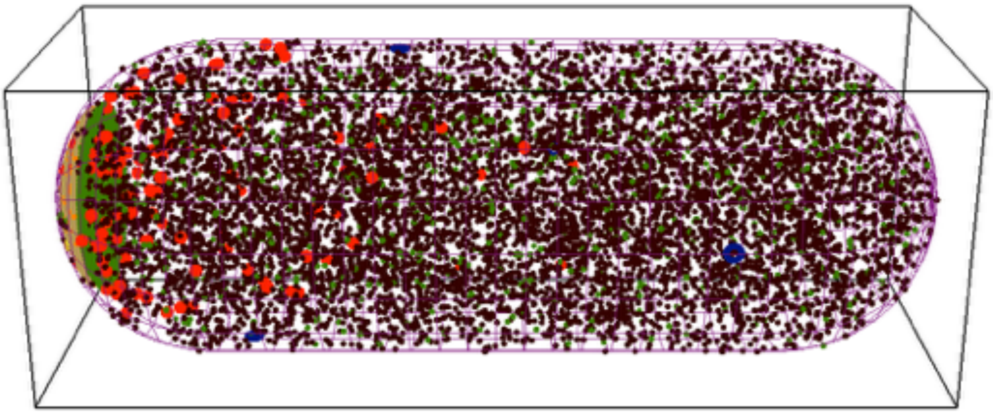
Answer 4: Compartments!

11. Compartments

Make a compartment "cytoplasm".

Place diffusing molecules into this compartment.

Solution:



(ctseb11.txt)

12. Better graphics

Remove the frame around the boundaries.

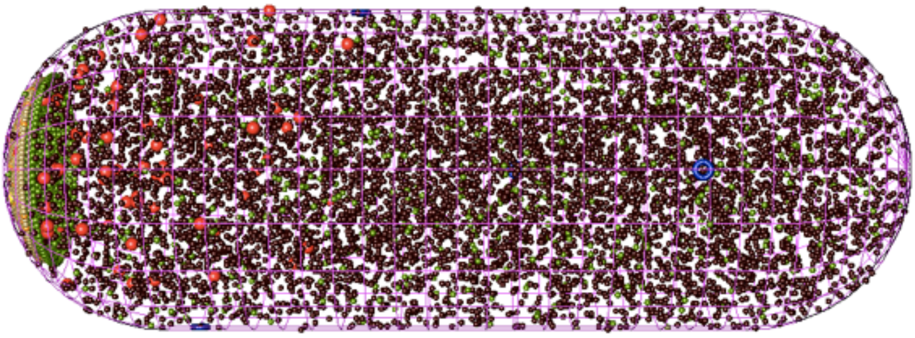
Set the graphics quality to opengl_better.

Set the background to the RGB colour white.

Add the following light settings:

```
light 0 position -50 50 0
light 0 diffuse 1 1 1
light 0 ambient 0.05 0.05 0.05
light 0 specular 1 1 1
```

Solution:



(ctseb12.txt)

13. Recording the system and plotting results

Define output file out2.txt

Count number of all molecules, using the `molcount` command at 10 ms intervals, and let the results be written into out2.txt.

Comment out all (!) graphics statements, incl. `frame_thickness`.

Set the time step length to 1 ms.

Change the equilmol statements:
Remove cmd b pause.
For the first 2000 ms, equilibrate A and Am every 10 ms to 13.2% Am;
from 2000 to 4000 ms, equilibrate A and Am every 10 ms to 6.6% Am;
from 2000 to 4000 ms, equilibrate A and Am every 10 ms to 13.2% Am.

Solution:
A file out2.txt will appear in your current directory and be filled with 601 lines, 10 columns of numbers.

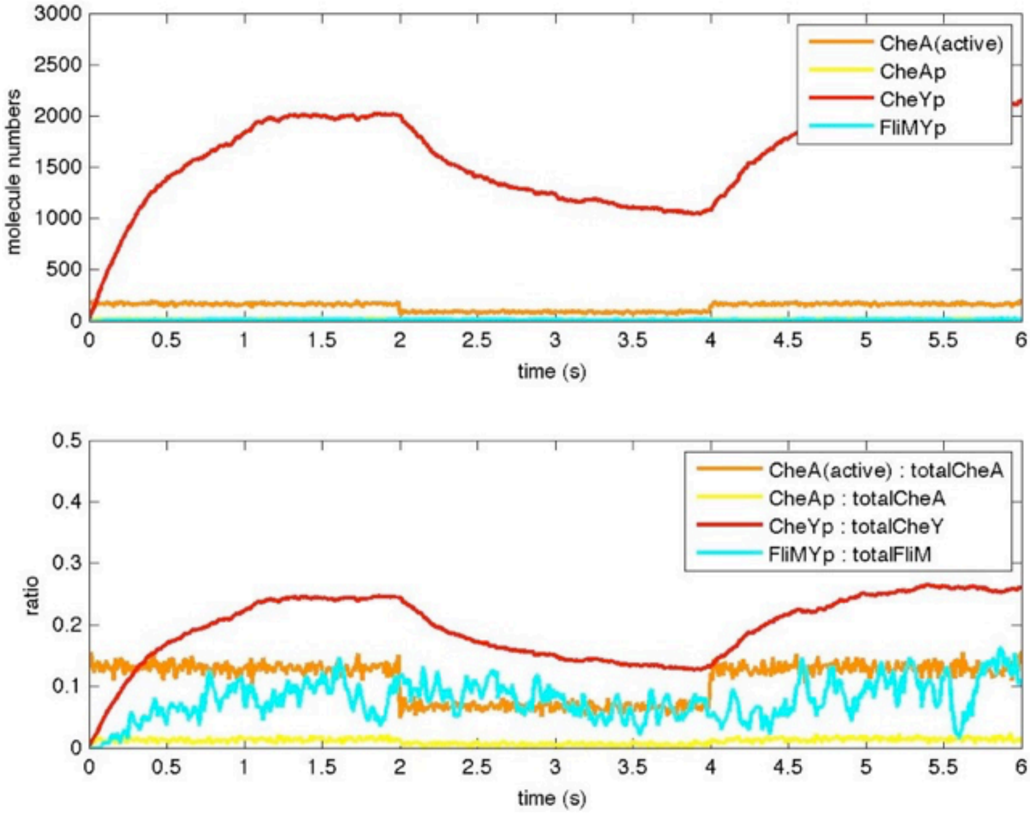
(ctseb13.txt)

Open *MATLAB* by double clicking the desktop icon or by typing `matlab` into a new Terminal window.

Copy the file Smoldyn1/m-files/Mctseb.m to your working directory (MyFiles).

Plot the simulation results in MATLAB using Mctseb.

Solution:



Read Mctseb.m and understand it. Why does it make sense to plot the second subplot?

This is the point you need to reach in the first practical. If you still have time, please continue. Otherwise, please continue here tomorrow.

14. More quantitative simulations: Prepare for spatial distributions

Within your MySims directory, create a new folder for each simulation, and copy all required files into it.

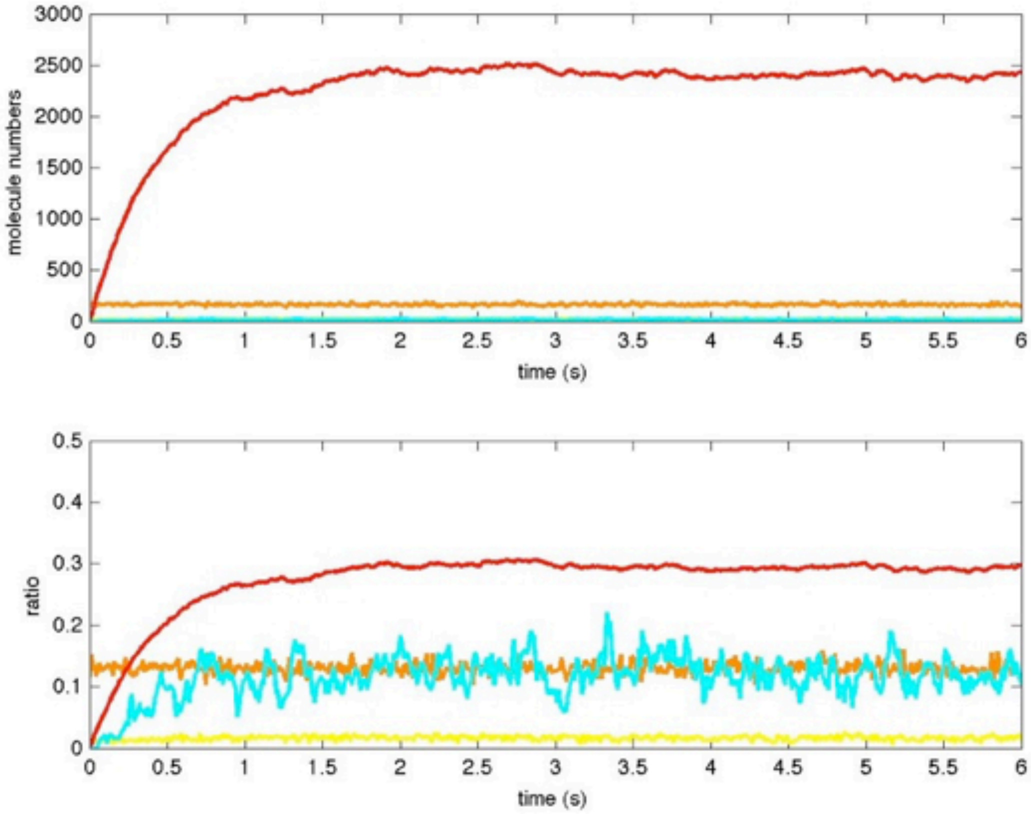
To get a more pronounced Yp gradient, make all Z diffusible.

To let the cell reach steady state, change the CheA activation profile to always have on average 13.2% Am.

Record the time profile of molecule numbers to check when steady state is reached.

Plot with MATLAB (Mctseb.m).

Solution:
(ctseb14.txt)



Note that the steady state is reached at c. 2000 ms.

15. Record spatial distributions

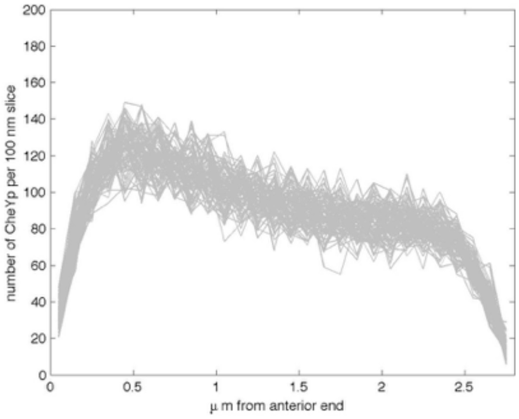
Shorten the simulation duration to 3000 ms.

Record histograms of the spatial distribution of Yp along the x-axis, using the command `molcountspace`:
- 100 snapshots every 10 ms at steady state: save in a file `out3_Yp.txt`
- 30 x 100ms-averages during buildup: save in a file `out4_Yp.txt`

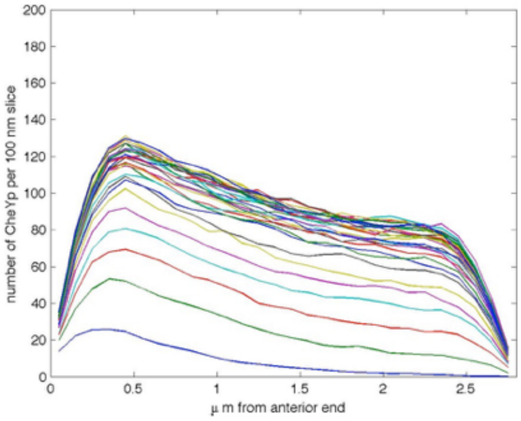
Plot out3_Yp.txt with m-file Yprofile.m and out4_Yp.txt with YprofileTime.m.

Solution:
(ctseb15.txt)

molcountspace (out3_Yp) / Yprofile.m:
snapshots at steady state (2000-3000 ms)



molcountspace (out4_Yp) / YprofileTime.m:
100ms-averages during Yp buildup (0-3000 ms)



16. Record normalised spatial distributions

Question: Why do the Yp molecule numbers go towards zero at both ends?

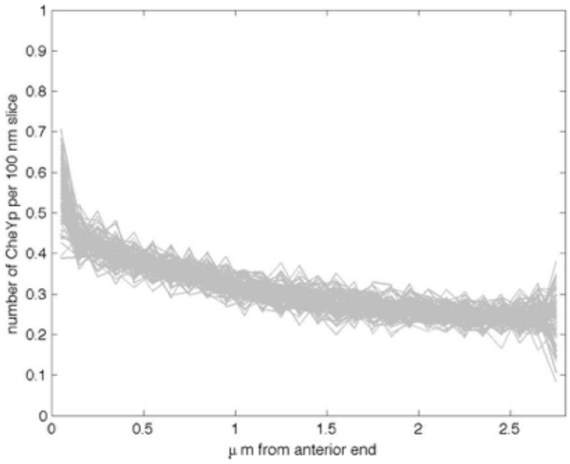
Answer: The cell has hemispherical ends, with less space for molecules at the tips than in the centre.

To normalise for total CheY concentration and varying volume along the x-axis, also record profile of Y molecules.

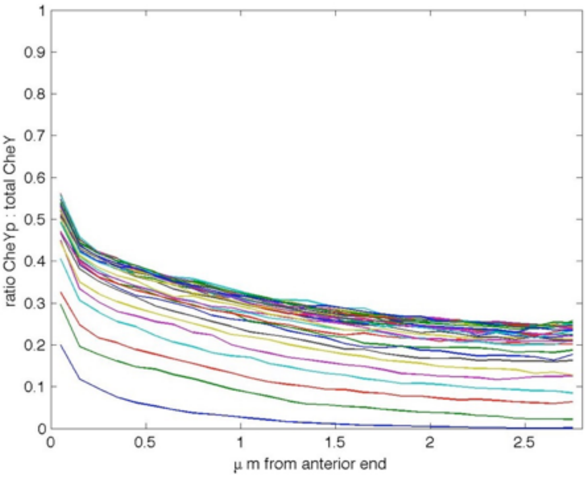
Calculate the ratio Yp/Y and plot using MATLAB: YprofileRatio.m and YprofileRatioTime.m.

Solution:
(ctseb16.txt)

molcountspace (out3_Y and out3_Yp) / YprofileRatio.m:
snapshots at steady state (2000-3000 ms)



molcountspace (out4_Y and out4_Yp) / YprofileRatioTime.m:
100ms-averages during Yp buildup (0-3000 ms)


ratio CheYp : total CheY
µ m from anterior end