Escherichia Coli simulation using the Dynamic Energy Budget theory applied on a multi-agent environment
(To be submitted to the Journal of Theoretical Biology)

Authors
Sara Cruz, Pedro Santos, Carlos Martinho, Gonçalo Marques, Tiago Domingos

Abstract

Escherichia Coli is a Gram-negative, rod-shaped bacterium, and one of the oldest microorganisms that maintains a symbiotic relation with Men. It is very commonly used in research and, therefore, it has been long-time established as a reference in microbiology studies. In this work, the results of a multi-agents simulation of E coli, using the Dynamic Energy Budget theory, are presented, with the goal of reproducing the life cycle behaviors and movement patterns that characterize this bacterium. The work also includes an implementation of a set of mutations based on the alteration of simulation variables and DEB parameters. The ION Framework was used as a development tool in order to generate the bacterial environment and the corresponding interaction protocols with the agents. The aim of this implementation is to test the possibility of using the DEB theory applied in a multi-agent scenario to obtain realistic results comparable with previously reported experimental investigation.

Introduction

Systems’ computer modeling is a very important area of research nowadays. By translating knowledge obtained through experimentation so it can be processed by a computer, it is possible to watch a realistic simulation of the world it represents and easily alter its parameters dynamically [20].

The Dynamic Energy Budget theory [7] is a mathematical theory that aims to reproduce the life cycle of all living beings, from bacteria to plants and animals, at an individual level. Many multi-disciplinary teams around the world are, at the moment, investing work towards modeling tens or even hundreds of organisms using DEB theory, and have been able to do so quite successfully for many examples [4] [5] [6] [15] [17] [18].

The Escherichia Coli bacterium has been long-time established as a reference in microbiology studies, because of the simplicity and organization of its DNA. Hundreds of different E. Coli strains have been studied and sequenced. This takes particular relevance, since many Escherichia Coli strains are pathogenic and can be mortal for humans. Serious outbreaks have occurred throughout history. Most recently, the E. Coli O104:H4 [14] strain outbreak that started in Germany in May of 2011 contaminated more than 3000 people, causing the death of 50 of them. Therefore, it becomes relevant for academic purposes and useful for such public health problems to have an accurate model that allows the prediction of these bacteria’s behaviors.

It is simple to see the usefulness of uniting such a wide-range mathematical model as the DEB theory with a multi-agent computer program. It would give scientists the power to quickly model potentially any living organism, and easily run realistic computer simulations that provide data with a detail level that would be very hard to obtain with a regular laboratorial experiment.

The Life-Engine tool, developed by the Biodroid Company, seeks to take this approach towards creating a biological DEB engine that allows the addition of realistic biological creatures to video-games as well as way to model them for academic purposes. However, it is not yet complete. There are some functionalities that are not yet implemented, and the E. Coli parameters were not yet calculated.

The present work represents a step towards the ambitious goal of creating a powerful biological simulation engine that allows the easy creation of DEB agents, their worlds and interactions between them.

The main goal of the implementation was to develop an Escherichia Coli simulation by using a multi-agent approach and applying the Dynamic Energy Budget theory. The developed application is able to accurately imitate the chemotactic movements displayed by this bacterium as well as its life cycle and interactions with the environment it dwells in. It also provides a set of mutated E. Coli strains as well as an objectively defined process to include new mutations in the existing model. The obtained results match the ones reported in experimental projects using real E. Coli for the movement patterns [3] [10], DEB modeled life cycle [11] [13] and mutations [1] [8] [9].

Dynamic Energy Budget Theory

The Dynamic Energy Budget (DEB) [7] theory aims to describe the life cycle of all living organisms, whether they are uni or multi-cellular, as the energy fluxes between them and the environment, as well as the internal energy fluxes between their sub-components.

The organism’s development state is defined by 3 main state variables:
Reserve ($E$) refers to the energy absorbed from the substrate present in the environment and stored to be used for metabolic purposes. The energy contained in this variable will be redirected to the other components in order to fulfill the organism’s needs.

Structure ($V$) is related to the physical size of the organism. Any increase in structure (growth) will later require energy to be mobilized from the reserve towards its maintenance.

Maturity ($E_B$) relates to the complexity (development level) of the organism. When a being is created, its maturity ($E_B$) has value 0 and it is designated as an Embryo. Birth occurs when maturity reaches the birth threshold ($E_{b}$), and is defined as the point where the organism starts feeding (assimilation). It becomes a Juvenile. When the puberty threshold is reached ($E_{p}$), the individual gets the ability to generate offspring, becoming an Adult. Until this moment, the energy invested towards maturity increase was devoted to maturation, to augment the individual’s complexity. DEB theory considers that when an individual becomes an adult, it has reached the maximum level of complexity, making it unnecessary to invest more energy towards maturation.

From puberty on, reserve energy will be applied to reproduction, so that an organism can generate offspring. The energy that was formerly mobilized to maturity will be "stored" in the Reproduction Buffer ($E_R$). Each time the mother lays an egg, $E_R$ will decrease. It is as if the mother releases energy into the world, in the form of an egg. A new egg is only made of reserve, it has no structure. Until birth, the energy will be converted to structure and the embryo will not feed, living of the reserve energy given by the mother. When an organism reaches the maturity level at birth, it becomes a juvenile. At the moment of birth, the reserve density [$E$] ($E/V$ – the relation between reserve and structure) is the same as the mother’s when it laid the egg. This is called the Maternal Effect.

**Standard DEB model**

Figure 1 shows how energy flows through the DEB components of an organism, in the form of its various fluxes. $p_x$ is the ingestion flux. It contains the energy that enters the organism from the environment as the result of ingestion and has the chemical composition of food.

After assimilation, the energy will be stored in the Reserve ($E$). This is flux $p_A$, the assimilation flux, which results from ingestion but already has the chemical composition of reserve.

From the reserve, energy will be mobilized to the other variables to fulfill the individual’s biological needs. This is defined as the mobilization flux $p_C$. One part of the reserve energy is applied to soma, and the other to maturity. The proportion in which reserve energy is applied to these two components is defined by $\kappa$, a constant that takes values between 0 and 1. Therefore, a fraction $\kappa$ of $p_C$ will be used for growth $p_G$ and somatic maintenance $p_S$, while the rest of it goes to maturity increase, reproduction ($p_R$) and maturity maintenance ($p_I$). A higher maturity value will mean that the individual has greater complexity, and a higher structure value will mean an organism is bigger. DEB theory considers that one part of the somatic maintenance costs is proportional to volume and the other part is proportional to the surface area, and maturity maintenance is proportional to the organism’s complexity. Therefore, growth and maturation will imply greater maintenance costs from that moment on.
Sometimes, there isn’t enough food in the medium to support the organism’s activities. When there’s substrate shortage, growth continues for a short period of time, using the reserve. Afterwards, if there isn’t enough substrate available for somatic maintenance, the structure will diminish, causing the individual to “shrink”. When structure reaches a minimum threshold value, death occurs. If there isn’t enough available substrate to assure maturity maintenance, this will cause the value of this variable to decrease, “rejuvenating” the organism.

**Time-dependent differential equations**

Biological simulations are usually applied to reproduce the behavior of a given environment during a given time interval. The standard way to apply a mathematical model, such as DEB, to a computer simulation is to define a clear set of system variables, input and output parameters and how they evolve through time.

We have our three main state variables: Reserve, structure and maturity, with which we can define an individual’s state and recreate the main events in its life cycle. These values can be differentially defined as functions of time, by using other time-dependent variables: the DEB energy fluxes, in the following way:

Reserve: \[
\frac{dE}{dt} = \dot{p}_A - \dot{p}_C
\]

Structure: \[
\frac{dV}{dt} = \frac{\dot{p}_G}{E_G}
\]

Maturity: \[
\frac{dE_H}{dt} = \dot{p}_R , \quad E_H < E_H^p
\]

Reproduction Buffer: \[
\frac{dE_R}{dt} = \dot{p}_R , \quad E_H > E_H^p
\]

At this point, all that there is left is to have the set of equations needed to create a DEB computer model is to define these fluxes as functions of parameters that can be obtained experimentally or by browsing the literature:

\[
\dot{p}_A = V^{2/3} \left\{ \dot{p}_{Am} \right\} f
\]

\[
\dot{p}_C = E \left[ \frac{\dot{p}_M}{E_G} + \ddot{v} / L + \ddot{p}_S / V \right] / \kappa E / V + [E_G]
\]

\[
\dot{p}_S = V \left[ \dot{p}_M \right] + V^{2/3} \left\{ \dot{p}_T \right\}
\]

\[
\dot{p}_G = \kappa \left( \ddot{p}_C - \ddot{p}_S \right)
\]

\[
\dot{p}_R = \kappa E_H
\]

\[
\dot{p}_J = (1 - \kappa) \ddot{p}_C - \ddot{p}_J , \text{where,}
\]

- \left\{ \dot{p}_{Am} \right\}, in 1/(\mu m^2 s) is the surface-area-dependent maximum assimilation rate, i. e., the maximum amount of energy a being with a given surface-area can assimilate in a given amount of time;
- \( f \) is the functional response. A being is not always able to assimilate according to its maximum assimilation rate, mainly because the amount of food present in the environment may not allow it. The functional response is multiplied by the maximum assimilation rate to specify how much of that energy can be assimilated according to environmental or other limitations;
- \( [E_G] \), in J/\mu m^3 is the volume-specific costs of structure, which defines how much energy must be mobilized from the reserve in order to add a given amount of volume to the structure value;
- \( \ddot{v} \), in \mu m/s is called energy conductance and specifies how fast energy travels through an individual with a given length;
- \( L (\mu m) \) is the structural length \( V^{1/3} \);
- \( \left[ \dot{p}_M \right] \), in 1/(\mu m^3 s) is the specific volume-linked somatic maintenance rate, which expresses the amount of energy that must be allocated from reserve in order to maintain a given amount of structure for a given time interval;
- \( \left\{ \dot{p}_T \right\} \), in 1/(\mu m^2 s) is the specific surface-area-linked somatic maintenance rate, which has the same meaning of \( [\dot{p}_M] \), but refers to maintenance costs which relate to the surface-area of an individual, such as maintaining the body-temperature.
- \( \ddot{p}_J \) (s^-1) is the maturity maintenance rate coefficient

**V1-morphs**

For simple beings, such as prokaryotic cells, the cell-cycle stages can be explicitly related to the structure value \( V^{1/3} \). The E. Coli bacterium, for instance, reproduces through division when it has duplicated its entire DNA. The organism’s complexity doesn’t change throughout its life, and all the cell cycle events (birth, reproduction and death) result from changes in structure alone. Cell division, for example, occurs when the structure threshold at division is reached \( V_d \), producing two daughter cells with half the structure of the mother. To represent these kinds of beings using the DEB model, it is usually considered that they are always juveniles and that the constant \( \kappa \) is equal to 1, meaning that there is no energy invested into maturation or the reproduction buffer. This allows a simplification of the

![Energy fluxes in simplified model for unicellular organisms](image)
conceptual model, displayed in Figure 2.

Another concern one must have when modeling using the DEB theory is to take into account the shape of the organism. The standard DEB equations assume that the organism is an isomorph, which means that it grows proportionally in all three dimensions. Individuals with this kind of shape are considered $V^{2/3}$-morphs, since their volume is proportional to their surface area to the power $2/3$. This relation is very important since it affects the substrate uptake process and maintenance. However, some beings such as certain bacteria grow in length alone. Such organisms are called V1-morphs, because their body volume is directly proportional to one dimension: length. There is one other shape possible: V0-morphs, that maintain their surface area throughout their entire life, which does not mean that there is no structure-wise growth. Structure increases throughout their lives, just as with other organisms, but it gets constricted, maintaining the same surface area from birth until death.

Eschericia Coli is a rod. Its body grows in length alone, but it maintains two constant size caps in the extremities. As such, it is considered a mixture between a V0 and V1 morph.

To deal with these possibilities, DEB theory defines the shape correction function $M(V)$, which must be inserted in the equations that involve energy conductance $\dot{v}$ and the maximum assimilation rate $\left\{P_{Am}\right\}$. For beings that are not isomorphs, $\dot{v}$ should be replaced by $\dot{v} \times M(V)$, and the same method should be applied to $\left\{P_{Am}\right\}$. The shape correction function for rods is:

$$M(V) = \frac{\delta}{\delta + 2} \left(\frac{V}{V_d}\right)^{2/3} + \frac{2}{\delta + 2} \left(\frac{V}{V_d}\right)^{1/3}$$

The $\delta$ constant is called the shape coefficient and it determines the contributions of the V0 and V1-morph components. When $\delta \rightarrow \infty$, the being becomes close to a V0-morph. The E. Coli has a shape coefficient of 0.26.

### Results and Discussion

**Chemotaxis**

The bacterial response to attractants (chemotaxis) was evaluated in three ways: the distribution of the time intervals spent in each state, the paths defined by the bacterial movements and the emergent behavior that arises from the final implementation as a response to an attractant gradient. All the results presented in this section were obtained with the same set of chemotaxis parameters:

Simulation step = 50 ms (time interval between simulation updates)

Sensor step = 100 ms (time interval between sensor updates, which include the analysis of the substrate concentration and adequate chemotactic response)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumble Prob.</td>
<td>80 %</td>
<td>15 %</td>
<td>10 %</td>
<td>5 %</td>
</tr>
</tbody>
</table>

Table 1 Tumble probabilities according to the detected differences in substrate concentration (Current – Previous substrate concentration)

**Motor rotations**

The bacteria were placed in a medium without attractant, to reproduce the conditions described for the experiment. Results were obtained after a period of 150 minutes.

The flagellar motors can turn clockwise (CW), which produces a tumble, or counterclockwise (CCW), which produces a run. When switches between the rotation modes occur, the time spent in the previous mode is calculated, and 1 is added to the frequency registered for that time interval, in the corresponding mode. The charts in Figure 3 display the distribution of the relative frequencies of the time intervals spent in each of the flagellar motor orientations.

![Figure 3](image)

**Figure 3** Fraction of time intervals spent in CW and CCW mode. Real, from the work of Korobkova et al. [10] (left) and Simulated (right)

The results obtained with the simulated chemotaxis closely resemble those displayed by real bacteria. The approach followed to reproduce the movement patterns was very simple, since it didn’t directly implement the protein chain responsible for the chemotactic response. The charts represent the result of the final implementation, after fine-tuning the tumble
probabilities for each gradient detected, and the sensing intervals of the E. Coli bacteria.

Path
The images in Table 2 represent the 2D projections of the paths followed by one simulated bacteria as well as a traced sequence of loci obtained experimentally with real E. Coli bacteria

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Simulated with the developed simulation engine</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Table 2 Paths of Real and simulated bacteria. Experimental data from Berg and Brown [3]

Visually, the rotation angles of the tumble movements seem to accurately reproduce those displayed by real bacteria. The run perturbations implemented through the Brownian diffusion method also provided very realistic results. The time spent in both run and tumble mode tested in the previous section, besides making sense numerically, also provided paths that match the ones observed in nature.

Emergent behavior
The final goal of the implemented chemotactic reactions was to have the bacterial cells respond to a glucose concentration gradient in the simulation environment. The ultimate test for the bacterial movement would be to obtain an emergent reaction to a simple circular gradient that guided the agents to its center, in the form of a bulk of bacteria visibly moving away from the periphery of the simulation world, and converging in the center, where the substrate concentration is higher.

A simulation was set in a space with 1.6 * 1.6 * 1.6 mm (4.1 mm³), and after 250 s, 1994 individuals, moving at an average speed of 20 μm/s had gathered in the center of the simulation space.

Life-cycle
The life-cycle implementation of the E. Coli simulation, applying DEB theory, will be divided into three stages: individual and population growth with unlimited substrate, and population growth in starvation.

The tests were carried out using the Br/A (standard) strain of the E. Coli, with a doubling time (τ) of approximately 100 min, fed on glucose. The simulation space had dimensions 1.6 * 1.6 * 1.6 mm, and all the simulations started with around 2000 individuals. The chemotaxis movement parameters applied were the same as defined in the previous section.

Individual growth
According to the DEB theory, the isomorphic creatures usually have von Bertalanffy growth curves. The E. Coli, however, does not display this kind of growth, since it is a mixture between a V1 and a V0 morph, with shape coefficient δ=0.26. Therefore, it is important to test if the growth curve resulting from the insertion of the shape correction function M(V) displays realistic growth results.

![Image](image7.png)

Figure 4 E. Coli individual growth, real (up - dots) [11] and simulated (down). The line displayed in the upper figure is a linear regression obtained from the displayed dots. The dots themselves correspond to data directly obtained from experimentation.

Figure 4 displays E. Coli growth curves obtained from experimental results [11] and from the simulation. The simulated bacterial growth in volume appears as an almost linear function of time, verifying the correction of the implementation and of the shape correction function, when applied to an individual cell. The slight growth acceleration could be adjusted by slightly modifying the shape coefficient (δ) or it might be due to an imprecision in the DEB equations. However, the result came quite close to reality.

Population growth
To test the evolution of the E. Coli population through time, the experience from the work of Martínez-Salas et al. [13] was recreated using the simulation program. In order to do this, the bacterial colony had to be synchronized by size, as it happens in the corresponding laboratorial experience. However, to create a bacterial colony in which all the bacteria were initialized with the same size was not enough. An approach as simple proved unrealistic, since the separation of bacterial cells according to their dimensions is not an absolutely
precise method. Using this approach, the number of individuals in a population doubled almost instantly. To eliminate this error, the cells were synchronized by size, but a small error coefficient was added. The initial structure value of all the bacteria in the population was defined as:

\[ V_i = \frac{V_d}{2} + \text{RandomBetween}\left(0, \frac{V_d}{10}\right) \]

The population growth curves for both tests exhibits periods with no division, characterized by straight flat lines in the chart, followed by multiplication periods, where the relative cell number appears to increase linearly as a function of time (this effect happens because of the logarithmic scale applied).

![Figure 5](image5.png)  
*Figure 5 Simulation of E. Coli Br/A growth with unlimited substrate (t=100 min) using the developed simulation engine*

![Figure 6](image6.png)  
*Figure 6 E. Coli Br/A growth, from the laboratorial work of Martínez-Salas et al. [13] (t=45 min)*

The growth curve displayed by the simulated bacterial colony is very similar to the real population in shape. However, the scale is different from the data gathered in the experimental work. This is to be expected, since the bacteria used for each of the tests had different doubling times (t = 45 min for the real E. Coli and t=100 min for the simulation strain). The results can be considered to be positive despite this fact, since the cycles are equivalent when matched to the doubling times for each of the colonies. The first doubling step ends at time t = t for both colonies, and the second starts at approximately t = 1.9 t, which means that the results obtained are verified by the experimental data.

**Starvation**

Since the model created responds to the amount of food present in the environment through the application of the DEB theory, it is possible to carry out starvation tests using the proposed solution. These were achieved by placing the standard bacterial strain in an environment with no substrate.

**Figure 7** displays the simulated population curve obtained for the Br/A colony without the presence of glucose in the environment. The chart displays a short growth and multiplication period in the first 10 hours, in which the reserve is applied towards structural increase, allowing the division of the bacteria whose structure threshold for duplication is crossed.

![Figure 7](image7.png)  
*Figure 7 Population curve in starvation*

This is followed by a stagnation interval, represented by a straight line. At this point, continuous shrinking occurs. This means that the structure value will decrease without crossing the structure death threshold. The overall structure of the population becomes smaller, but no deaths can be observed. The work of Barcina et. al. [2] states that continuous shrinking can be observed in bacteria placed in starvation conditions, which corroborates the results obtained through the application of the DEB theory. Afterwards, the cell number begins to decrease, as the structure threshold for death starts to be crossed. The population can survive for about 5 days, completely disappearing after this time. The small fluctuations in the downward curve might result from the implemented necrofagy mechanism, which keeps some of the bacteria alive feeding on dead colony members throughout this period of time.

![Figure 8](image8.png)  
*Figure 8 Population curve in starvation, with several E. Coli strains, from the work of Farida K. Vasi and Richard E. Lenski [19]*

When compared to a population curve in starvation from experimental data, it is possible to notice some differences between what happens in nature and the results of the simulation. The real populations last longer, they display no stagnation period and the general shape of the curves is different. This might be due to several factors.
First and foremost, the bacterial strains used in the simulation and in the laboratory are quite different. The simulation applies a standard Br/A E. Coli strain, while the strains used in the experiment are mutants selected from previous tests, which gives them a greater endurance to adverse conditions.

The chart in Figure 9 also displays evident cell number increase in certain points after the descending period. This is probably a cause of the naturally-occurring mutations that take place in the E. Coli bacterium when it is subjected to non-optimal conditions. Although the developed simulation allows the creation of mutations that could possibly reproduce the naturally-occurring ones, it does not allow a colony to change in the course of a single experiment. It is not possible with the implemented solution to have a colony adapt to its environment in the course of a test.

Another known fact is that living beings tend to decrease their metabolic rates when subjected to substrate shortage. The DEB theory does not yet predict this effect in its equations. The rate at which reserve is utilized does not depend on substrate availability. This factor has a tendency to cause DEB populations to extinguish faster than real ones [16].

Lastly, the application does not allow the simulation of numbers of individuals comparable to a laboratorial sample. In this example, we have a sample of 1 ml (1 cm^3), and it is hard to have a simulation sample bigger that 1 cm^3, without having performance issues manifesting in a standard machine.

In general, the starvation mechanisms that are now implemented do not suffice to reproduce what happens in reality, requiring further work towards implementing new features and refining those that are already working.

**Constant glucose input**

The final standard E. Coli scenario was the constant substrate input test. In a laboratory, a certain number of E. Coli is placed in a medium with a small quantity of glucose and, at constant time intervals, a certain quantity of glucose is added to the environment. This allows the bacteria to multiply, by not indefinitely, since the population growth is limited by the amount of substrate in the environment.

Figure 9 displays the behavior of a simulated bacterial colony in the afore-described conditions, using the developed simulation engine.

The graph shows sinusoidal fluctuations of the glucose amount and E. Coli population, with a difference of approximately T/2 between the two curves. This means that, when the E. Coli population increases, the amount of glucose decreases. Which is to be expected, since a higher number of individuals feeding makes the substrate amount in the environment decrease.

At a certain point, the substrate present in the environment is no longer sufficient to feed all the individuals in the simulation, and the population size begins to fall, decreasing the speed at which the glucose is consumed. This makes the substrate concentration rise again. These shifts in population and substrate happen periodically, with a tendency towards stabilization.

**Figure 9 Simulated population growth with constant glucose input.** The environment was initially set with 450 B/rA E. Coli individuals and 1.28 picograms of glucose, in a simulation space of 0.4 X 0.4 X 0.4 mm. The test lasted for 50 days, and an amount of 1.28 pg of glucose was added each 60 minutes. The dots represent numerical data directly extracted from the simulation and the lines are 6th degree polynomial regressions obtained from those point sets.

These results show that, although the starvation implementation is not perfect, the population displays realistic and plausible results in substrate-limited-growth conditions.

**Strains and Mutations**

After having the standard E. Coli model fully functional (B/rA), the process applied to generate the parameters for the rest of the bacterial strains was very direct.

Firstly, the length at division (L_d) was defined from literature data. From its value, the structure value at division was computed using a proportionality constant with the standard strain at the power 3, as follows:

\[ V_{d,A} = \left( \frac{L_{d,A}}{L_{d,S}} \right)^3 \times V_{d,S} \]

where A is the strain we want to model and S is the standard strain. This approach can be followed as long as the initial reserve density [E] for all bacterial strains is equal to the B/rA reserve density at birth [E_0], guaranteeing that the reason between structure and reserve remains coherent.

From there, the metabolic factor is applied, to generate realistic mean division time (t) values. This is done by sequentially trying out metabolic rates, and stopping at the one that provides most realistic results. The mathetic factor is applied by multiplying all the time-dependent variables by its value.

This approach was followed to define the set of DEB parameters, usable with the Life Engine for 3 more slow growing E. Coli strains, and 2 fast growing ones. Each of the strains is coded as a C# object that implements the mutation class.
Table 3 displays the results obtained for all the strains simulated, as well as the set of DEB parameters and literature data used for comparison. The average and standard deviation for the simulation values were generated by a population of 10 tests, where the target strain was placed alone in the simulation environment. For each of those tests, the simulation ran for an amount of time that allowed, at least, 5 division cycles. The bacteria were fed on glucose with no limit in terms of substrate.

As explained before, the metabolic rate affects all the parameters that are time dependent. In DEB notation, those which are dotted: \( \{ \dot{p}_T \}, \{ \dot{p}_M \}, \dot{v} \) and \( \{ \dot{p}_{Am} \} \).

The results obtained for the length at division (Ld) value are completely accurate. This was to be expected, since this value is directly mapped to a DEB parameter, the structure threshold of division (Vd).

As for the mean division time, by manipulating the metabolic factor, it was possible to obtain values that came very close to those displayed by real bacteria. The accuracy of the results is verified by the small variations obtained for every \( \tau \) value, when compared to the average.

After modelling and verifying the correctness of the 6 E. Coli strains, some tests were carried out to see how the different strains reacted when placed in the same environment.

### Unlimited growth

Firstly, a simple simulation was ran, with 4 E. Coli strains with different doubling times and no substrate limit. The simulation started with 64 individuals for each colony and ran for about 9 hours.

As it was to be expected, every colony displayed exponential growth, with much faster cell number increase for those that had smaller doubling times, with the exception of the B/rK strain, which in the end displayed a slightly smaller growth than the B/rA colony.

However, this is just a result of the randomness of the simulation, since the difference is not significant, and the B/rK strain even exhibited greater cell number increase around minute 60.

#### Individual structural growth

In the same test conditions, another simulation was run to compare the structure evolution at an individual level.

![Image](image-url)

**Figure 11 Structure evolution through time for strains B/rA (Ld = 2 \( \mu \text{m} \), \( \tau = 109 \text{ min} \)), B/rK (Ld = 3.6 \( \mu \text{m} \), \( \tau = 100 \text{ min} \)) and MC1000 (Ld = 2 \( \mu \text{m} \), \( \tau = 21 \text{ min} \))**

<table>
<thead>
<tr>
<th>Strains</th>
<th>Slow Growing</th>
<th>Fast Growing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>B/rA</td>
<td>B/rK</td>
</tr>
<tr>
<td>( \tau ) (min)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulated Average</td>
<td>106,3</td>
<td>99,9</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>15,35</td>
<td>7,79</td>
</tr>
<tr>
<td>Ld (( \mu \text{m} ))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulated</td>
<td>2</td>
<td>3.6</td>
</tr>
</tbody>
</table>

| DEB    | | | | | | |
| \( \kappa \) | 1 | 1 | 1 | 1 | 1 | 1 |
| \( \{ \dot{p}_T \} \) | 0 | 0 | 0 | 0 | 0 | 0 |
| \( \{ \dot{p}_M \} \) | 1,13E-14 | 2,26E-14 | 1,53E-14 | 5,31E-15 | 4,27E-14 | 5,54E-14 |
| \( E_G \) | 5,08E-09 | 5,08E-09 | 5,08E-09 | 5,08E-09 | 5,08E-09 | 5,08E-09 |
| \( V_d \) | 0,30 | 1,75 | 0,52 | 0,40 | 0,30 | 0,30 |
| \( \dot{v} \) | 1,25E-04 | 2,50E-04 | 1,69E-04 | 5,88E-05 | 4,73E-04 | 6,13E-04 |
| \( \{ \dot{p}_{Am} \} \) | 7,37E-13 | 1,47E-12 | 9,95E-13 | 3,46E-13 | 2,79E-12 | 3,61E-12 |
| \( E_M \) | 5,91E-09 | 5,91E-09 | 5,91E-09 | 5,91E-09 | 5,91E-09 | 5,91E-09 |
| Metabolic Factor | 1 | 2 | 1,35 | 0,47 | 3,78 | 4,9 |

Table 3 DEB parameters used as input in the Life Engine DEB Library.

* no data was found in literature about the length at division of the fast-growing strains, so the standard E. Coli length at division was used.
between different strains.

Figure 10 displays the results of individual growth for 3 bacterial strains, with different doubling times and structure division thresholds. The peaks in the three lines represent the points where the bacterium divided. At those points, a daughter cell was created with half the structure and reserve of the mother. As a result, the mother also becomes half its size at division. The B strains have approximately the same division time, so they have the same number of peaks. However, strain B/rK is larger and has greater size fluctuations. The MC1000 strain has around 5 times the number of peaks of the slow growing bacteria, since its doubling time is around 5 times smaller.

**Starvation**

Afterwards, the reaction of 4 different bacterial strains to starvation was tested. 121 cells of each strain were placed in an environment without substrate for 11 days. As it had happened in the starvation test with the B/rA strain, the chart displays a clear population growth during the first half of the first day, followed by stagnation, for every strain.

The B/rK colony, however, having a length at division almost 2 times larger than the rest of the bacterial colonies, did not increase the structure value enough for a single individual to reproduce, stepping directly into the stagnation phase.

After the stagnation period, all strains go into a descending phase. The CM1000 colony is the first to be extinct, because it needs to keep a high metabolic rate in order to divide almost 5 times as fast as the B/rA and B/rK strains. The B/rK disappears in second because, although its division time is the same as the B/rA strain, it needs to maintain a higher metabolic rate in order to get to almost the double of the size of the B/rA strain in the same time interval. This means that the reserve is consumed in a higher rate than the rest of the bacteria, as well as the structure when the reserve runs out. This mechanism also explains why the B/rF26 *E. Coli* is the last to disappear, since it has the average size of an *E. Coli* and a division time that is more than the double of the rest of the slow growing cells.

**Glucose-limited growth**

As it was said before, one of the main advantages of using biological simulations is to obtain a broad range of result data faster and with fewer costs than through laboratorial experimentation. The simulation performed in this section is an example of how these properties can be applied.

The most direct way to generate competition between different bacterial strains is to place them in substrate-limited conditions. Figure 13 displays the behavior of 4 different *E. Coli* strains, one fast-growing population and three slow-growing, subjected to similar conditions to the ones described for the isolated B/rA strain. As it had happened with the B/rA strain, the substrate concentration displays an approximately sinusoidal behavior.

The fast-growing MC1000 strain had a clear advantage over the slow-growing ones. In the initial growth phase, these individuals multiplied much faster than it happened with the rest of the strains, spreading through the simulation space as they grew, consuming almost all the substrate available. After that point, the B/rA and B/rF26 populations stabilized. The MC1000 continued to display the sinusoidal behavior, inverse to the glucose concentration, as it had happened before with the isolated B/rA strain. The B/rK *E. Coli* has the double of the size of the other strains, and is also slow-growing. It had no chance to get enough substrate to allow any population increase, and became extinct in the first population-descendant period.

After synchronizing completely with the glucose, around the 700th simulation hour, the MC1000 strain showed a tendency towards stabilizing, as did the glucose amount, displaying smaller and smaller cell number amplitudes. To carry out such an experimental work, to obtain similar results, it would be necessary to set up an experiment with constant glucose input, 4 different strains and find a way to measure the dimensions of the
population of each strain throughout the entire experiment, that would last for 55 days. By using the proposed simulation solution, all that was needed was to define the quantity of each strain present in the environment, the rate and amount of glucose input, and the displayed data was generated. Furthermore, these results were obtained after running the simulation for only 30 minutes.

**Conclusions**
The main goal of the developed application was to implement a biological simulation that accurately reproduced the behavior of the Escherichia Coli bacterium in a multi-agent environment.
The DEB theory was chosen to define the bacterial life-cycle. This was very useful for the project, since the implementation laid upon a stable background theory, with well-defined methods for recreating the desired behaviors. It is also useful for the DEB community to count on a working computer model that uses the theory and helps to ensure its accuracy and applicability in different conditions. To expedite the development process, instead of writing all the equations from the beginning, Biodroid’s Life Engine was improved and applied to the project. Although almost all the metabolic DEB equations were already written in the DEBLib component of the engine, it still lacked in the using of the state variables to trigger life-cycle events such as birth, division and death. This part of the functionality was added by the author in the course of this work.
In parallel with this part of the work, the basis for the simulation itself was developed in C#. In the end, it incorporated a fully configurable 3D multi-agent environment with a visual interface. The events defined by the DEB state variables thresholds were triggered using the ION framework, as well as the basis for the communication between objects that share the virtual world. The C++ DEBLib functionalities were integrated into the C# code, creating a simulation engine that has all the features of DEB theory connected to a virtual world with a visualization layer.
The simulation was tested in two phases, to evaluate two different features.
In a first phase, the chemotactic reactions were evaluated by comparing the results from the simulation with laboratorial experiments. The time intervals spent in each of the rotation modes were very similar to the ones reported in [22]. The paths described by the moving individuals closely resembled those defined by real cells [5]. The emergent behavior of the bacterial colony reproduced the displacement of E. Coli bacteria navigating towards a concentration gradient, without adding rules more complex than simply detecting concentration differences in the environment. A more complex algorithm would be unrealistic having in mind the simplicity of the object of study. The approach followed was very simple, when compared to the related work, where the whole protein chain responsible for the chemotactic reactions was virtually recreated. However, the results presented were just as accurate in the tests performed. The high level of abstraction means a greater effort towards parameter estimation, since the tumble probabilities are hard to directly extract from an experiment. However, in terms of performance, the simplicity of the solution takes great relevance.
In the second phase, the life-cycle of the bacteria was tested and compared to results from the literature. The bacterial growth was very similar to the laboratorial results, at an individual [24] and population [28] levels. This validates the applicability of DEB theory to similar situations and confirms the accuracy of the implemented model. The starvation tests did not come out as positive, since the population curves extracted were different from the ones obtained in [42]. Although the genetic profile of the individuals is quite different, this is not enough to account for the disparity of results.
Further work is needed to perfect this situation, in terms of adding auto-mutation functionalities and changing the DEB metabolism rules in substrate shortage. Despite this fact, the results of the substrate-limited growth simulation were very coherent.
Using the implemented solution, it was also possible to develop a set of bacterial strains, both slow and fast growing. The numeric results obtained from a relevant number of tests came very close to the results gathered from literature [21] [20] [1]. From there, 4 different simulations were carried out, to showcase the different ways to use the solution.
Generally speaking, the solution presented in this work adequately answered the problem, which was a coherent, accurate biological simulation. Tools were also created to allow the extensibility of the model to incorporate other functionalities, such as different strains and mutations. All the relevant simulation variables are exposed to the user, so that a great number of tests can be performed without the need to recompile the code. Lastly, the solution is connected to a biological simulation engine, and the respective theory, which are in constant evolution. This means that the developed tool, with minor adjustments, has the potential to be as up to date in any time the future, as it is in the current moment.

**Acknowledgements**
This work was partially financed by Programa Operacional Regional de Lisboa (QREN SI I&DT) in the framework of project Life Engine.

**References**


