Theoretical Investigation of the Role of Protein-DNA Search in Quorum Sensing Cells

A Thesis

submitted to
Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme
by

Kinjal Mondal



Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April 2022
Supervisor: Dr. Srabanti Chaudhury

All rights reserved

Certificate

This is to certify that this dissertation entitled **Theoretical Investigation of the Role of Protein-DNA Search in Quorum Sensing Cells** towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Kinjal Mondal at Indian Institute of Science Education and Research under the supervision of Dr. Srabanti Chaudhury, Professor, Department of Chemistry, during the academic year 2021-2022.

S. Chaudhury

Dr. Srabanti Chaudhury

Committee:

Dr. Srabanti Chaudhury

Dr. Apratim Chatterji

Acknowledgments

I would like to asknowledge all who have been involved in helping me write my final year MS Thesis.

The greatest contributor towards my thesis would be my supervisor Dr. Srabanti Chadhury for her excellent guidance, mentorship, and support in my research. She has mentored me for almost 3 years now. She has been a constant source of motivation and inspiration for me all throughout. She has taught me many research techniques in theoretical chemistry and biophysics, how to approach a new research problem, how to do literature review to find an interesting research problem, how to question any obvious scientific fact and think more about it. I feel quite blessed to have an advisor like her.

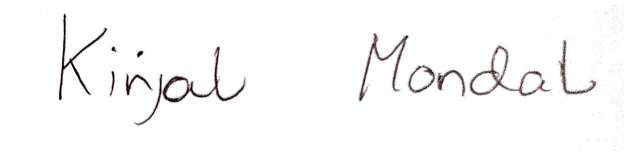
I would also like to thank my TAC member Dr. Apratim Chatterji for his courses in my 3rd and 4th years which helped me set up a strong background towards my thesis, as well as useful discussions throughout the year.

I finally would like to acknowledge my lab members of C-202 Dr. Divya Singh, Dr. Bappa Ghosh, Rugwed, Bhawakshi, Akash, Akshay, and Pankaj for making my life at lab fun. Also, I would like to thank my D Block roomates Pranay, Kshitij, Kaustaubh, Avinash, Sourajeet, and Gitesh for making my hostel life fun.

This thesis is dedicated to my Beloved Mother, Father, and Brother

Declaration

I hereby declare that the matter embodied in the report entitled **Theoretical Investigation Protein-DNA Search in Quorum Sensing Cells** are the results of the work carried out by me at the Department of Chemistry, Indian Institute of Science Education and Research, Pune, under the supervision of Dr. Srabanti Chaudhury and the same has not been submitted elsewhere for any other degree.



Kinjal Mondal

5/13/2022

Contents

1. Abstract 9

2. Introduction 10 -12

3. Theoretical and Simulational Methodologies for Protein-DNA model in Quorum Sensing 13-20

   3.1 Model 13-14

   3.2 Theoretical Methodology: Backward Master Equations 14-18

   3.3 Simulation Methodology 18-20

4. Results, Discussions, and Conclusions from the Discrete State Model of Quorum Sensing 21-35

   4.1 Search Time Phase Diagram 21-24

   4.2 Role of the dissociation and formation rate of the AHLi-LuxR complex in target search kinetics 24-25

   4.3 Role of the interior and exterior diffusion rate of the AHL on the target search kinetics 25-28

   4.4 Role of positive feedback loop 28-29

   4.5 Role of the number of cells in the medium on the search time 30-32

   4.6 Effect of the Various Kinetic Parameters Visualized Using Heat 32-34

   4.7 Discussions and Conclusions 34-35

5. Theoretical Investigation of the Transitions of Active Inertial Particles 36-45

   5.1 Introduction 36-37

   5.2 Theoretical and Simulation Methodology 37-41

   5.3 Results 41-45

   5.4 Conclusion 45

List of Figures

1. Pictorial representation of a model of protein-DNA interaction in quorum sensing cells in a typical Gram-negative bacteria 13
2. Average search time (in *min*) versus λ 23
3. Average search time (in *min*) versus λ at different values of the (a) dissociation rates of the protein complex to free AHL and LuxR and (b) rates of association of AHL and LuxR to form the protein complex. 24
4. Average search time (in *min*) versus λ at different values of the (a) rate of diffusion of AHL from the inside to the outside of the cells and (b) rate of diffusion of AHL from the outside to the inside of the cells 26
5. Pictorial representation of a model of protein-DNA interaction in quorum sensing cells in a typical Gram-negative bacteria with a positive feedback loop 28
6. (a)Average search time versus λ at different the number of cells in the medium (b) Average number of AHL molecules present in a cell versus λ at different values of the number of cells present in the medium 30
7. Heat Map Representing the search time versus λ and (a) dissociation rates of the protein complex to free AHL and LuxR and (b) rates of association of AHL and LuxR to form the protein complex 33
8. Heat Map Representing the search time versus λ and (a) dissociation rates of the protein complex to free AHL and LuxR and (b) rates of association of AHL and LuxR to form the protein complex. 33
9. Heat Map Representing the search time versus λ And the Number of Cells Present in the Medium 34
10. Schemaic Representation of Transition Path and Transition Path Time in a double

Well Potential 37

11. Theory- Simulation Comparison Plot of $\frac{p(t)}{\Omega}$ vs Ωt 41

12. Effect of the strength of active force on the transit time 41

13. Effect of inertia (effect of mass) 42

14. Effect of inertia (effect of friction) 42

15. Comparison of Actual Expression of Transit Time in the short time tail with Approximate Expression 43

16. Comparison of the Actual Average of the Transit Time Distribution(obtained by Numerical Integration of the full distribution) with the approximate expression of the average time 44

1. **Abstract**

Quorum sensing (QS) is a cellular interaction mechanism in bacteria that regulates gene expression. The search and binding of the AHL-bound LuxR-type proteins to specific sites on DNA in QS cells of Gram-negative bacteria is a complex process and we have theoretically investigated it using a simple discrete-state model. We have shown that several factors such as the rate of formation of the AHL-bound LuxR protein within the cells and its dissociation to freely diffusing autoinducer molecule (AHL), the diffusion of the latter in and out of the cells, the presence of a feedback loop and the cell population density is very important in the protein target search and can control the gene regulation processes. We also explain these observations physically. We have also performed stochastic simulations to validate our theoretical results.

We have also looked at a different problem where we calculated transit time distributions in inertial active particles. Transition path times are defined as the time taken by a particle to make a transition between two stable molecular conformers. Typically, active forces cause enhanced diffusion, superdiffusion. ATP dependent forces present in a cell disrupt the equilibrium in the environment, thereby causing active forces. We have obtained an analytical expression of the transit times for such systems and have also performed simulations to validate our analytical results. We have also found approximate expressions in the short and the long time limits. In addition to this, we have derived approximately the mean transition path time.

2. **Introduction**

The search of protein for targets on the DNA is very important in many cellular processes like DNA transcription, DNA replication, DNA unwinding[1]. This protein-DNA search is an extremely complicated phenomenon and it has been intensely investigated both experimentally and theoretically. Early experimental studies show that lac repressor finds its target much faster than the time predicted by a pure diffusion theory[2]. This has led to a series of theoretical and experimental studies for predicting the mechanism of protein-DNA search. The most famous of them is the facilitated diffusion theory that predicted that a protein searching for a target on the DNA can undergo a random walk motion on the DNA chain (1-D diffusion) along with a 3-D diffusion in the bulk[3, 4]. This phenomenon has been supported both via several spectroscopic and single-molecule experimental techniques[5, 6]. Kolomeisky et. al. have theoretically investigated the important physical features of this protein target search process through an analytical discrete state model[7]. Using this discrete state approach, several aspects of protein-DNA search in a more realistic cellular environment have been explored such as multiple targets on the DNA chain[8], the effect of heterogenous sequences[9], presence of multiple DNA conformers[10], and so on. However, despite several efforts to study these target search processes some aspects of protein-DNA search are still not well understood.

Bacterial gene expression is regulated by a cellular communication mechanism known as quorum sensing(QS). QS[11, 12] is a cellular communication mechanism in which bacteria regulate gene expression according to their neighbouring population through the use of small chemical signalling proteins known as autoinducers. QS is very importrant in various physiological processes like bioluminescence, formation of biofilms, production of antiobiotics and secretion of virulence factor[13-15]. The process of QS has been investigated extensively both experimentally and theoretically[16-23]. The bacteria produce and release small signaling proteins known as autoinducers. At a critical cellular population density, once the autoinducers attain a minimum threshold concentration $\sim 1 - 10 \mu g/ml$, they activate transcription and translation thereby causing gene expression. This mechanism of QS is true in Gram-negative as well as Gram-positive bacteria. In Gram-negative bacteria the most common signalling autoinducer is Acyl-Homoserine lactone(AHL) which was first discovered in the bacterium *Vibrio fischeri*[24]. On the other hand, Gram-positive bacteria use processed oligopeptides[11, 20]. In Gram-negative bacteria the cell walls are permeable to AHL and these signaling

molecules diffuse freely through the cell walls and their concentration increases inside the cells with an increase in the cellular population. At a minimum threshold concentration $\sim 1 - 10 \mu g/ml$[25], the AHL molecule reversibly binds to the transcriptional regulator proteins like LuxR inside the cells. This AHL-bound LuxR protein complex then finds and activates the specific promoter-target site on the DNA thus causing gene expression which can now generate various forms of cellular response like luminescence and also produce a positive feedback loop producing more amount of AHL protein[11, 20]. However, in Gram-positive bacteria, the processed oligopeptides bind to the receptors present in the cell membrane which can then phosphorylate proteins present in the cytoplasm and the phosphorylated proteins can bind to the DNA[11, 20]. Thus as evident, an important event in QS is the binding of the target on the DNA by a protein to activate gene expression. The protein-DNA target binding phenomenon in several QS systems has been explored in many experiments. The molecular mechanism of the binding of the complex to the DNA molecule for a QS system in *Sinorhizobium meliloti*(a Gram-negative bacteria) in the presence of different AHL molecules with different side chains was studied using dynamic force spectroscopy experiments and it was found that the mean lifetime of the bound protein-DNA complexes varied with different AHLs present in the protein complex[26]. In *Agrobacterium tumefaciens*(Gram-negative bacteria), the crystal structure of the transcriptional regulator TraR in complex with the autoinducer and its bound DNA were determined[27]. It has been also shown that QS antagonists which mimic the autoinducer in *Pseudomonas aeruginosa* (Gram-negative) can prevent autoinducer complex binding to the promoter thus inhibiting QS[28] and can act as potential therapeutic agents[29]. In *Bacillus cereus* (a Gram-positive bacteria), the binding of the peptide molecule PapR onto its target DNA molecule which activates the quorum-sensing regulator PICR was looked at extensively to propose a molecular activation mechanism for this regulator[30]. In bioluminescent marine bacterium *V. fischeri*, upon reaching a threshold concentration of AHL, the LuxR protein forms a complex with AHL. This LuxR-AHL complex finds and binds to the promoter on the DNA thereby initiating transcription of the luminescence luciferase operon [LuxICDABE]. This leads to an increase in autoinducer synthesis and production via the increase in transcription and translation of luxICDABE thereby establishing a positive feedback loop[31]. The effect of cellular concentrations, QS anatagonists, postive feedback loops and many more can be investigated if one can understand the microscopic details of how QS systems influence protein-DNA interactions and affect gene expression.

Our thesis is organised as follows. We develop a simple theoretical model and solve Master Equations and perform some numerical simulations based on this simple theoretical

model according to the methodologies described in Chapter 3. We then give the results from this model and conclude in Chapter 4.

In Chapter 5, we then digress to a different problem of Transit Time Distributions. We calculate analytically the full form of the transit time distributions and perform some numerical simulations to validate this analytical expression. We further investigate transit time distributions in the long and short time tails. Next, we calculate approximately the mean of the transition path times and finally conclude .

3. Theoretical and Simulational Methodologies for Protein-DNA model in Quorum Sensing

3.1 **Model**

We here describe a simple theoretical model to study protein-DNA interaction in QS in Gram-negative bacteria. To build a minimal model on such a protein-DNA target search process, we consider a system of *N* cells in the medium (say Cell 1, Cell 2, Cell 3, Cell X,…., Cell N) as illustrated in Figure 1a. We assume that there is a sufficient amount of AHL present intracellularly and extracellularly such that it can interact with the transcriptional regulator protein LuxR in the intracellular medium and form the AHL -bound LuxR protein complex. For simplicity, we also assume that all the cells are similar as they are present in the same extracellular environment. An AHL protein (AHLi) which is initially synthesized in the interior of any of the cells (e.g. Cell X), can diffuse into the extracellular medium with rate $k_{AHLi\rightarrow AHLe}$ or can also associate to a LuxR protein thus forming a AHLi-LuxR complex with a rate $k_{AHLi\rightarrow AHLi-LuxR}$. The AHL protein (AHLe) present in the exterior of the cell can reversibly diffuse into any of the cells X (X=1 to *N*) with a rate $k_{AHLe\rightarrow AHLi}$. The AHLi-LuxR complex can dissociate to form a free AHLi molecule with rate $k_{AHLi-LuxR\rightarrow AHLi}$ or it can bind to a DNA chain present inside the cell with rate k_{on} to initiate the target search process of finding the specific target. After associating with the DNA, this protein complex can dissociate from the chain with a rate k_{off} and bind to some other site or can slide on the chain to the next site with a velocity *u* as considered in a standard protein-DNA model of target search until it identifies the specific target site[7]. The DNA chain is assumed to be a highly intertwined structure lying in the bulk solution, and the protein complex usually diffuses very fast near the DNA. So once the protein complex (AHLi-LuxR complex) dissociates from the DNA chain, it does not keep track of where on the DNA chain it was previously associated. So, it is equally likely for all sites on the DNA to bind to the protein complex[7, 32, 33]. Thus one can ignore the memory associated with the location of the protein near the DNA chain when the complex dissociates from any non-specific site on the DNA. Such an assumption has been used in developing many theoretical models and explaining several experimental results[7]. The formation of AHLi from bacterial LuxI type protein is an irreversible event such that the rate of formation of the AHLi and the protein-DNA interaction will not affect each other. Since all the cells are assumed to be similar, all the above-mentioned rate constants in our model are considered to be the same for all the cells.

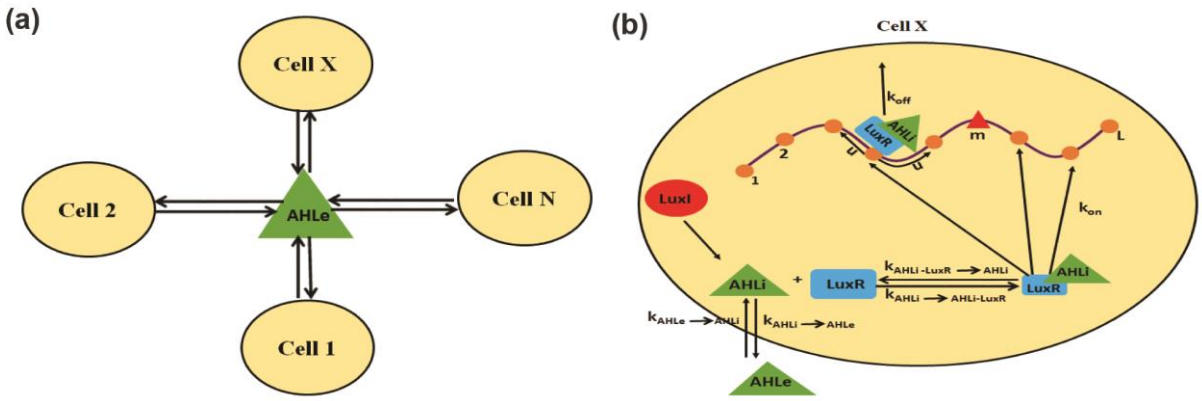


Figure 1: Pictorial representation of a model of protein-DNA interaction in quorum QS cells. There are system of N cells in the medium as shown in figure 1a. The AHL molecules can diffuse in and out of the cells. An AHL molecule present inside the cells can bind to a LuxR to form the AHL-LuxR complex which can then find the target on the DNA.

We investigate this simple protein-target search model QS systems as illustrated in Figure 1b. The DNA molecule consists of L binding sites with m being the target site. In general, there are many LuxR and autoinducer molecules within a cell. In this study, we model the activity of any one of the autoinducer molecules with one of the many LuxR molecules present inside each cell. We assume that the dynamics followed by this AHL molecule is the same as the dynamics followed by any other AHL molecule which is present that cell or any other cell. We investigate this simple theoretical model using a first passage time approach.

3.2 Theoretical Methodology: Backward Master Equations and their Solutions

We consider the simple discrete-state model figure 1 to investigate protein search dynamics in QS cells using a method of the first-passage probabilities[7-10, 32-37]. To begin with, we introduce a general probability density function (PDF)[7] $G_k^X(t)$ to reach the target site m at time t in any of the N cells for the first time if the system started initially from the state k (k=AHLi, AHLi-LuxR) in the X^{th} cell. Similarly, $G_{iDNA}^X(t)$ denotes the PDF of finding the target site at any of the cells at time t if initially, the AHL-bound LuxR complex started from the i^{th} site on the DNA (i=1 to L) at the X^{th} cell. $G_{AHLe}(t)$ is the PDF of to reach the target site at time t provided the AHL molecule initially started from the extracellular medium.

For simplicity we use the following notations that are denoted as: $AHLi \equiv A, AHLe \equiv B, AHLi - LuxR \equiv C$. All these PDF's time evolve according to the following backward master equations[38]

$$\frac{dG_A^X(t)}{dt} = k_{A\to B}G_B(t) + k_{A\to C}G_C^X(t) - (k_{A\to B} + k_{A\to C})G_A^X(t) \quad \text{for} \quad 1 \le X \le N \qquad (3.1)$$

Here the first term on the RHS of Eq. 3.1 $k_{A\to B}G_B(t)$ denotes that the AHL protein can freely diffuse from the interior of the X^{th} cell to the exterior with rate $k_{A\to B}$. The second term $k_{A\to C}G_C^X(t)$ denotes that the AHL protein in any Xth cell can bind with the LuxR present inside the cell with rate $k_{A\to C}$. The last term $-(k_{A\to B} + k_{A\to C})G_A^X(t)$ denotes that there is a (1-$-(k_{A\to B} + k_{A\to C})dt$) probability of the AHL molecule staying in the same state after a small interval of time dt.

$$\frac{dG_B(t)}{dt} = k_{B\to A}(G_A^1(t) + G_A^2(t) + G_A^3(t) + \cdots) - Nk_{B\to A}G_B(t) \qquad (3.2)$$

Here $k_{B\to A}G_A^1(t)$ denotes that the AHL protein from the exterior of the cell can go into the interior of cell 1 with rate $k_{B\to A}$, the term $k_{B\to A}G_A^2(t)$ denotes that the AHL protein state from the cell exterior can enter cell 2 with rate $k_{B\to A}$, and so on till the cell *N*. The last term denotes that there is a (1- $Nk_{B\to A}$dt) probability of the AHL molecule to stay in the cell exterior after a small interval of time dt.

$$\frac{dG_C^X(t)}{dt} = k_{C\to A}G_A^X(t) + \frac{k_{on}}{L}\sum_{i=1}^{L} G_{iDNA}^X(t) - (k_{C\to A} + k_{on})G_C^X(t) \;\; (1 \le X \le N) \qquad (3.3)$$

Here the first term $k_{C\to A}G_A^X(t)$ denotes that the AHLi-LuxR bound state in the X^{th} cell can dissociate to give free AHL and free LuxR in the X^{th} cell with rate $k_{C\to A}$ and the second term $\frac{k_{on}}{L}\sum_{i=1}^{L} G_{iDNA}^X(t)$ considers that the AHLi-LuxR bound state in the X^{th} cell can bind to the L sites on the DNA in the X^{th} cell with a rate $\frac{k_{on}}{L}$. . The last term denotes the probability (1-$(k_{C\to A} + k_{on})$dt) of the AHLi-LuxR complex to stay in the same state after time dt.

For sites at the two ends of the DNA chain (*i* = 1 and *i* = *L*)

$$\frac{dG_{1DNA}^X(t)}{dt} = uG_{2DNA}^X(t) + k_{off}G_C^X(t) - (u + k_{off})G_{1DNA}^X(t) \quad (1 \le X \le N) \qquad (3.4)$$

$$\frac{dG_{LDNA}^X(t)}{dt} = uG_{(L-1)DNA}^X(t) + k_{off}G_C^X(t) - (u + k_{off})G_{LDNA}^X(t) \; (1 \le X \le N) \qquad (3.5)$$

For all other sites,

$$\frac{dG_{iDNA}^X(t)}{dt} = uG_{(i+1)DNA}^X(t) + uG_{(i-1)DNA}^X(t) + k_{off}G_C^X(t)(t) - (2u + k_{off})G_{iDNA}^X(t) \quad (2 \le i \le m-1, m+1 \le i \le L-1, 1 \le X \le N). \qquad (3.6)$$

At *t* = 0, if the AHLi-LuxR complex is already at the target site, the search process ends immediately and then

$$G_{mDNA}^X(t) = \delta(t) \quad (1 \le X \le N) \qquad (3.7)$$

Laplace Transforming the first passage PDF's (Eqn 2.1 to 2.6), we get

$$(s + k_{A\to B} + k_{A\to C})\hat{G}_A^X(s) = k_{A\to B}\hat{G}_B(s) + k_{A\to C}\hat{G}_C^X(s) \text{ for } (1 \le X \le N) \qquad (3.8)$$

$$(s + Nk_{B\to A})\hat{G}_B(s) = k_{B\to A}(\hat{G}_A^1(s) + \hat{G}_A^2(s) + \cdots .. + \hat{G}_A^N(s)) \quad (3.9)$$

$$(s + k_{C\to A} + k_{on})\hat{G}_C^X(s) = k_{C\to A}\hat{G}_A^X(s) + \frac{k_{on}}{L}\sum_{i=1}^{L}\hat{G}_{iDNA}^X(s)\ (1 \le X \le N) \quad (3.10)$$

$$\left(s + u + k_{off}\right)\hat{G}_{1DNA}^X(s) = u\hat{G}_{2DNA}^X(s) + k_{off}\hat{G}_C^X(s)\ \ (1 \le X \le N) \quad (3.11)$$

$$\left(s + u + k_{off}\right)\hat{G}_{LDNA}^X(s) = u\hat{G}_{(L-1)DNA}^X(s) + k_{off}\hat{G}_C^X(s)(1 \le X \le N) \quad (3.12)$$

$$\left(s + 2u + k_{off}\right)\hat{G}_{iDNA}^X(s) = u\hat{G}_{(i+1)DNA}^X(s) + u\hat{G}_{(i-1)DNA}^X(s) + k_{off}\hat{G}_C^X(s) \quad \text{for} \quad (\ 2 \le i \le m-1, m+1 \le i \le L-1, 1 \le X \le N) \quad (3.13)$$

$$\hat{G}_{mDNA}^X(s) = 1 \qquad (1 \le X \le N) \quad (3.14)$$

We can solve these equations by assuming an ansatz for $\hat{G}_{iDNA}^X(s) = Ay^n + By^{-n} + C, y = \frac{s+2u+k_{off}-\sqrt{(s+2u+k_{off})^2-4u^2}}{2u}, C = \frac{k_{off}\hat{F}_C^X(s)}{s+k_{off}}$. The two constants can be found by applying the appropriate boundary and the initial conditions (3.11 and 3.14 for $1 \le n \le m$) and (3.12 and 3.14 for $m \le n \le L$). This problem can further be simplified by exploiting the symmetry of the problem which yields $\hat{G}_A^1(s) = \hat{G}_A^2(s) = \cdots = \hat{G}_A^N(s) = \hat{G}_A(s)\ , \hat{G}_C^1(s) = \hat{G}_C^2(s) = \cdots = \hat{G}_C^N(s) = \hat{G}_C(s)$ (3.15)

and $\hat{G}_{iDNA}^1(s) = \hat{G}_{iDNA}^2(s) = \cdots = \hat{G}_{iDNA}^N(s) = \hat{G}_{iDNA}(s)$ (for $1 \le i \le L$). This yields the final solution

$$\hat{G}_C^1(s) = \hat{G}_C^2(s) = \cdots = \hat{G}_C^N(s) = \hat{G}_C(s) = \frac{k_{on}S(s)D(s)(s+k_{off})}{L(s+k_{off})A(s)+k_{on}k_{off}S(s)D(s)-k_{on}k_{off}LD(s)} \quad (3.16)$$

where the new auxiliary function

$$S(s) = \frac{y(1+y)(y^{-L}-y^L)}{(1-y)(y^{1-m}+y^m)(y^{m-L}+y^{1+L-m})} \quad (3.17)$$

$$A(s) = s^3 + s^2(k_{A\to B} + Nk_{B\to A} + k_{A\to C} + k_{on} + k_{C\to A}) + s\big(k_{A\to C}Nk_{B\to A} + k_{on}(k_{A\to B} + Nk_{B\to A} + k_{A\to C}) + k_{C\to A}(k_{A\to B} + Nk_{B\to A})\big) + k_{on}k_{A\to C}Nk_{B\to A} \quad (3.18)$$

$$D(s) = s^2 + s(k_{A\to B} + Nk_{B\to A} + k_{A\to C}) + k_{A\to C}Nk_{B\to A} \quad (3.19)$$

$$\hat{G}_A^1(s) = \hat{G}_A^2(s) = \cdots = \hat{G}_A^N(s) = \hat{G}_A(s) = \frac{(s+Nk_{B\to A})\hat{F}_C(s)k_{A\to AHLi-C}}{D(s)} \quad (3.20)$$

Initially, we choose the system to be in the bulk and it corresponds to the state when an AHL molecule that is synthesized inside any one of the cells is ready to initiate the search process. These analytical expressions for these PDF's in the Laplace form allow us to obtain a full theoretical description of the target search phenomenon. They yield the average search time from the required initial states.

So, the average target search time on any cell when the initial state is an AHL molecule which is synthesized in any of the cells is given by

$$T_A = -\frac{d\hat{G}_A(s)}{ds}\Big|_{s\to 0} = T_C + \frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}} + \frac{1}{k_{A\to C}} \tag{3.21}$$

where T_C denotes the time required to find the target on the DNA within any cell starting from the AHLi-LuxR complex. Before initiating the search process, the AHL molecule which is present in the interior of any of the cells can diffuse to the extracellular medium and enter into any of the cells repeatedly, and finally bind to a LuxR. The average time involved in this process is the second term on the RHS $\frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}}$. Also the last term, $\frac{1}{k_{A\to C}}$ is the average time taken by the AHL molecule to bind directly to a LuxR to form a AHLi-LuxR complex and start its search for the target.

The average time required to perform the target search on the DNA chain within any cell starting from the AHLi-LuxR complex is

$$T_C = \lim_{s\to 0} -\frac{d\hat{F}_C(s)}{ds} = \frac{1}{k_{off}}\left(\frac{L}{S(0)} - 1\right) + \left(\frac{L}{S(0)}\right)\left(\frac{1}{k_{on}} + \frac{k_{C\to A}(k_{A\to B}+Nk_{B\to A})}{k_{on}k_{A\to C}Nk_{B\to A}}\right) \tag{3.22}$$

Physically $S(0)$ denotes the number of sites scanned on the DNA on an average during each association. Thus, $\frac{L}{S(0)}$ denotes the number of associations the AHLi-LuxR complex has to make to the DNA while trying to find the specific target site. Since the search process gets over once the AHLi-LuxR complex finds the target site, the number of associations ($\frac{L}{S(0)}$) is one more than the number of dissociations($\frac{L}{S(0)}$-1) and $\frac{1}{k_{off}}$ denotes the time spent in the 1D motion on the DNA chain before dissociating from the chain. So the time required for each dissociation event times the number of dissociations i.e. $\frac{1}{k_{off}}\left(\frac{L}{S(0)} - 1\right)$ denotes the overall time spent in a 1D motion on the DNA chain while searching for its target.

The AHLi-LuxR complex present in the bulk can directly bind to the DNA or it may dissociate before binding to the DNA chain and has to associate again to form the complex and participate in the search process. $\frac{1}{k_{on}}$ denotes the average time taken by the AHLi-LuxR complex to directly associate to the DNA chain without dissociating into free AHL and LuxR. Alternatively, the AHLi-LuxR complex can dissociate into free AHL and LuxR and the free AHL molecule may diffuse in and out of the cells repeatedly before it forms the protein complex again and binds to the DNA. The mean time associated with all these possible events is given by $\frac{k_{C\to A}(k_{A\to B}+Nk_{B\to A})}{k_{on}k_{A\to C}Nk_{B\to A}}$. So, the average time the AHLi-LuxR complex spends in the bulk

by doing a 3D search before each association to the DNA chain is the sum of these two times. With $\frac{L}{S(0)}$ number of possible associations, the total time spent in the bulk solution (in 3D search) in searching for the target site is the product of the number of association events times the time required for each association i.e. $\left(\frac{L}{S(0)}\right)\left(\frac{1}{k_{on}}+\frac{k_{C\to A}(k_{A\to B}+Nk_{B\to A})}{k_{on}k_{A\to C}Nk_{B\to A}}\right)$. One can then consider an effective binding rate

$$k_{oneffective}=\frac{k_{on}}{1+\frac{k_{C\to A}(k_{A\to B}+Nk_{B\to A})}{k_{A\to C}Nk_{B\to A}}} \quad (3.23)\ .$$

The protein (complex) in this quorum-sensing system spends a longer time in the bulk before associating to the DNA chain as compared to a regular protein-DNA target search process. This is because in QS cells the protein complex can dissociate before binding directly to the chain. This free AHL protein has to rebind to the LuxR for forming the AHLi-LuxR complex to initiate the search process. This will add to the average search time. The effect of the formation/dissociation rate of the protein complex and the diffusion rates of the AHL molecule in and out of the cells are all incorporated in this effective rate constant $k_{oneffective}$ given in Eq. 3.23 and Eq. 3.21 and 3.22 can be rewritten in terms of $k_{oneffective}$

$$T_C=\frac{1}{k_{off}}\left(\frac{L}{S(0)}-1\right)+\left(\frac{L}{S(0)}\right)\left(\frac{1}{k_{oneffective}}\right) \quad (3.24a)$$

$$T_A=\frac{1}{k_{off}}\left(\frac{L}{S(0)}-1\right)+\left(\frac{L}{S(0)}\right)\left(\frac{1}{k_{oneffective}}\right)+\frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}}+\frac{1}{k_{A\to C}} \quad (3.24b)$$

3.3 Simulation Methodology

To validate the analytical results obtained for the search time, Monte-Carlo simulations were performed. To perform the simulations, the following general methodology is adopted[39]. Firstly, the system is assumed to start from the state of the AHL molecule. Then the system evolves with time till the final target state is achieved. To evolve the particle over a particular time interval dt, the following general methodology is adopted.

Suppose the system is currently present at a particular state X. From this particular state, it can go to *n* different states: Y_1, Y_2, ….Y_n with rates $k^1_{X\to Y}$, $k^2_{X\to Y}$,…., $k^N_{X\to Y}$. So after a small interval of time *dt*, the system currently at X can be in the states Y_1, Y_2, ….Y_n with probabilities $k^1_{X\to Y}$dt, $k^2_{X\to Y}$dt,…., $k^N_{X\to Y}$dt respectively. The system can also remain in state X with probability 1-($k^1_{X\to Y}+k^2_{X\to Y}$ +….+ $k^N_{X\to Y}$)dt. To simulate this step starting from the state X, an uniformly distributed random number is generated between [0,1). Next, this [0,1) interval is separated into several subintervals each of whose length is equal to the probabilities of the various transitions. An example of such a division is ([0, $k^1_{X\to Y}$ dt), [$k^1_{X\to Y}dt, (k^1_{X\to Y}+$

$k^2_{X\to Y}$)dt),…..,[$(k^1_{X\to Y} + k^2_{X\to Y} + \ldots.+ k^{N-1}_{X\to Y})$ dt, $(k^1_{X\to Y} + k^2_{X\to Y} + \ldots.+ k^{N}_{X\to Y})$ dt), [$(k^1_{X\to Y} + k^2_{X\to Y} + \ldots.+k^{N}_{X\to Y})$dt,1)). In this case, the length of the first sub-interval equals the probability of the transition X→ Y_1 after time dt, the length of the second sub-interval equals the probability of the transition X→ Y_2 after time dt, and so on. Similarly, the length of the second last sub-interval equals to the probability of the last transition X→ Y_n after time dt. Finally, the length of the last subinterval denotes the probability of staying in the same state after time dt. Next, if the particular random number r falls exactly within any of such intervals, the particular transition is assumed to happen. The time is updated and the system is evolved until it hits the final target state.

Several realizations of this process are simulated to achieve a good enough average of the search time.

Additionally we also perform simulations to calculate the number of internal AHL molecules present in the cell at any time. To perform these simulations, we initially consider N basal LuxI(one LuxI per cell) proteins initially in the system. We also fix a time T when we would calculate the average number of AHL molecules present in the cells. We also simplify our system by taking care of the fact that our entire model of N cells could be effectively reduced to only a single cell model. This helps us in the way that we could simply get away without considering in which exact cell the AHL molecule is present but oncly considering whether the AHL molecule is present in the interior of the cells or the exterior. Then we follow the following general methodlogy to simulate the system over time.

1.First we calculate the times at which each of N basal LuxI's produce AHL's. To do so we intialise a N× $\frac{T}{dt}$ matrix consisting of only -1's. This matrix will be used to store the state of AHL molecules. Then we time evolve each of the N LuxI molecules to produce AHL's till time <T by the methodology discussed above. We store the time at which a new AHL is produced by the k^{th} LuxI in the k^{th} row of this matrix. The row number of the matrix tells us which LuxI produced the AHL and the column number (if not equal to -1) tells us at what time that AHL was produced.

2. Now as we have obtained this initial structure of the initial starting configuration of the AHL's we time evolve the AHL's till they hit the final promoter state and form a new LuxI protein(in this case we label the matrix element by the time at which the new LuxI was formed), or the time T is reached (in this case we label the matrix element by -2 if the AHL was in the interior or -3 if the AHL was in the exterior of the cell). From this matrix, we calculate the number of elements>0, which gives us the number of new LuxI's formed. If the number of new

LuxI's formed were>0(suppose k), we proceed to the next step. Otherwise we stop our simulations. Also, by calculating the number of -2 elements in this matrix, we can get an count of the number of internal AHL molecules at this time T.

3. From the number of new LuxI's formed thorugh a feedback loop, we again create a k$\times \frac{T}{dt}$ matrix consisting of only -1's. Then we store the times at which the new AHL's are produced till time<T.

4. Now if there are no new AHL's formed(basically no new -1 elements in the above matrix), we can stop our simulations. Otherwise we go to step 2.

5. Finally, we can calculate the total number of internal AHL molecules from step 2 and then divide it by the number of cells to get an average.

6. We can average over this obtained number of average internal AHL molecules over an ensemble of systems to obtain our actual average.

4. Results, Discussions, and Conclusions from the discrete state Model of QS

4.1 Search Time Phase Diagram

Our study aims to explore the interaction of the DNA-binding protein (in our case the AHLi-LuxR complex) and the DNA chain within QS cells. We investigate the protein-DNA target search process in QS systems by calculating the mean search times as a function of the scanning-length λ($\lambda = \sqrt[2]{\frac{u}{k_{off}}}$). We define the scanning length (the protein-DNA interaction parameter) that equals the average number of sites that the compex moves on the DNA chain before it detaches, i.e., the average distance that the protein molecule travels during each search cycle.

This parameter λ has been considered in several protein-DNA search problems to identify the dynamic search regimes.[7-9, 32, 33, 36]. λ changes with the variation of k_{off}. $\lambda < 1$ denotes the region of 3-D diffusion (the hopping regime) in which the protein complex on average occasionally binds to the DNA and dissociates back into the bulk before sliding on the DNA chain. When the protein-DNA interaction strength is between the length of each site on the DNA i.e. 1 and the length of the DNA chain i.e. L so that $1 < \lambda < L$ (sliding regime), the complex dissociates several times before it rebinds and rescans the DNA to find the target. The search process in this regime occurs via a combination of 1D and 3D search mechanisms and the protein can scan some length on DNA in every search cycle. When the protein-DNA interaction strength is greater than the DNA length, $L < \lambda$, (called a random walk regime) the protein on an average finds its target on the DNA by random walking on the DNA chain.

On the basis of our discrete-state model, we provide a comprehensive illustration of this target search process in quorum-sensing systems. Figure 2 represents the average target search time in the three distinct dynamic regimes depending on the scanning parameter λ.

For $\lambda < L$, $S(0) \approx \sqrt[2]{1+4\lambda^2}$ and

$$T_A = \frac{k_{oneffective}+k_{off}}{k_{oneffective}k_{off}} \frac{L}{\sqrt[2]{1+4\lambda^2}} + \frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}} + \frac{1}{k_{A\to C}} \quad (4.1)$$

Figure 2 shows that when $\lambda \ll 1$, the protein complex is unable to slide on the DNA because of high rate of dissociation k_{off} (essentially a 3D search regime), initially the target search time is independent of the interaction strength and then decrease with an increase in λ. When $4\lambda^2 \ll 1$,

$$T_A \approx \frac{L}{k_{oneffective}} + \frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}} + \frac{1}{k_{A\to C}} \quad (4.2)$$

i.e the search time is independent of λ. In this region, the complex after associating to any non-target site of the DNA from the solution will immediately dissociate from that site without hopping to any nearby site. So the mean search time does not vary with λ. As λ increases such that $4\lambda^2 > 1$,

$$T_A \approx \frac{k_{oneffective}+k_{off}}{k_{oneffective}k_{off}}\frac{L}{2\lambda} + \frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}} + \frac{1}{k_{A\to C}} \quad (4.3)$$

The relative probability that the protein will hop along the chain before it dissociates will increase. This will increase the number of sites scanned on the DNA per association event and will decrease the number of associations required to perform the target search process thereby decreasing the search time. Via the method of derivatives, it can be shown that T_C is a decreasing function of λ and reaches a minimum at $\lambda = \sqrt{\frac{u}{k_{oneffective}}}$. When $\lambda < \sqrt{\frac{u}{k_{oneffective}}}$, from the definition of λ, $k_{oneffective} < k_{off}$ which implies that the search process before attaining this minimum will be dominated via bulk diffusion. As $k_{oneffective} < u$, this minimum gets attained beyond $\lambda = 1$ where $T_A = \frac{L}{\sqrt[2]{uk_{oneffective}}} + \frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}} + \frac{1}{k_{A\to C}}$. In the sliding regime $(1 < \lambda < L)$, the AHLi-LuxR complex that assciates to the DNA, performs a random walk motion on it, dissociates from it into the solution repeatedly till it finds the target and so, is, therefore, a combination of multiple 1D and 3D search regimes. As λ increases such that $\lambda > \sqrt{\frac{u}{k_{oneffective}}}$, i.e $k_{off} < k_{oneffective}$, the 1D diffusion starts dominating over the 3D diffusion. This causes a redundant random walk behavior and the average search time increases with the simultaneous increase in the protein-DNA interaction strength.

When $\lambda > L$, (random walk regime), $S(0) \approx L + \frac{L[6m(1+L-m)-(1+3L+2L^2)]\lambda^{-2}}{6} + O(\lambda^{-3})$ and

$$T_A \approx \frac{L^2}{12u} + \frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}} + \frac{1}{k_{A\to C}} \quad (4.4)$$

In this regime, the affinity of DNA to bind to the complex is strong enough and the protein finds its target on the DNA by a random walking during one search cycle. There is a larger time taken in this regime as the protein complex repeatedly visits the same non-specific DNA sites in a given search cycle until the target is found. The mean search time here has a squared scaling with the DNA length L. Since λ is varied by changing k_{off}, the search time in this regime is independent of λ.

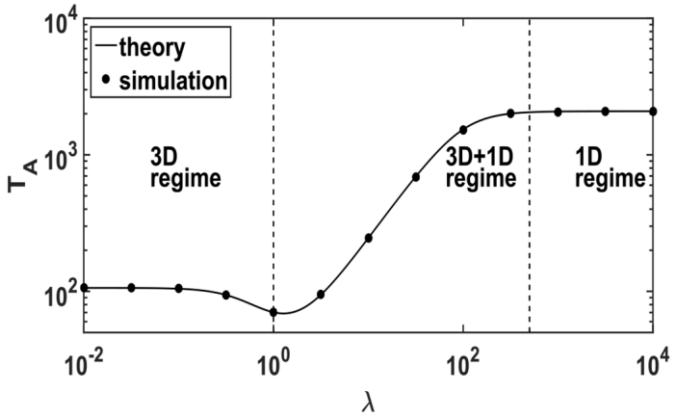


Figure 2: Average search times (in *min*) versus λ. The solid curve corresponds to our predicted analytical result. The symbols represents our stochastic simulation data. The parameters are $k_{A\to B} = k_{B\to A} = 60\ min^{-1}$, $k_{A\to C} = k_{C\to A} = 1\ min^{-1}$, $k_{on} = 10\ min^{-1} = u$, L = 500, m = 250, N = 10. The transition rates were determined from Ref. 23 in genetically modified *E. coli.*

The AHLi–LuxR complex that binds to the DNA to initiate gene expression can also dissociate into free AHL and LuxR. The former in turn can again diffuse across the cell and needs to re-enter into the cells to activate the LuxR regulators. So, it is conceivable that dissociation rate of AHL-LuxR protein complex and the diffusion of AHL molecules in and out of the cells can directly affect the interaction of the protein complex with the DNA and influence the target search process. As we have identified that there are mainly three dynamic phases involved in finding the target by the AHL molecule on the DNA (3D, 1D+3D, 1D), one would like to investigate how these factors involved in QS systems affect the target search process or equivalently the protein-DNA interaction in each of the three dynamic regimes. As evident from the analytical expression of the search times, the target search process and the different dynamic search regimes are determined by the entire mechanism of QS like the association of the AHL and LuxR to form the complex and vice versa, the reversible rate of diffusion of the autoinducer from the inside to the outside of the cell and so on. Also, it has been shown that the effect of microfluidic environments like increased amount of encapsulation by biofilms, less amount of oxygen levels, distribution of molecules that are limited by diffusion, and larger population density can strongly affect rate constants involved in this QS mechanism[23, 40-42]. This can, in turn, affect the target search process. So, it is important to study the effect of these several rate constants on target search phenomenon. Also the presence of a feedback loop is common to many bacteria that exhibit gene expression mediated by QS[31]. We would monitor the target search time as a function of all these parameters involved in the QS in the various dynamic search regimes.

4.2 Role of the dissociation and formation rate of the AHLi-LuxR complex in target search kinetics

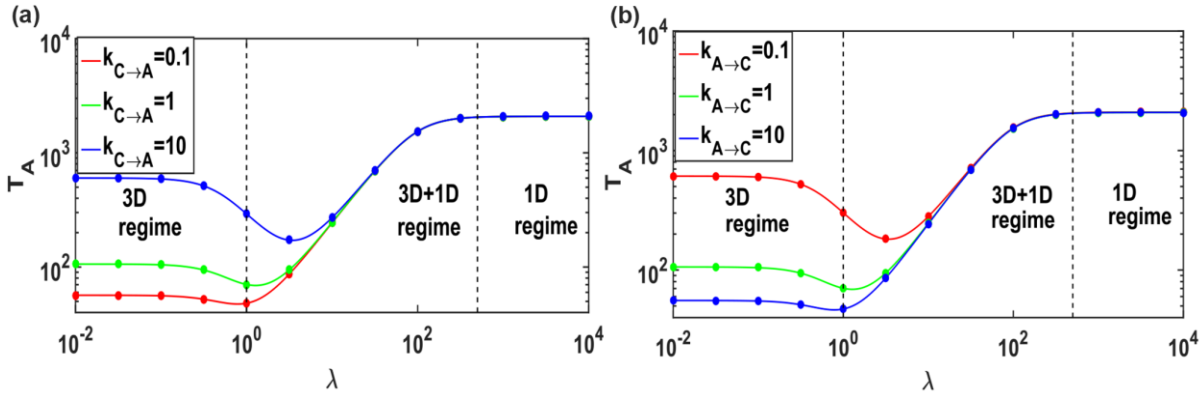


Figure 3: Average search time (in *min*) versus λ at various values of the (a) dissociation rates of the protein complex to free AHL and LuxR and (b) rates of association of AHL and LuxR to form the complex. The solid lines represent the analytic results for different values of $k_{C\to A}$ and $k_{A\to C}$(in min^{-1}) respectively whereas symbols represent Monte Carlo simulation data. In (a) $k_{A\to C}$ is kept at 1 min^{-1} and in (b) $k_{C\to A}$ is kept at 1 min^{-1}. The following set of common parameters were used for the theory and simulations: $k_{A\to B} = k_{B\to A} = 60\ min^{-1}$, $k_{on} = 10\ min^{-1} = u$, $L = 500$, $m = 250$, $N = 10$.

The transition rates were determined from Ref. 23 in genetically modified *E. coli.*

The effect of the rate of association of the AHL and LuxR to form the complex and dissociation rate of the protein complex on the search dynamics is shown in Figure 3. When the protein complex dissociates to form the free AHL and LuxR, the AHL has to associate again with the LuxR molecule to participate in the search process. Thus, increasing the rate of dissociation $k_{C\to A}$ will decelerate the search process and increase the target search time. Also, an increase in the rate of association of AHL and LuxR (keeping other parameters fixed) will facilitate the complex formation and hence will speed up the target search process as shown in Figure 3b. Since such binding-dissociation events happen only in the bulk medium, it will affect search time in the 3D regime and to some extent the 3D+1D regime. At a very small value of $k_{C\to A}$, the complex dissociates to a very small extent. A very large value of $k_{A\to C}$ implies that as soon as the complex dissociates it is immediately formed again. Thus, with a decrease or increase in $k_{C\to A}$ or $k_{A\to C}$ respectively, $\frac{1}{k_{oneffective}}$ in Eq. 3.24 converges to $\frac{1}{k_{on}}$ and the search time in Eq. 3.24b converges to

$T_A \approx \frac{1}{k_{off}}\left(\frac{L}{S(0)}-1\right)+\left(\frac{L}{S(0)}\right)\left(\frac{1}{k_{on}}\right)+\frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}}+\frac{1}{k_{A\to C}}$ at small values of $k_{C\to A}$ $\left(k_{C\to A} \ll \frac{k_{A\to C}Nk_{B\to A}}{k_{A\to B}+Nk_{B\to A}}\right)$ and

$T_A \approx \frac{1}{k_{off}}\left(\frac{L}{S(0)}-1\right)+\left(\frac{L}{S(0)}\right)\left(\frac{1}{k_{on}}\right)$ at large values of $k_{A\to C}$ $\left(k_{A\to C} \gg \frac{k_{C\to A}(k_{A\to B}+Nk_{B\to A})}{Nk_{B\to A}}\right)$.

Also at high values of $k_{C\to A}$ or very less values of $k_{A\to C}$, $\frac{1}{k_{oneffective}}$ diverges and the search time also diverges. A very large value $k_{C\to A}$ implies that the complex dissociates as soon as it is formed and a very small value of $k_{A\to C}$ implies that the initial complex formation will be difficult or if the complex dissociates it cannot associate again. Both these two cases will hinder the complex formation and its subsequent association to the DNA chain which is essential for the initiation of the target search process and hence the search time increases. Experimentally it has been shown that anti-activator proteins can act as antagonists and inhibit the binding of QS receptors such as LasR and RhlR to the autoinducers thereby delaying the expression of QS target genes in *Pseudomonas aeruginosa* cells[43]. Also lack of a functional Las system has been shown to decrease virulence in *Pseudomonas aerugonisca* in several different studies[44, 45]. One can provide a qualitative understanding of these phenomena using our simple theoretical model. The presence of QS antagonists or an insufficient amount of LasR(LuxR type) protein in *Pseudomonas aerugonisca* will reduce the binding rate of the autoinducer to the QS receptor ($k_{A\to C}$). As evident from our theory, this will influence the search dynamics in the 3D diffusion regions and the onset of the sliding phase and will slow down the average search process. This will in turn decrease the rate of virulent gene expression which is in qualitative agreement with experiments (at least in the 3D regimes of the target search).

4.3 Role of the interior and exterior diffusion rate of the AHL on the target search kinetics

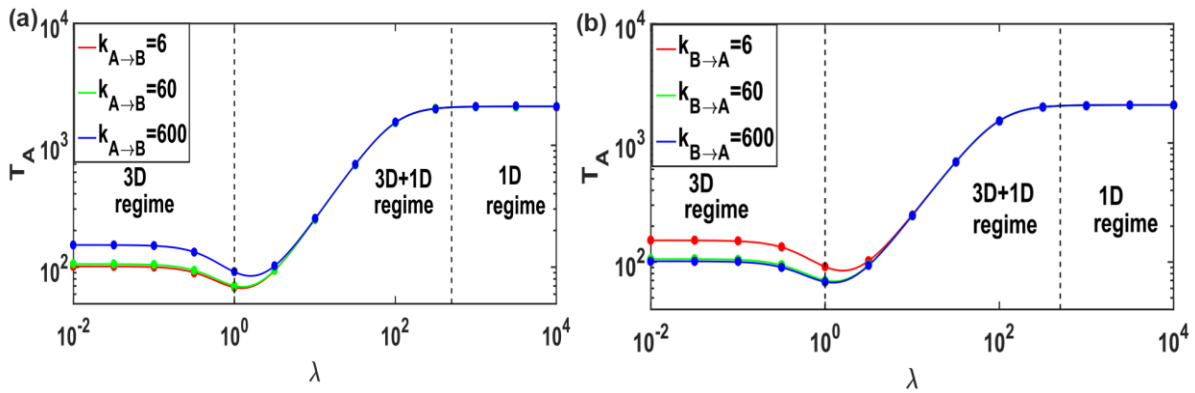


Figure 4: Mean search time (in *min*) versus λ at different values of the (a) diffusion rate of AHL from the outside to the inside of the cells and (b) diffusion rate of AHL from the inside to the outside of the cells. The solid lines correspond to the analytic results for different values of $k_{A\to B}$ and $k_{B\to A}$ respectively (in min^{-1}) while the symbols correspond to Monte Carlo simulations. In (a) $k_{B\to A}$ is fixed at $60\,min^{-1}$ and in (b) $k_{A\to B}$ is fixed at $60\,min^{-1}$. The following set of common parameters were used for the theory and simulations: $k_{A\to C} = k_{C\to A} = 1\,min^{-1}$, $k_{on} = 10min^{-1} = u$, L = 500, m = 250, N = 10. The transition rates $k_{A\to B}$, $k_{B\to A}$, $k_{A\to C}$, $k_{C\to A}$, and k_{on} were determined from Ref. 23 in genetically modified *E. coli.*

The AHL molecule initially may diffuse out of the cell, or after binding with LuxR, the AHL-bound LuxR protein complex in any of the cells may dissociate to free AHL and LuxR and this AHL molecule may diffuse to the extracellular enviornment. By increasing the rate of diffusion from the inside to the outside of the cells, the search process will slow down as the AHL molecule needs to come back to the cell interior to bind to LuxR and form the complex. If the rate at which the AHL re-enters any one of the cells increases, the complex formation within the cells will happen faster and the search process will be completed in a shorter time. Such diffusion of AHL molecule inside and outside the cells happens only in the bulk surrounding the DNA molecule and hence only the bulk search regimes will be affected (3D and the onset of the 3D+1D regime) as presented in Figure 4. Physically, a very low $k_{A\to B}$ value corresponds to the situation that the complex after dissociating into the free AHL molecule does not leave the cell. Similarly, a very high $k_{B\to A}$ value corresponds to the situation that the free AHL present in the exterior of the cells immediately re-enters any one of the cells. From Eq. 3.23, one can show that with a decrease in $k_{A\to B}$ or increase in $k_{B\to A}$, $\frac{1}{k_{oneffective}}$ converges to $\frac{1+\frac{k_{C\to A}}{k_{A\to C}}}{k_{on}}$ and the search time in Eq. 3.25b converge to

$$T_A \approx \frac{1}{k_{off}}\left(\frac{L}{S(0)} - 1\right) + \left(\frac{L}{S(0)}\right)\left(\frac{1+\frac{k_{C\to A}}{k_{A\to C}}}{k_{on}}\right) + \frac{1}{k_{A\to C}} \qquad (4.5)$$

This happens when $k_{A\to B} \ll \left(\frac{k_{A\to C}}{k_{C\to A}} + 1\right) N k_{B\to A}$ or when $\frac{k_{C\to A}k_{A\to B}}{N(k_{A\to C}+k_{C\to A})} \ll k_{B\to A}$). This search time in Eq. 4.5 corresponds to the time required for the target search process when the AHL molecule stays within the cell and does not diffuse out of the cellular medium. ($k_{A\to B}$= 0 in Eq 7).

On the other hand, $\frac{1}{k_{oneffective}}$ diverges with an increase in $k_{A\to B}$ or decrease in $k_{B\to A}$ and the search time also diverges. A very high value of $k_{A\to B}$ implies that the free AHL molecule present in the interior of one of the cells will always tend to diffuse out of the cell and a very low value of $k_{B\to A}$ implies that the AHL molecule present in the exterior of the cells will face difficulty in entering the cells. Both these situations will hinder the formation of the AHLi-LuxR complex within the cells and this will increase the search time.

Experimentally it was reported that the length and the amount of substitution of the *N*-acyl side chain of the autoinducer molecule in QS systems is important to determining whether autoinducers can diffuse freely or diffuses via an active efflux.[46] The transport of two different autoinducers, namely *N*-(3-oxododecanoyl) homoserine lactone ($3OC_{12}$-HSL-long chain HSL) and *N*-butyryl homoserine lactone (C_4-HSL-short chain HSL) was studied in *Pseudomonas aeruginosa* cells by using tritium labeled signals. These studies indicated that cellular concentration of tritium labeled C_4-HSL reaches a steady-state quickly and was nearly equal to the extracellular concentration whereas the tritium labeled $3OC_{12}$-HSL required longer to achieve steady-state cellular concentration and in the presence if efflux pumps the cellular concentrations of tritium-labeled $3OC_{12}$-HSL are 3 times higher than external levels. It was proposed that the high cellular level of tritium-labeled OC_{12}-HSL is due to the partitioning of cell membranes. Short chain HSLs' can be easily modelled in our theory by taking the rate of the AHL diffusion from the interior to the exterior of the cell to be equal to the rate of the AHL diffusion from the exterior to the interior of the cell ($k_{A\to B} = k_{B\to A}$). Long chain HSLs' can be modeled considering the rate of diffusion of the AHL molecule from the exterior to the interior of the cell to be greater than the diffusion rate from the interior to the exterior of the cell ($k_{B\to A} > k_{A\to B}$). As predicted from our theory, this would lessen the overall search time mainly in the 3D and the onset of the 3D+1D diffusive regimes. However, it has been shown in experiments that tritium labelled $3OC_{12}$-HSL would have shown even higher concentrations in the cell interior than in the extracellular environment when efflux pumps present in the cell membrane are deactivated($k_{B\to A} \gg k_{A\to B}$). Thus from our theory, it can be predicted that the efflux pump that drives the long chain HSL's from the inside to the cell exterior increases the search time more than as compared with long chain HSLs' without efflux pumps.

4.4 Role of positive feedback loop

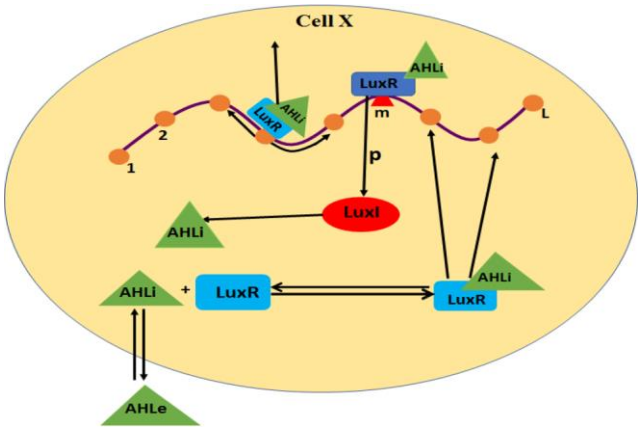


Figure 5: Pictorial representation of a model of protein target search process in QS cells in a typical Gram-negative bacteria with a positive-feedback loop. The AHL molecule can diffuse in and out of the cells. Inside the cells, the AHL molecule can bind with the LuxR to form the AHL-LuxR complex that can find the target on the DNA. After the AHL-LuxR complex finds the target on the DNA, it induces transcription and translation that can further produce LuxI protein molecules, which can produce AHL proteins.

This light production by V. fischeri is correlated with cellular population density and that is controlled by QS. The LuxR bound autoinducer complex associates to a target site on the DNA and activates transcription of *luxI*CDABE operon that enhances light production in V. Fischeri. This results in an exponential increase of both autoinducer production via the increase in transcription of LuxI and light emission via the increase in transcription of luciferase operons. This regulatory setup generates a positive feedback loop leading to bioluminescence. To understand the effect of QS–induced emission of light in Vibrio Fischeri via this postive feedback loop on the DNA- protein interactions, we extend our simple discrete state model with a feedback loop as shown in Figure 5. We assume that the target site is the LuxICDABE promoter so that when the LuxR- bound AHL complex binds to the target site (or the LuxICDABE promoter), it induces gene transcription. After this step, further translation of the same results in an increase in transcription of LuxI and produces light. The process of transcription and translation into the protein LuxI from the gene LuxICDABE is assumed to happen with a rate p. This LuxI can again produce AHL molecules which can then participate in the target search process, and the cycle continues. For sake of simplicity and illustration, we consider a cycle where one AHL molecule produces one LuxI protein via a postive feedback loop. One can calculate the time required to form a new LuxI protein as a result of this postive feedback. This time T_{total} is the mean time required to achieve the target state LuxICDABE

for the first time from the AHL molecule (T_A) in addition to the time needed to achieve the final state LuxI starting from the target state LuxICDABE.

$$T_{total} = T_A + \frac{1}{p} \tag{4.6}$$

where the process of transcription and translation into the protein LuxI from the gene LuxICDABE is assumed to occur with a rate p.

Eq.15 shows that T_{total} will have a a similar trend with respect to the scanning length parameter which is measure of our protein-DNA interaction strength as depicted by T_A (p is independent of λ). A shorter target search time would reduce the time for a feedback cycle to take place and vice versa. Thus in a given time interval, shorter DNA search time will increase the number of feedback cycles which will in turn increase the number of AHL molecules present inside the cell (either in the free state or in the bound state with LuxR) which are being produced by the LuxI. The overall enhancement in the concentration of the AHL molecules in the cell will also increase the probability of the LuxR to binds to the autoinducer which further activates LuxR-autoinducer complex to bind to a target site on the DNA and triggers the feedback loop hereby promoting the target search process. In order to validate our understanding, one can also monitor the number of AHL molecules within a cell to understand how the protein search for DNA target process is impacted by QS. Our discrete state stochastic model with positive feedback cannot provide a analytical solution to this problem but we can numerically evaluate the number of internal AHL molecules in a cell and obtain a quantitative trend of the variation of the number of AHL molecules with the variation in protein DNA interaction strength. from our theoretical study.As shown in Figure 6, one can conclude that T_{total} as well as T_A follows an inverse relation with the number of AHL molcules with respect to its reponse to the DNA-Protein interaction strength.

4.5 Role of the number of cells in the medium on the search time

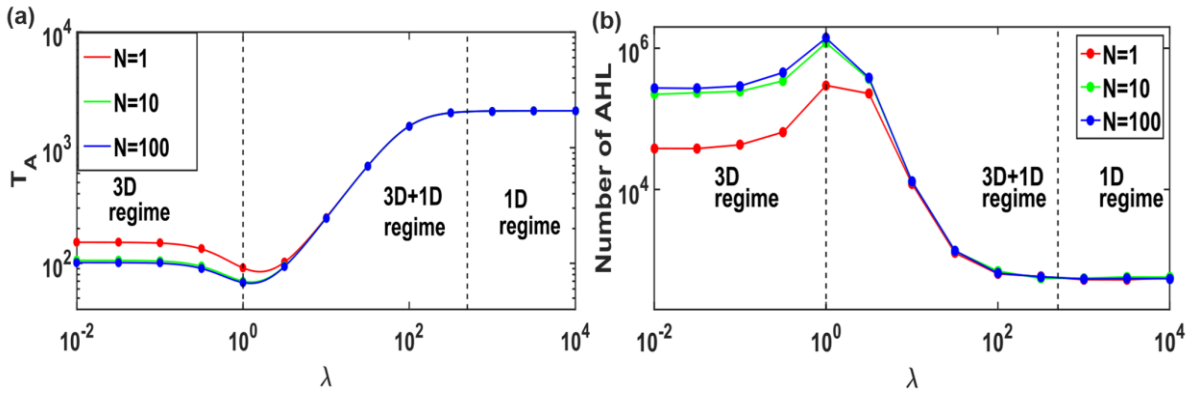


Figure 6: (a) Mean search time (in *min*) versus λ at different values of the number of cells present in the medium. Analytical results have been represented by solid lines. Monte Carlo simulations data are shown by symbols. The following parameters were chosen for the theory and simulations: $k_{A\to B} = k_{B\to A} = 60\, min^{-1}$, $k_{A\to C} = k_{C\to A} = 1\, min^{-1}$, $k_{on} = u = 10\, min^{-1}$, L = 500, *m* = 250. The transition rates were determined from Ref. 29 in genetically modified *E. coli.* (b) Average number of AHL molecules present in a cell versus the protein-DNA interaction strength λ at different values of the number of cells present in the medium. Monte Carlo simulations data are shown by symbols. The parameters chosen for simulations: $k_{A\to B} = k_{B\to A} = 60\, min^{-1}$, $k_{A\to C} = k_{C\to A} = 1\, min^{-1}$, $k_{on} = u = 10\, min^{-1}$, L = 500, *m* = 250,rate of transcription of AHL from LuxI=0.03 min^{-1}, and rate of translation of LuxI=3.08 min^{-1}.

As discussed earlier, QS is a process used by cells regulate gene expression in response to cellular population fluctuations. The QS bacteria generate signalling molecules called autoinducers that rise in concentration with increasing cellular population density. These autoinducers bind with the regulatory protein LuxR and the LuxR-autoinducer complex activates target gene transcription. So, a very important question to ask is how changes in the number of cells or the cellular population density can affect the positive feedback loop, as well as the target search time dynamics. Figure 6 shows that increasing the number of cells will decrease the target search time significantly only in the 3D search regimes. The protein complex inside one of the cells has a probability to dissociate into free AHL and LuxR and the former can diffuse to the exterior of the cell. The AHL molecule in the exterior of the cell has a probability (1-$Nk_{B\to A}dt$) to remain in the extracellular environment after some time *dt*. So with an increase in the number of cells, *N*, this probability decreases, and the AHL molecule will then diffuse quickly to one of the cells and complete the search process. Again, all these events happen only in the bulk medium surrounding the DNA and hence only the bulk search regimes (3D, initial 3D+1D regime) will be affected.

At a larger cellular population density, the free AHL molecule has a low probability to exist in the extracellular medium after time dt. So, in that case, this AHL molecule after diffusing to the exterior of the cell will immediately diffuse back into one of the cells. As *N* increases, $\frac{1}{k_{oneffective}}$ converges to $\frac{1+\frac{k_{C\to A}}{k_{A\to C}}}{k_{on}}$ (when $\frac{k_{C\to A}k_{A\to B}}{(k_{A\to C}+k_{C\to A})k_{B\to A}} \ll N$), the corresponding search time converge to Eq. 4.5 and is equal to the time needed by the AHLi-LuxR complex to

find the target in a situation in which the AHL molecule does not to diffuse into the extracellular medium. Due to the positive feedback loop, the number of internal AHL molecules in a cell increase with the cell population density, as shown in figure 6b. But LuxR production rate and the rate of production of AHL from an individual LuxR molecule are independent of the cellular environment and does not change with the cell population density[23].

Since the number of AHL molecules present inside all the cells increases proportionately with an rise in the cellular density, the overall flux of proteins onto a specific target within a specific cell is not likely to change with increasing *N*. Again, if the time taken by an individual AHL molecule to find a target in any of the cells decreases, the time after which that particular target will get activated will also decrease. This will lead to an overall decrease in the time each target waits before it gets activated. With the rise in the cellular density, the overall probability of an individual protein molecule to enter into a cell and thereby activate a target after a ceratin amount of time increases. This decreases the target search time of an AHL for a target as well as a decrease in the time each target waits before it gets activated. Thus our theoretical results verify that cell population density is indeed an important parameter in determinimg how fast the protein searches for its target on the DNA. It is also noteworthy to mention that a model consisting of only a single cell with the rate of diffusion into the cell given by an effective rate constant $k' = Nk_{B\to A}$ would give us the same results. However, the idea of including N cells is to give a more realistic physical picture to the model. Our results add a new physical dimension to the problem, which can be probed in experimental biological systems, i.e. the effect of the number of cells on the QS.

In *vibrio parahaemolyticus* population, it was observed that there is a high level of T6SS2 gene expression in a relatively more dense cellular medium and a low level of T6SS2 gene expression in a relatively less dense cellular medium[47]. This is in qualitative agreement with our theoretical predictions, which suggests that there will be an upregulation of gene expression with an increase in cell density in the 3D regime.

The effect of the formation/ dissociation rates, the diffusion rates, and the number of cells on the target search process in QS cells can also be explained mathematically. For realistic chain lengths and parameter values, $\frac{k_{A\to B}}{k_{A\to C}Nk_{B\to A}} + \frac{1}{k_{A\to C}}$ term that denotes the initial time taken by the AHL molecule to form the AHL-LuxR complex is very less than the average search time, and T_A is governed by changes in $k_{oneffective}$. Using the expression of $k_{oneffective}$ in Eq. 3.24, one finds that as $k_{C\to A}$ or $k_{A\to B}$ increases, $k_{oneffective}$ decreases, and hence the search time increases. Also increasing $k_{A\to C}$, $k_{B\to A}$, or N will increase $k_{oneffective}$ and will

decrease the average search time. As seen from Eq 3.1 and 3.2, average search time will change in the 3D and the sliding regime with changes in the $k_{oneffective}$. As shown in Eq. 3.4, in the 1D random walk regime, the average search time is independent of $k_{oneffective}$ and is unaffected by the formation and the dissociation rates of AHLi-LuxR complex, the rate of diffusion of the AHL molecule from the inside to the outside of the cells and vice versa and also the number of cells. To understand better the effect of the various kinetic parameters on the different dynamic search regimes, we have shown heat maps for the search times in below.

4.6 Effect of the Various Kinetic Parameters Visualized Using Heat Maps

The effect of the various kinetic parameters on the target search process in the process of QS can be more clearly understood by studying the heat map of the ratio of the mean search time for the target on the DNA chain for different sets of parameter values and the mean search time for a standard choice of parameter set (Figure 2). As discussed earlier, the search time increases/decreases only in the 3D and the onset of the 3D and 1D regime. As we have explained in our main manuscript with an increase in $k_{C\to A}$ and decrease in $k_{A\to C}$, the mean search time increases due to which the color changes from deep blue to orange for lower λ values whereas there is no color change for higher ranges of λ (as there is no change in the search time). Also, the change in the color is very significant only in the higher ranges of $k_{C\to A}$ and smaller values of $k_{A\to C}$ and otherwise remain constant at blue. This highlights the convergence effect of the search time at lower values of $k_{C\to A}$ and higher values of $k_{A\to C}$ which has been discussed earlier. Similar heat maps can be constructed to understand the effect of $k_{A\to B}$, $k_{B\to A}$, and the cell population density N as shown in Figure 8 and 9 respectively.

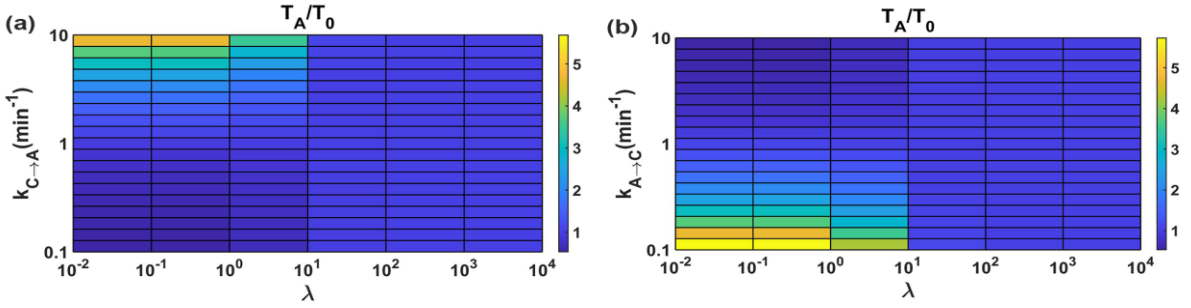


Figure 7: The heat map for the ratio $\frac{T_C}{T_0}$ is plotted against the scanning length λ and the kinetic parameter (a) $k_{C\to A}$ and (b) $k_{A\to C}$ (in min^{-1}). For calculating T_C , in (a) $k_{A\to C} = 1min^{-1}$ and (b) $k_{C\to A} = 1min^{-1}$. The following set of parameters was used for the calculation of T_0: $k_{A\to B} = k_{B\to A} = 60\ min^{-1}$. The common set of parameters are

$k_{on} = 10min^{-1} = u$, L=500, m=250, N=10.

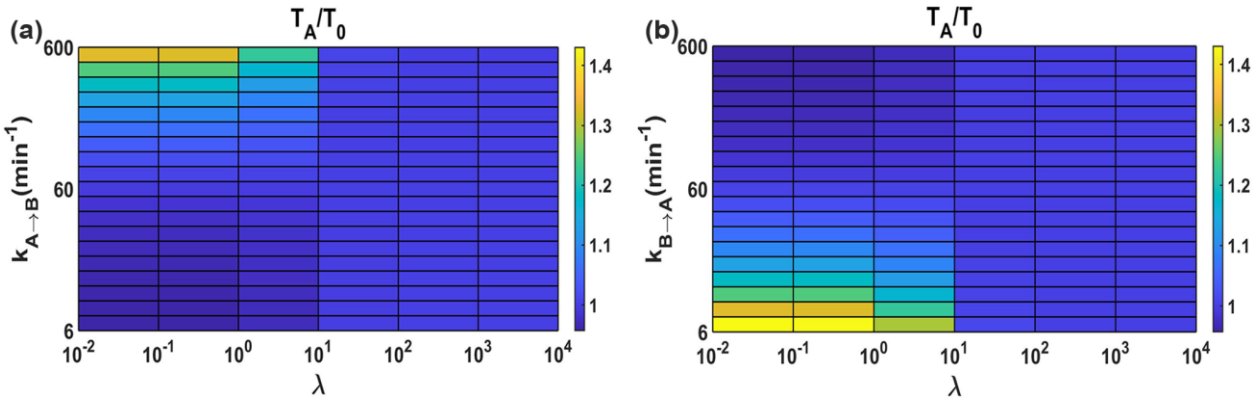


Figure 8: The heat map for the ratio $\frac{T_C}{T_0}$ is plotted against the scanning length λ and the kinetic parameter (a) $k_{A\to B}$ and (b) $k_{B\to A}$ (in min^{-1}). For calculating T_C , in (a) $k_{B\to A} = 60min^{-1}$ and (b) $k_{A\to B} = 60min^{-1}$. The following set of parameters was used for the calculation of T_0 , $k_{A\to C} = k_{C\to A} = 1min^{-1}$. The common set of parameters are $k_{on} = 10min^{-1} = u$, *L*=500, *m*=250, *N*=10.

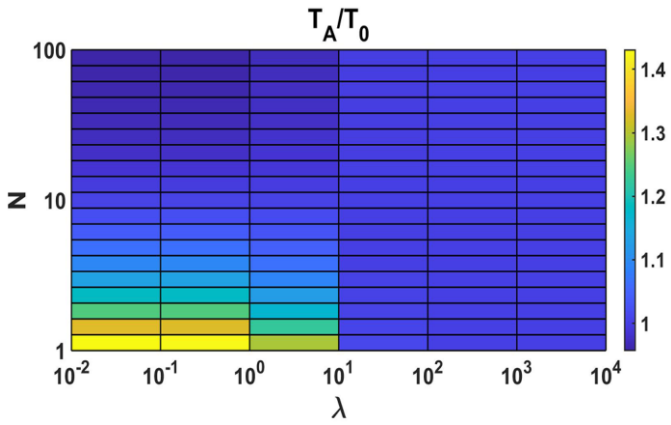


Figure 9: $\frac{T_C}{T_0}$ is plotted as a color map against the scanning length λ and the kinetic parameter *N*. The set of parameters chosen for T_C are $k_{A\to B} = k_{B\to A} = 60\ min^{-1}$, $k_{A\to C} = k_{C\to A} = 1min^{-1}$, $k_{on} = 10min^{-1} = u$, L=500, m=250. The set of parameters used for T_0 are $k_{A\to B} = k_{B\to A} = 60\ min^{-1}$, $k_{C\to A} = k_{A\to C} = 1min^{-1}$, $k_{on} = 10min^{-1} = u$, *L*=500, *m*=250.

4.7 Discussions and Conclusions

We have formulated a theoretical methodology to explicitly investigate role of QS on the protein target search process on the DNA using a discrete state model. It sheds some light on the complex mechanisms of protein target search in QS cells in Gram-negative bacteria. We observed that although the entire mechanism of QS can dictate the target search process, it effectively reduces AHLi-LuxR protein complex search for the target on the DNA. We have explored the three dynamical search regimes and concluded that the search process is the fastest

in the sliding regime. Also, these dynamical search regimes show that several factors involved in QS such as the binding rate of the AHL and LuxR to form the complex and its dissociation, the diffusion rate of the AHL to and from the cells, and the cell population density can hinder or accelerate the search process and can overall affect the search dynamics. Our results show that their effects are observed only in the 3D jumping regime and in the initial stages of the sliding motion where the 3D diffusion dominates the 1D diffusion. We also identify a saturation point beyond which the search process becomes independent of each of these factors. These results already prove that this theoretical model can be extended to incorporate more complicated biological scenarios. Our minimal model provides a qualitative understanding of previously reported experimental results without changing the basic physics underlying such search processes. However due to the non- availability of experimental data on the rate parameters in *Pseudomonas aeruginosa* and *vibrio parahaemolyticus* cells, exact comparisons cannot be made directly using our model.

We also perform stochastic simulations to support our theoretical calculations and physical-chemical arguments and can be probed in experiments.

While the presented theoretical approach provides a satisfactory minimal description of the important biophysical features of the protein-DNA search mechanism in QS cells, it is important to show its limitations. Our theoretical approach is very oversimplified and several factors like the heterogeneity of the cellular medium, the absence of crowders in the cellular medium, presence of different conformations of the AHL[48], presence of QS antagonists mimicking the autoinducer which can prevent the binding of the autoinducer complex to the DNA thus inhibiting QS[28], effects of degradation of AHL[49]and LuxR[21]are not considered at all. In addition, this study is only restricted to protein–DNA dynamics in QS cells in Gram-negative bacteria whereas the process is more complicated in Gram-positive bacteria[11]. However, despite our theoretical approach being very oversimplified, the conclusions that we have drawn from our theoretical model is generally true. The underlying physics in a more realistic QS system is the same as that of the simplified system discussed in our work.

5. Theoretical Investigation of the Transitions of Active Inertial Particles

5.1 Introduction:

In the last 3 chapters of our thesis, we investigated the problem of facilitated diffusion in QS. In this chapter we attempt to model a different type of problem in Biological Physics, the problem of Transit Path Times(TPT).

Protein Folding and other conformational rearrangements in biomolecules are often modeled by a one-dimensional diffusion in a energy landscape characterized by two minima on both sides of a maxima. Transition paths include small parts of the path followed by a particle crossing the potential barrier connecting these two minima. Single-molecule(SM) experiments recently have achieved significant capabilities to measure transitions in nucleic acids and protein-folding[50, 51]. The mean TPT for these protein-folding and their distribution were investigated by SM-FRET (Single-Molecule Forster Resonance Energy Transfer) measurements in protein folding[52, 53], DNA hairpin[54] and single-molecule force spectroscopy[51, 55, 56]. Such SM experiments allows one to monitor the transition path time that involves the duration the molecule spends in crossing the energy barrier between reactants and products[57, 58] and provides valuable information about folding-unfolding events of such biomolecules. Such experimental studies have led to a flurry of theoretical and computational studies on transition path times[59, 60]. The most common theoretical model used for analysing the dynamics of such folding events of biomolecules is the diffusion of a reaction coordinate over a one-dimensional (1D) energy landscape. Analogous to experiments, this reaction coordinate is some experimental observable, such as the mechanical extension of the molecule in force spectroscopy.

Active forces have been found to cause out-of-equilibrium randomness in cellular enviornments. Folding of biopolymers occurs in a non-equilibrium environment due to the presence of ATP dependent forces present in the cell[61]. Similar non-equilibrium processes have also been observed in cytoskeletons inside cells driven by motor proteins and cell membranes because of events like ATP hydrolysis[62]. Crystallization of active particles have also been investigated in the overdamped and the inertial limits. Several of these studies like the folding of biopolymers, and the phenomenon of crystallization involve the involve the transitions of particles from one particular stable state to the next. All these studies together point out the importance of TPT's by including active forces. Indeed, the influence of active forces on TPT's have been studied by Orland et al[60]. However, this study has been done in the overdamped regime, a limit where inertial effect are negligible compared to frictional

effects. As pointed by Orland et al, for shorter transition events, inertia may play a vital role and they have calculated the full TPT distribution along with the mean TPT[59]. These studies suggest us that it would be interesting to consider the effect of inertia along with active forces on TPT's. Recent studies have also shown that macroscopic active particles that self-propel by converting energy into mechanical motion, particles in the high Reynold's number regime, microparticles moving through a dusty plasm, colloidal particles in air, and so on are often inadequately described by overdamped Langein dynamics. Thus inertial dynamics are very much necessary to study these types of particles. Indeed, inertia coupled with active Langevin dynamics have been found to generate very interesting physics. Inertial effects in such particles have been found to cause delay in translational and orientational dynamics by bringing in memory. In addition to this, inertial effects have been found to hinder crystallization of active particles. [63, 64] Motivated by all these processes we proceed to calculate transit time distributions in the presence of active forces and inertia.

5.2 Theoretical and Simulation Methodology:

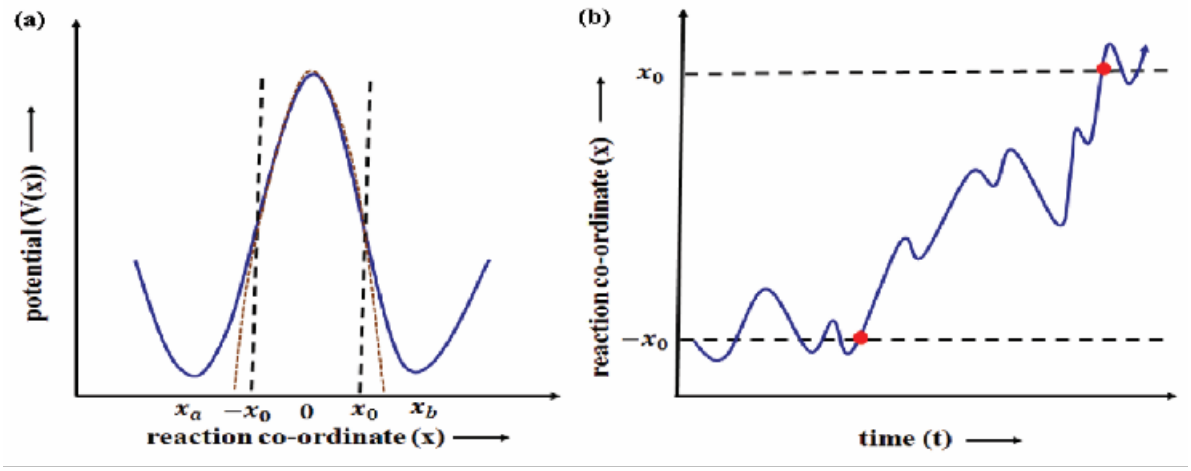


Figure 10: (a) We consider a double-well potential as shown above.The top of the double well potential is modelled as an inverted parabola. x_a and x_b denote the most stable conformers involved in the transition. The particle is assumed to make a transit from $-x_0$ to x_0.(b) Time taken for a particle to transit from $-x_0$ to x_0 giving us the transit path time.

We consider a particle performing random motion in an inverted parabola of the form $V = -\frac{kx^2}{2}$ as shown in figure 11. At time t=0, the particle starts its motion from the point $-x_0$. Transit paths are those that reach the point x_0 without ever recrossing the points $-x_0$ and x_0. To compute the transition times for these paths, we need to invoke absorbing boundary conditions at the end points. However, absorbing boundaries are very difficult to implement mathematically, as an exact analytical solution has been found in very rare cases. Instead, we

can invoke free boundary conditions at $-x_0$ and x_0 and we should expect that the solution with these free boundaries will very well approximate the solution with absorbing boundaries when the barrier height is high, i.e. $\beta E \gg 1$ i.e. $\beta = \frac{1}{k_B T}$, $E = \frac{k x_0^2}{2}$.

Now we start from the Generalized Langevin Equation of Motion[60]:

$$m\ddot{x}(t) = -\gamma \dot{x}(t) + kx(t) + \xi(t) + \eta(t) \quad (5.1)$$

where $\xi(t)$ and $\eta(t)$ are independent Gaussian Processes with mean 0 and correlation $< \xi(t)\xi(t') >= 2\gamma k_B T \delta(t - t')$ and $< \eta(t)\eta(t') >= C exp(-\left|\frac{t-t'}{\tau_A}\right|)$. Now, $\xi(t) + \eta(t)$ can be replaced by a overall Gaussian Random process R(t) with mean 0 and correlation $< R(t)R(t') >= 2\gamma k_B T \delta(t - t') + C exp(-\left|\frac{t-t'}{\tau_A}\right|)$. Now, following this equation we can proceed with calculations as has been given in Reference 10.

We Laplace transform Eqn 5.1 to obtain an expression of the average and the variance of $x(t)$ given the particle started from $-x_0$ with an initial velocity $\dot{x}(0) = v_0$. They are given by the following expressions

$$< x(t) >= \frac{m}{\sqrt{\gamma^2 + 4mK}} (e^{\lambda_+ t}(\lambda_- x_0 + v_0) - e^{\lambda_- t}(\lambda_+ x_0 + v_0)) \quad (5.2)$$

where $\lambda_+ = \frac{-\gamma + \sqrt{\gamma^2 + 4mK}}{2m}$ and $\lambda_- = \frac{-\gamma - \sqrt{\gamma^2 + 4mK}}{2m}$

and

$$< x(t) - < x(t) >^2 >= \phi^2(t) = \frac{2k_B T\gamma}{\sqrt{\gamma^2 + 4mK}} \left(\frac{e^{2\lambda_+ t} - 1}{2\lambda_+} - 2\frac{e^{(\lambda_+ + \lambda_-)t} - 1}{\lambda_+ + \lambda_-} + \frac{e^{2\lambda_- t} - 1}{2\lambda_-} + \right.$$

$$\frac{C e^{-\frac{t}{\tau_A}} \tau_A}{\tau_A(-1 + \lambda_+\tau_A)(-1 + \lambda_-\tau_A) 2k_B T\gamma} \left(\frac{e^{-\frac{t}{\tau_A}}}{\lambda_+}(1 - \lambda_-\tau_A) + \frac{e^{t\left(2\lambda_+ + \frac{1}{\tau_A}\right)}(-1 + \lambda_-\tau_A)}{\lambda_+} + \frac{e^{\frac{t}{\tau_A}}}{\lambda_-}(1 - \lambda_+\tau_A) + \right.$$

$$\left. \frac{e^{t\left(2\lambda_- + \frac{1}{\tau_A}\right)}(-1 + \lambda_+\tau_A)}{\lambda_-} - 2\frac{e^{t\left(\lambda_+ + \lambda_- + \frac{1}{\tau_A}\right)}(-2 + (\lambda_+ + \lambda_-)\tau_A)}{\lambda_+ + \lambda_-} + 2\frac{e^{\frac{t}{\tau_A}}}{\lambda_+ + \lambda_-}(-2 + (\lambda_+ + \lambda_-)\tau_A \right) + 2(-\lambda_+ +$$

$$\lambda_-) \frac{{\tau_A}^2 (e^{t(\lambda_-)}(1 - \lambda_+\tau_A) + e^{t(\lambda_+)}(-1 + \lambda_-\tau_A)}{(1 + \lambda_-\tau_A)(1 + \lambda_+\tau_A)} +$$

$$\left. 2e^{\frac{t}{\tau_A}}\tau_A \frac{(-e^{t(\lambda_-)}(1 - \lambda_+\tau_A) + e^{t(\lambda_+)}(1 - \lambda_-\tau_A)(-e^{t(\lambda_-)}(1 - \lambda_+\tau_A) + e^{t(\lambda_+)}(1 - \lambda_-\tau_A))}{(1 + \lambda_-\tau_A)(1 + \lambda_+\tau_A)} \right)$$

(5.3)

Now due to the linear structure of eqn 5.1 and since R(t) is a gaussian random process we can say that the distribution of x(t) at any particular time will be gaussian too with an average $< x(t) >$ and variance $\phi^2(t)$. This gives us the functional form of the Probability Density

Function $P(x, t/-x_0, v_0, 0)$ which is the PDF of being at position x at time t given that the particle initially started from $-x_0$ with velocity v_0. The functional form of $P(x, t/-x_0, v_0, 0)$ can be given by

$$P(x, t/-x_0, v_0, 0) = \frac{1}{\sqrt{2\pi\phi^2(t)}} \exp\left(-\frac{(x-<x(t)>)^2}{2\phi^2(t)}\right) \quad (5.4)$$

From this distribution, assuming that the initial velocities follow an equilibrium distribution, we can incorporate the equilibrium distribution and integrate over the initial velocities v_0. This gives us

$$P(x, t/-x_0, 0) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{(x-<y_0(t)>)^2}{2\sigma^2(t)}\right) \quad (5.5)$$

Where

$$< y_0(t) >= \frac{m}{\sqrt{\gamma^2+4mK}} (e^{\lambda_+ t}(\lambda_- x_0) - e^{\lambda_- t}(\lambda_+ x_0)) \quad (5.6)$$

And

$$\sigma^2(t) = \phi^2(t) + \frac{mk_BT}{\gamma^2+4mK} (e^{\lambda_+ t} - e^{\lambda_- t})^2 \quad (5.7)$$

From this function, we can obtain the absorption function $Q_A(t)$, which gives the total probability that a particle is present beyond the region $x > x_0$. So, mathematically

$$Q_A(t) = \int_{x_0}^{\infty} P(x, t/-x_0, 0)dx \quad (5.8)$$

All these trajectories have a transit time smaller than t, and so $Q_A(t)$ is proportional to the probability that the transition time is small than t. Thus, the difference $Q_A(t + dt)$-$Q_A(t)$ is thus equal to the fraction to trajectories that cross the region $>x_0$ within the small time interval dt. Then we have our transit time distribution function $p(t)$

$$p(t) \propto \frac{dQ_A(t)}{dt} \; or \; p(t) = C\frac{dQ_A(t)}{dt} \quad (5.9)$$

We note that this expression of transit time is only approximate, as we are also counting trajectories that may recross the boundaries at $-x_0$ and x_0. However, with steeper barrier heights, we expect that these recrossing will become rare and our obtained analytical expression will become exact at high barriers.

The constant C can be found by appropriate normalization condition over $p(t)$. This yields the final expression of

$$p(t) = -\frac{2}{\sqrt{\pi}} \frac{G'(t)e^{-G^2(t)}}{1-\mathrm{erf}\,(\beta^* E)} \quad (5.10)$$

where

$$G(t) = \frac{x_0-<y_0(t)>}{\sqrt[2]{2\sigma^2(t)}} \quad (5.11)$$

and

$$\beta^* E = \frac{\frac{m^2 x_0{}^2 \lambda_-{}^2}{\gamma^2+4mK}}{\frac{2}{\gamma^2+4mK}\left(-\frac{C\tau_A}{2\tau_A(-1+\lambda_+\tau_A)(-1+\lambda_-\tau_A)}\left(\frac{-1+\lambda_-\tau_A}{\lambda_+}+\frac{2\tau_A(1+\lambda_-\tau_A)(1-\lambda_-\tau_A)}{(1+\lambda_+\tau_A)(1+\lambda_-\tau_A)}\right)+\frac{C\tau_A\left(\frac{1}{\lambda_+}+\frac{\lambda_-\tau_A}{\lambda_+}\right)}{2(1+\lambda_+\tau_A)(1+\lambda_-\tau_A)}+\frac{4k_BT\gamma}{(\gamma^2+4mK)2\lambda_+}+\frac{2m\,k_BT}{\gamma^2+4mk}\right.}$$

is the effective barrier height. (5.12)

It can be mathematically shown that this expression reduces to the expression given in Reference 60 when m=0(overdamped case) and to the expression given in Reference 59 when C=0.

To validate the analytical expression obtained in the high barrier limit, we can perform numerical simulations. To perform the numerical simulations, we numerically integrate the Generalised Langevin Equation given in Equation 5.1. To perform these numerical integration, we discretise the time i.e. take a small time step h and assume that the random forces are constant over each such time interval. The Langevin equation (1) has been numerically integrated in Reference 65.[65] We quote the result directly.

$$x(t+h)(1+\gamma\frac{h}{2}) = 2x(t) - x(t-h)\left(1-\gamma\frac{h}{2}\right) + h^2(F(t)+\frac{R_+ + R_-}{2}). \quad (5.13)$$

Here R_+ denotes the total random noise in the time interval [t,t+h] and R_- denotes the toal random noise in the time interval [t-h,t]. F(t) is the deterministic force which equals kx(t) in our case.

To implement this equation we must calculate the random forces at each time interval. The total random noise is a sum of two random noises, one of them being the Gaussian white noise $\xi(t)$ and the other one being the active noise $\eta(t)$. To produce the Gaussian white noise $\xi(t)$ within any interval of time [t,t+h], we generate a Gaussian N(0, $\frac{2\gamma k_B T}{h}$) random variable. To produce the active noise, we follow the procedure given in Reference 66[66].

(1) We choose $E = e^{-\frac{h}{\tau_A}}$

(2) We choose a starting value $\eta(0) = N(0, C)$

(3) For every i^{th} time step, we choose

$\eta(i) = E\eta(i-1) + N(0, (1-E^2)C)$.

Now that we have produced both our random forces, we are now ready with our final algorithm.

(1) We initialise our particle at timer t=0 at $-x_0$. We initialise the velocity v_0 from a Gaussian N(0, $\frac{k_B T}{m}$) distribution. We also calculate the random forces produced at the time interval [t,t+h](R_0) and the deterministic force F_0=-kx_0.

(2) We then apply the velocity-velet algorithm $x_h = x_0 + v_0 h + \frac{(R_0+F_0)h^2}{2m}$

(3)Then we calculate the total random forces R_h within the time interval [h,2h], and apply Equation 5.13 to calculate x_{2h}.

(4)Now we reiterate step 3 to calculate the positions x_{3h}, x_{4h},….., and so on.

Now using this algorithm we calculate our transit time distributions. We compute the transit time distributions from the simulations with absorbing boundary conditions at x_0 and $-x_0$. We initialize the position of the particle at - x_0. We initialize the velocity from an equilibrium distribution of velocities. Then we evolve the particle till it reaches the final postion x_0. To ensure absorbing boundaries, we resimulate the particle if it enters the region $\leq -x_0$. We repeat this process until we evolve the particle to the region $x \geq x_0$ without entering into the region $x \leq -x_0$ and we record the TPT. By repeating this methodology for an ensemble, we get our entire TTD.

5.3 Results:

(a) Theory-simulation comparison:

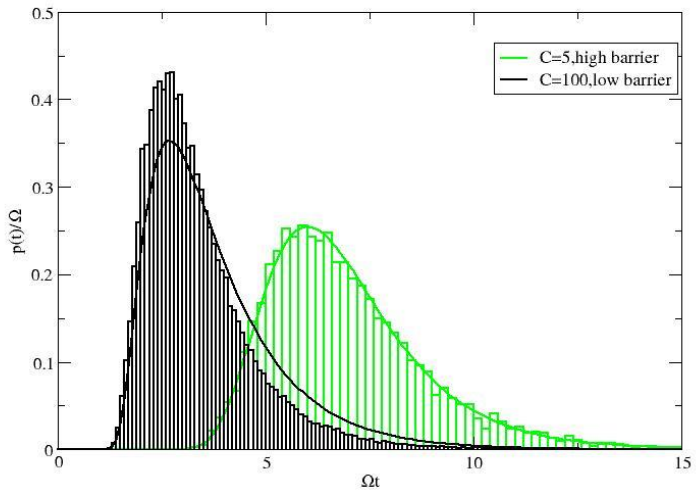


Figure 11: Plot of $\frac{p(t)}{\Omega}$ vs Ωt ($\Omega = \frac{K}{\gamma}$): Theory is given by the solid lines and the histograms denotes the data generated by performing simulations. Black Plot denotes the transit time distribution with γ=1.0, m=1, k=1, C=100, $\tau_A = 5$, x_0=10, $\beta^* E$ =0.65318, Green Plot denotes the transit time distribution with γ =1.0, m=1, k=1, C=5, $\tau_A = 5$, x_0=10, $\beta^* E$ =10.4656.

From this graph, we can see that the agreement between the theoretical plots and the simulation plots are excellent at high barrier(green plot). However there is a disagreement at a lower barrier(black plot).

(b) Effect of the strength of active force on the transit time:

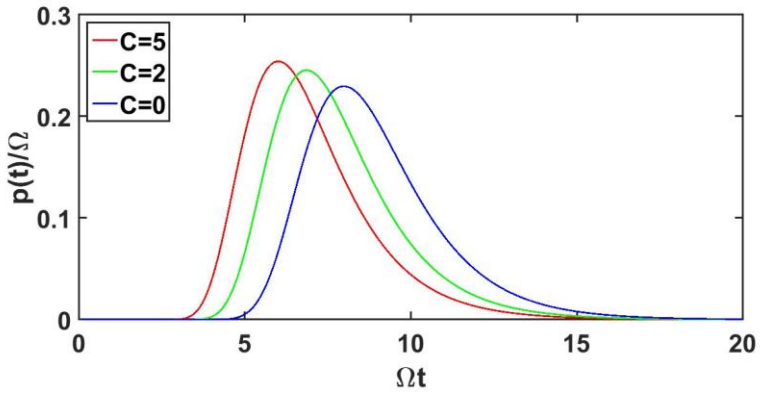


Figure 12: Plot of $\frac{p(t)}{\Omega}$ vs Ωt ($\Omega = \frac{K}{\gamma}$): Red Plot denotes the distribution with γ =1, m=1, k=1, kbT=1, C=5, $\tau_A = 5$,x_0=10, $\beta^* E$ =10.4656. Green Plot denotes the distribution with γ =1, m=1, k=1, kbT=1, C=2,$\tau_A = 5$,x_0=10, $\beta^* E$ =19.9122, Blue Plot denotes the distribution with γ =1, m=1, k=1, kbT=1, C=0,$\tau_A = 5$,x_0=10.

From the above graph, we can see that the active forces shorten the TPT. This is because if the strength of the active forces are strong, then all the active forces will be more positively correlated. All these positively correlated forces will tend to push the particle forward and shorten it's TPT's.

(c) Effect of Inertia:

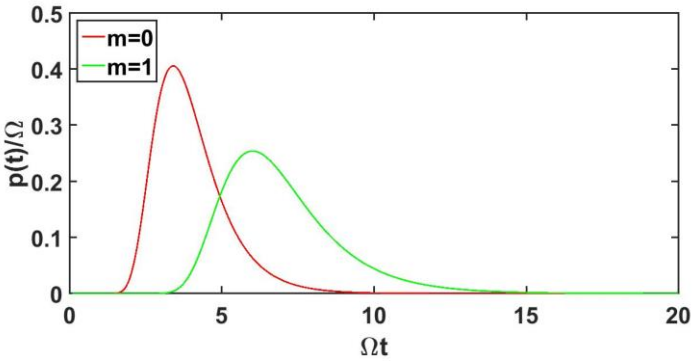


Figure 13: Plot of $\frac{p(t)}{\Omega}$ vs Ωt ($\Omega = \frac{K}{\gamma}$): Red Plot denotes the overdamped distribution(m=0) with γ =1, k=1, kbT=1, C=5, $\tau_A = 5$,x_0=10 $\beta^* E$ =9.667. Green Plot denotes the distribution with γ =1, m=1, k=1, kbT=1, C=5,$\tau_A = 5$, x_0=10, $\beta^* E$ =10.4656.

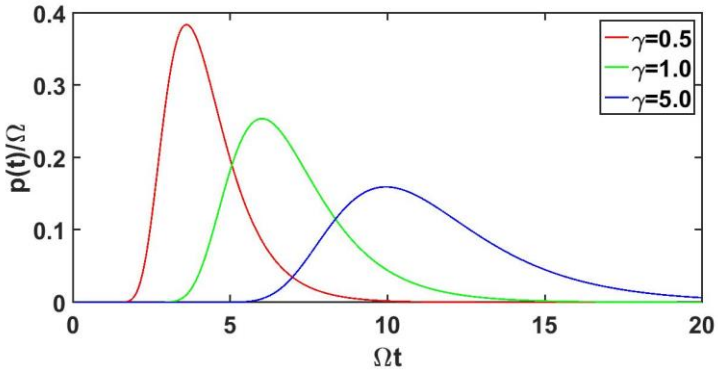


Figure 14: Plot of $\frac{p(t)}{\Omega}$ vs Ωt ($\Omega = \frac{K}{\gamma}$): Red Plot denotes the transit time distribution with γ =0.5, m=1, k=1, C=5, $\tau_A = 5$, x_0 =10, $\beta^* E$ =10.0394, Green Plot denotes the transit time distribution with γ =1.0, m=1, k=1, C=5, $\tau_A = 5$, x_0=10, $\beta^* E$ =10.4656, Blue Plot denotes the transit time distribution for the overdamped case with γ =5.0, m=1,k=1,C=5, $\tau_A = 5$, x_0=10, $\beta^* E$ =14.4812.

From these plots we can conclude that the effect of inertia is to increase the TPT's. This is because inertial effects will tend to cause the particle to feel heavy. This will make it difficult for the particle to move forward, that will increase the transition path time.

(d) Asymptotic results and Average Time:

We first get an approximate analytical expression of the short time tail of the TPT Distribution. For $t \to 0$, we note that $G^2(t)$ diverges at short times. So, the Transit time Distribution(TTD) is dominated at short times by $e^{-G^2(t)}$. To find an expression, we taylor expand $G^2(t)$ in mathematica by Taylor-Expanding each of the numerator and the denominator in Eqn 4.11 and retain the lowest powers of t. This yields us a short-time asymptote of TTD as $\lim_{t\to 0} p(t) \approx \exp\left(-\frac{2mx_0^2}{kbT} t^{-2}\right)$ (5.14). We note that this result reduces exactly to the overdamped limit result in case of $m \to 0$ and the inertial case of $C \to 0$. In the below plot we compare this asymptotic expression of TTD with the actual expression and we find that the agreement is good at shorter times.

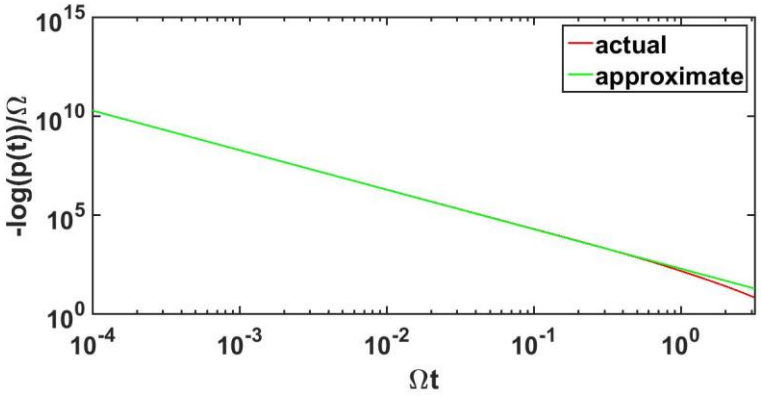


Figure 15: Plot of -Log($\frac{p(t)}{\Omega}$) (y axis) vs Ωt ($\Omega = \frac{K}{\gamma}$): Red Plot denotes the actual full distribution with γ =1, m=1, k=1, kbT=1, C=5, $\tau_A = 5, x_0$=10 $\beta^* E$ =10.4656. Green Plot denotes the approximate short time distribution with γ =1, m=1, k=1, kbT=1, C=5, τ_A = 5, x_0=8, $\beta^* E$ =10.4656

To find an expression of the average Transit Time, we need to evaluate a analytical form of the expression $\int_0^\infty tp(t)dt$.

Now, $\int_0^\infty tp(t)dt = \frac{\int_{\sqrt{\beta^* E}}^\infty t(G)e^{-G^2}dG}{\int_{\sqrt{\beta^* E}}^\infty e^{-G^2}dG}$ (5.15) (by substitution of Equation 5.10 and then a change of variables $G'(t)dt = dG$). Now, it is difficult to invert the full expression of G(t) to get t(G). However, it is possible to do so with the long time asymptotic of G(t). So, We first obtain a long time asymptotic of $G^2(t)$ at long times. This can be obtained by the procedure mentioned in Reference 59 to get $\lim_{t\to\infty} G^2(t) = \beta^* E(1 + \frac{4\sqrt{\gamma^2+4mK}}{\gamma+\sqrt{\gamma^2+4mK}} e^{-\lambda_+ t})$ (14) . Next we invert this equation and follow the same procedure as mentioned in Reference 59 to get our expression of the average Transit Time Distribution in the high barrier limit ($\beta^* E \gg 1$)

$< t_{TTD} >= \frac{1}{\lambda_+} Log(\beta^* E \frac{4e^C \sqrt[2]{\gamma^2+4mK}}{\gamma+\sqrt[2]{\gamma^2+4mK}})$ (C being the Euler-Mascheroni constant). (5.16)

We note that this analytical expression of the transit time reduces to the expression derived in Reference 60 in the overdamped case(m=0), and Reference 59 in the inertial case(C=0).

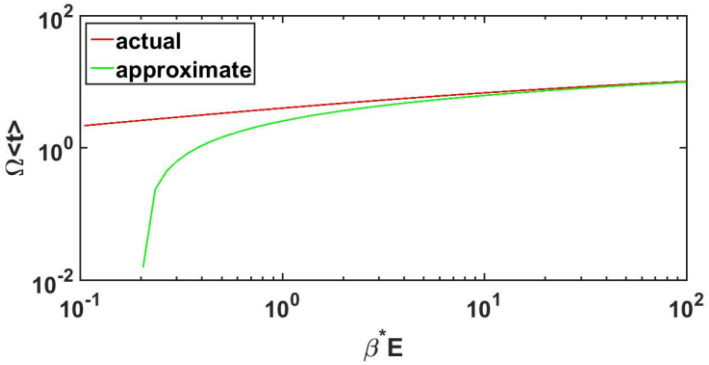


Figure 16: Plot of Ω<t> vs $\beta^* E$: Red Plot denotes the actual average from the numerical integration of the full distribution from Equation 5.15 with γ =1, m=1, k=1, kbT=1, C=5, $\tau_A = 2$. Green Plot denotes the value from the approximate average time expression (5.16) with γ =1, m=1, k=1, kbT=1, C=5,$\tau_A = 2$. x_0 is varied to vary our effective barrier height $\beta^* E$.

From the above plot we can see that our high barrier approximation works very well in the limit when $\beta^* E$>>1. Our analytical expression from equation 5.16 matches very well with the numerical integration of our full distribution from Equation 5.15 at high barriers ($\beta^* E$>>1).

5.4 Conclusion:

We have calculated an approximate analytical expression of Transit Time Distribution in the presence of active forces. We have also derived an approximate expression of the short-time tail, as well as a approximate expression of the average transit time. We validate our approximate theoretical results with exact numerical simulations. We observe that the agreement between the theory and the simulation is excellent at higher barriers. However, the agreement expectedly becomes less at lower barriers. Also, we have shown that the presence of inertia increases the average transition time in the presence of active forces. These qualitative results are in agreement with the observations made in Reference 64. Studies on crystallization made in this reference suggests us that the crystallization of active particles is more hindered in the presence of inertia as compared to the overdamped limit. Presence of inertia suppresses the effect of the self propulsion force and also widens the transition regime from the liquid to the solid, thus hindering the transitions of the particles from liquid to solid state.

Bibliography:

[1] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, P. Walter, Molecular Biology of the Cell, 6th ed., Garland Science, New York, 2014.
[2] A.D. Riggs, S. Bourgeois, M. Cohn, The Lac Represser-Operator Interaction: III. Kinetic Studies, J. Mol. Biol. 53 (1970) 401-417.
[3] M. Bauer, R. Metzler, Generalized Facilitated Diffusion Model for DNA-Binding Proteins with Search and Recognition States, Biophys. J. 102 (2012) 2321-2330.
[4] A.B. Kolomeisky, Physics of Protein–DNA Interactions: Mechanisms of Facilitated Target Search, Phys. Chem. Chem. Phys. 13 (2011) 2088-2095.
[5] S.E. Halford, J.F. Marko, How Do Site-Specific DNA-Binding Proteins Find Their Targets?, Nucleic Acids Res. 32 (2004) 3040-3052.
[6] D.M. Gowers, G.G. Wilson, S.E. Halford, Measurement of the Contributions of 1D and 3D Pathways to the Translocation of a Protein Along DNA, Proc. Natl. Acad. Sci. 102 (2005) 15883-15888.
[7] A. Veksler, A.B. Kolomeisky, Speed-selectivity paradox in the protein search for targets on DNA: is it real or not?, J. Phys. Chem. B 117 (2013) 12695-12701.
[8] M. Lange, M. Kochugaeva, A.B. Kolomeisky, Protein Search for Multiple Targets on DNA, J. Chem. Phys. 143 (2015) 105102.
[9] A.A. Shvets, A.B. Kolomeisky, Sequence Heterogeneity Accelerates Protein Search for Targets on DNA, J. Chem. Phys. 143 (2015) 245101.
[10] K. Mondal, S. Chaudhury, Effect of DNA conformation on the protein search for targets on DNA: A theoretical perspective, J. Phys. Chem. B 124 (2020) 3518-3526.
[11] C.M. Waters, B.L. Bassler, Quorum Sensing: Cell-to-Cell Communication in Bacteria, Annu. Rev. Cell Dev. Biol. 21 (2005) 319-346.
[12] N.A. Whitehead, A.M.L. Barnard, H. Slater, N.J.L. Simpson, G.P.C. Salmond, Quorum-Sensing in Gram-Negative Bacteria, FEMS Microbiol. Rev. 25 (2001) 365-404.
[13] R.P. Novick, E. Geisinger, Quorum Sensing in Staphylococci, Annu. Rev. Genet. 42 (2008) 541-564.
[14] W.-L. Ng, B.L. Bassler, Bacterial Quorum-Sensing Network Architectures, Annu. Rev. Genet. 43 (2009) 197-222.
[15] P. Williams, M. Cámara, Quorum Sensing and Environmental Adaptation in Pseudomonas Aeruginosa: a Tale of Regulatory Networks and Multifunctional Signal Molecules, Curr. Opin. Microbiol. 12 (2009) 182-191.
[16] H. Nealson Kenneth, T. Platt, J.W. Hastings, Cellular Control of the Synthesis and Activity of the Bacterial Luminescent System, J. Bacteriol. 104 (1970) 313-322.
[17] M. Schuster, D. Joseph Sexton, S.P. Diggle, E. Peter Greenberg, Acyl-Homoserine Lactone Quorum Sensing: From Evolution to Application, Annu. Rev. Microbiol. 67 (2013) 43-63.
[18] P.D. Pérez, S.J. Hagen, Heterogeneous Response to a Quorum-Sensing Signal in the Luminescence of Individual Vibrio fischeri, PLoS ONE 5 (2010) e15473.
[19] C. Lupp, G. Ruby Edward, Vibrio fischeri Uses Two Quorum-Sensing Systems for the Regulation of Early and Late Colonization Factors, J. Bacteriol. 187 (2005) 3620-3629.
[20] J. Pérez-Velázquez, M. Gölgeli, R. García-Contreras, Mathematical Modelling of Bacterial Quorum Sensing: A Review, Bull. Math. Biol. 78 (2016) 1585-1639.
[21] S. James, P. Nilsson, G. James, S. Kjelleberg, T. Fagerström, Luminescence Control in the Marine Bacterium Vibrio Fischeri: an Analysis of the Dynamics of Lux Regulation, J. Mol. Biol. 296 (2000) 1127-1137.
[22] J.D. Dockery, J.P. Keener, A Mathematical Model for Quorum Sensing in Pseudomonas Aeruginosa, Bull. Math. Biol. 63 (2001) 95-116.

[23] C.M. Austin, W. Stoy, P. Su, M.C. Harber, J.P. Bardill, B.K. Hammer, C.R. Forest, Modeling and validation of autoinducer-mediated bacterial gene expression in microfluidic environments, Biomicrofluidics 8 (2014) 034116.
[24] E.G. Ruby, Lessons From a Cooperative, Bacterial-Animal Association: the Vibrio Fischeri–Euprymna Scolopes Light Organ Symbiosis, Annu. Rev. Microbiol. 50 (1996) 591-624.
[25] A. Eberhard, A.L. Burlingame, C. Eberhard, G.L. Kenyon, K.H. Nealson, N.J. Oppenheimer, Structural identification of autoinducer of Photobacterium fischeri luciferase, Biochemistry 20 (1981) 2444-2449.
[26] F.W. Bartels, M. McIntosh, A. Fuhrmann, C. Metzendorf, P. Plattner, N. Sewald, D. Anselmetti, R. Ros, A. Becker, Effector-Stimulated Single Molecule Protein-DNA Interactions of a Quorum-Sensing System in Sinorhizobium meliloti, Biophys. J. 92 (2007) 4391-4400.
[27] A. Vannini, C. Volpari, C. Gargioli, E. Muraglia, R. Cortese, R. De Francesco, P. Neddermann, S. Di Marco, The Crystal Structure of the Quorum Sensing Protein TraR Bound to its Autoinducer and Target DNA, EMBO J. 21 (2002) 4393-4401.
[28] E.G. Suneby, L.R. Herndon, T.L. Schneider, Pseudomonas aeruginosa LasR·DNA Binding Is Directly Inhibited by Quorum Sensing Antagonists, ACS Infect. Dis. 3 (2017) 183-189.
[29] M. Malešević, F. Di Lorenzo, B. Filipić, N. Stanisavljević, K. Novović, L. Senerovic, N. Polović, A. Molinaro, M. Kojić, B. Jovčić, Pseudomonas aeruginosa Quorum Sensing Inhibition by Clinical Isolate Delftia Tsuruhatensis 11304: Involvement of N-Octadecanoylhomoserine Lactones, Sci. Rep. 9 (2019) 16465.
[30] R. Grenha, L. Slamti, M. Nicaise, Y. Refes, D. Lereclus, S. Nessler, Structural Basis for the Activation Mechanism of the PlcR Virulence Regulator by the Quorum-Sensing Signal Peptide PapR, Proc. Natl. Acad. Sci. 110 (2013) 1047-1052.
[31] M.B. Miller, B.L. Bassler, Quorum Sensing in Bacteria, Annual Review of Microbiology 55 (2001) 165-199.
[32] A. Shvets, M. Kochugaeva, A.B. Kolomeisky, Role of static and dynamic obstacles in the protein search for targets on DNA, J. Phys. Chem. B 120 (2016) 5802-5809.
[33] A.A. Shvets, A.B. Kolomeisky, Crowding on DNA in protein search for targets, J. Phys. Chem. Lett. 7 (2016) 2502-2506.
[34] K. Mondal, S. Chaudhury, A theoretical study of the role of bulk crowders on target search dynamics of DNA binding proteins, J. Stat. Mech. 2020 (2020) 093204.
[35] B. Punia, S. Chaudhury, Theoretical Study of the Role of Dynamic Bulk Crowders in the Protein Search for Targets on DNA, J. Stat. Mech. 2021 (2021) 073502.
[36] J. Shin, A.B. Kolomeisky, Molecular Search with Conformational Change: One-Dimensional Discrete-State Stochastic Model, J. Chem. Phys. 149 (2018) 174104.
[37] C. Felipe, J. Shin, A.B. Kolomeisky, DNA Looping and DNA Conformational Fluctuations Can Accelerate Protein Target Search, J. Phys. Chem. B 125 (2021) 1727-1734.
[38] N.G. Van Kampen, Stochastic Processes in Physics and Chemistry, 3rd ed., North Holland; Amsterdam, The Netherlands, 2007.
[39] L.A. Segel, L. Edelstein-Keshet, A Primer on Mathematical Models in Biology, Society for Industrial and Applied Mathematics, United States of America, 2013.
[40] S. Stewart Philip, Diffusion in Biofilms, J. Bacteriol. 185 (2003) 1485-1491.
[41] L.A. Vieira Helena, P. Freire, M. Arraiano Cecília, Effect of Escherichia coli Morphogene bolA on Biofilms, Appl. Environ. Microbiol. 70 (2004) 5682-5684.
[42] J.D. Bryers, F. Drummond, Local Macromolecule Diffusion Coefficients in Structurally Non-Uniform Bacterial Biofilms Using Fluorescence Recovery After Photobleaching (FRAP), Biotechnol. Bioeng. 60 (1998) 462-473.
[43] K.L. Asfahl, M. Schuster, Additive Effects of Quorum Sensing Anti-Activators on Pseudomonas aeruginosa Virulence Traits and Transcriptome, Front. Microbiol. 8 (2018) 2654.

[44] E.I. Lutter, S. Purighalla, J. Duong, D.G. Storey, Lethality and cooperation of Pseudomonas aeruginosa quorum-sensing mutants in Drosophila melanogaster infection models, Microbiology 158 (2012) 2125-2132.
[45] H. Zhu, R. Bandara, T.C.R. Conibear, S.J. Thuruthyil, S.A. Rice, S. Kjelleberg, M. Givskov, M.D.P. Willcox, Pseudomonas aeruginosa with LasI Quorum-Sensing deficiency during corneal infection, Invest. Ophthalmol. Vis. Sci. 45 (2004) 1897-1903.
[46] J.P. Pearson, C.V. Delden, B.H. Iglewski, Active Efflux and Diffusion Are Involved in Transport of Pseudomonas aeruginosa Cell-to-Cell Signals, J Bacteriol. 18 (1999) 1203–1210.
[47] L. Wang, D. Zhou, P. Mao, Y. Zhang, J. Hou, Y. Hu, J. Li, S. Hou, R. Yang, R. Wang, J. Qiu, Cell Density- and Quorum Sensing-Dependent Expression of Type VI Secretion System 2 in Vibrio parahaemolyticus, PLOS ONE 8 (2013) e73363.
[48] L. Soulère, M. Frezza, Y. Queneau, A. Doutheau, Exploring the Active Site of Acyl Homoserine Lactones-Dependent Transcriptional Regulators With Bacterial Quorum Sensing Modulators Using Molecular Mechanics and Docking Studies, J. Mol. Graph. Model. 26 (2007) 581-590.
[49] C.-S. Wong, C.-L. Koh, C.-K. Sam, J.W. Chen, Y.M. Chong, W.-F. Yin, K.-G. Chan, Degradation of Bacterial Quorum Sensing Signaling Molecules by the Microscopic Yeast Trichosporon loubieri Isolated from Tropical Wetland Waters, Sensors 13 (2013) 12943-12957.
[50] S. Chung Hoi, M. Louis John, A. Eaton William, Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 11837-11844.
[51] K. Neupane, F. Wang, T. Woodside Michael, Direct measurement of sequence-dependent transition path times and conformational diffusion in DNA duplex formation, Proc. Natl. Acad. Sci. U.S.A. 114 (2017) 1329-1334.
[52] S. Chung Hoi, K. McHale, M. Louis John, A. Eaton William, Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times, Science 335 (2012) 981-984.
[53] H.S. Chung, W.A. Eaton, Single-molecule fluorescence probes dynamics of barrier crossing, Nature 502 (2013) 685-688.
[54] K. Truex, H.S. Chung, J.M. Louis, W.A. Eaton, Testing Landscape Theory for Biomolecular Processes with Single Molecule Fluorescence Spectroscopy, Phys. Rev. Lett. 115 (2015) 018101.
[55] K. Neupane, A.N. Foster Daniel, R. Dee Derek, H. Yu, F. Wang, T. Woodside Michael, Direct observation of transition paths during the folding of proteins and nucleic acids, Science 352 (2016) 239-242.
[56] K. Neupane, N.Q. Hoffer, M.T. Woodside, Testing Kinetic Identities Involving Transition-Path Properties Using Single-Molecule Folding Trajectories, J. Phys. Chem. B 122 (2018) 11095-11099.
[57] G. Hummer, From transition paths to transition states and rate coefficients, J. Chem. Phys. 120 (2003) 516-523.
[58] W. E, E. Vanden-Eijnden, Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events, Annu. Rev. Phys. Chem. 61 (2010) 391-420.
[59] M. Laleman, E. Carlon, H. Orland, Transition path time distributions, J. Chem. Phys. 147 (2017) 214103.
[60] E. Carlon, H. Orland, T. Sakaue, C. Vanderzande, Effect of Memory and Active Forces on Transition Path Time Distributions, J. Phys. Chem. B 122 (2018) 11186-11194.
[61] N. Nikola, A.P. Solon, Y. Kafri, M. Kardar, J. Tailleur, R. Voituriez, Active Particles with Soft and Curved Walls: Equation of State, Ratchets, and Instabilities, Phys. Rev. Lett. 117 (2016) 098001.
[62] C.P. Brangwynne, G.H. Koenderink, F.C. MacKintosh, D.A. Weitz, Nonequilibrium Microtubule Fluctuations in a Model Cytoskeleton, Phys. Rev. Lett. 100 (2008) 118104.
[63] H. Löwen, Inertial effects of self-propelled particles: From active Brownian to active Langevin motion, J. Chem. Phys. 152 (2020) 040901.
[64] J.-j. Liao, F.-j. Lin, B.-q. Ai, Inertial effects on crystallization of active particles, Phys. A: Stat. Mech. Appl. 582 (2021) 126251.
[65] J.M. Thijssen, Computational physics, Cambridge University Press, Cambridge, 1999.

[66] H. Vandebroek, C. Vanderzande, The effect of active fluctuations on the dynamics of particles, motors and DNA-hairpins, Soft Matter 13 (2017) 2181-2191.