A comprehensive computational model to simulate TF binding in prokaryotes

Nicolae Radu Zabet and Boris Adryan

Department of Genetics and Cambridge Systems Biology Centre, University of Cambridge, UK

Introduction

Site specific transcription factors (TF) are proteins that orchestrate transcription by binding to specific target sites on the DNA. This binding can be both sequence- and conformation-specific. However, also non-specific binding with lower affinity can be observed. Once bound to the DNA the TF molecules perform an one dimensional random walk on the DNA until they either find a target site or unbind from the DNA template. In particular, during the one dimensional random walk on the DNA, a molecule will perform one of the three types of movements: (i) sliding , (ii) hopping and (iii) jumping. This combination of one and three dimensional diffusion is called *facilitated diffusion* and it is hypothesised that this speeds up the search process.

Model

Here, we present a comprehensive computational framework that allows the stochastic simulation of the search process of TFs for their target sites on the DNA. Each TF molecule is represented as an object (agent), which can move through three dimensional diffusion in the bacterial cytoplasm, but which also can bind to the DNA and perform an one dimensional random walk (see Figure 1).

Figure 1: *TF binding to the DNA. TF molecules bind to the DNA and mark several nucleotides as covered on the DNA. Two TF molecules cannot cover the same base pair on the DNA. The green position on the DNA marks the only position where the free TF molecule can bind.*

Results

First, we want to demonstrate how the molecules move on the DNA during a simulation run. Figure 2 shows an example of a random walk performed by 1 or 3 molecules on a 200 *bp* randomly generated DNA sequence. The molecules alternate the one dimensional movements (high density regions in Figure 2) with three dimensional excursions or hops (low density regions in Figure 2).

Furthermore, in the case of multiple molecules of the same TF species, the affinity and occupancy have a strong correlation, but not as good as in the case of 1 molecule (see Figure 3). This suggests that in the case of crowding and competition for DNA space, the affinity between TF molecules and DNA is not the sole determinant of the occupancy-bias. More specifically, the occupancy bias is not necessarily equivalent to the affinity landscape, in the sense that regions that are occupied most of the time are not necessarily the highest affinity ones.

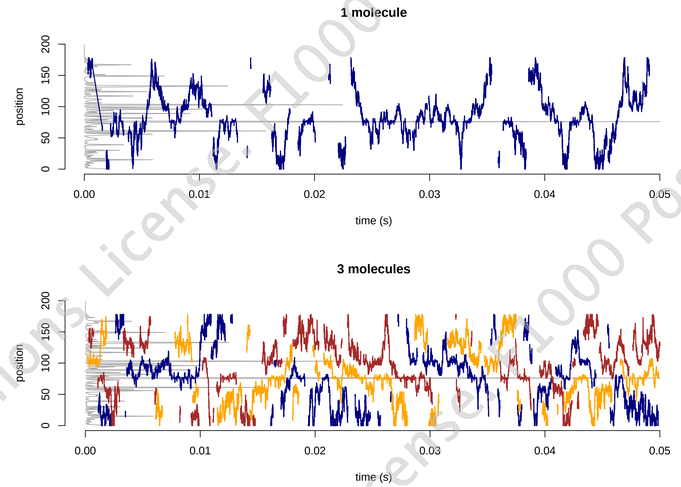

Figure 2: *Dynamic Behaviour of TF molecules.*

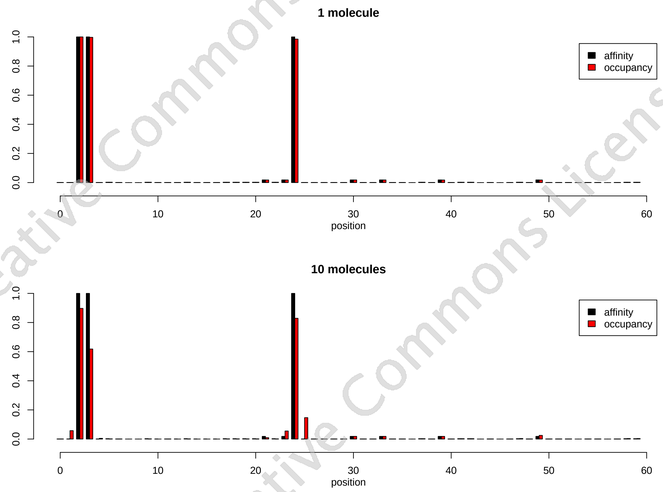

Figure 3: *Affinity vs Occupancy.*

We systematically investigated the accuracy of the proposed method to estimate model parameters. Figure 4 confirms that our proposed approach leads to the results of simulations deviating only negligible from the ones predicted mathematically. In particular, we considered four parameters: (i) observed sliding length, (ii) residence time and (iii) proportion of time bound to the DNA, (iv) 1D diffusion coefficient. Figure 4 shows that only the residence time has higher variability, but the average residence time of multiple (or longer) simulations matches well the value computed analytically. This variability can be reduced by running the simulations for longer times, which would lead to more accurate averages for the measures.

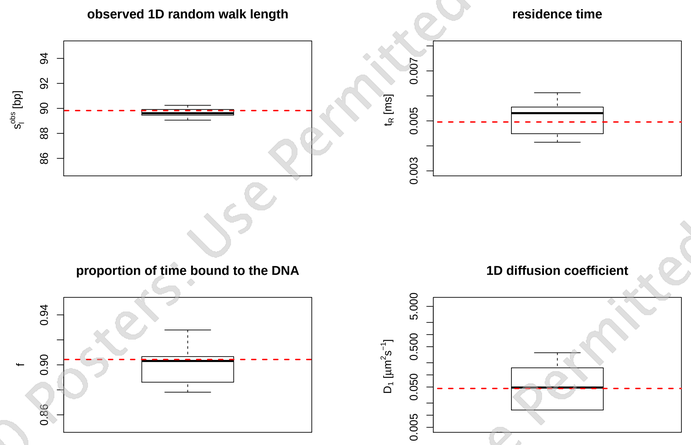

Figure 4: *Validation of our parameters estimation approach.*

Finally, we also compared the one dimensional diffusion coefficient from our simulations to the one proposed in (Elf et al., 2007) and we found that our simulations reproduce the value proposed by Elf et al. (2007) with negligible error.

Conclusions

Our model represents an ideal entry point for stochastic simulations on transcription factor target finding in prokaryotic systems. In addition, we provide an implementation in Java 1.6 of this computational model, which is called GRiP and is available at `http://logic.sysbiol.cam.ac.uk/grip/`. GRiP represents a new and efficient implementation of the TF search process which simulates at least ≈ 4 times faster than previous software (Chu et al., 2009; Barnes and Chu, 2010, 2011) and, consequently, allows for simulation of bigger systems for longer time intervals.

Publications:

- Zabet, N. R. & Adryan, B. (2012), 'GRiP: a computational tool to simulate transcription factor binding in prokaryotes', *Bioinformatics*. doi: 10.1093/bioinformatics/bts132
- Zabet, N. R. & Adryan, B. (2012), 'A comprehensive computational model of facilitated diffusion in prokaryotes' , *Bioinformatics*. doi: 10.1093/bioinformatics/bts178

Acknowledgements: This work was supported by the Medical Research Council [N.R.Z.]. B.A. is a Royal Society University Research Fellow.