**Project Specification**

Self replication is impo. Explain why.

It happens when the product of a reaction is a template for more reactions

Team have devised a reaction sequence that is capable of self-replicating, called Replication by Handhold Mediated Strand Displacement.

The process derives steps first introduced in two processes called Toehold Mediated Strand Displacement and Handhold Mediated Strand Displacement.

Toehold mediated strand displacement (TMSD) is a molecular mechanism which allows an engineer control over the kinetics of molecular arrangement [dr ouldridges paper]. With the use of TMSD, the occurrence of specific steps in molecular machinery can be defined.

Similarly, Handhold mediated strand displacement (HMSD) is a reaction mechanism for producing out-of-equilibrium products [javis paper].

Before processes are created in the lab, it is beneficial for an engineering to simulate them and observe their thermodynamic favourability. The simulation code used to simulate TMSD was a coarse grained dna model called oxDNA [cite original ox dna or tmsd paper?]. The same simulation code is currently being used to simulate copying by HMSD [Do I need to cite?]. This paper aims to use the oxDNA to simulate replication by HMSD.

**Preliminaries and oxDNA**

The simulation of replication by HMSD requires the simulation of several distinct steps, therefore, familiarity with the simulation code, oxDNA, and its capabilities is necessary. The exercises in the oxDNA primer by A. Sengar and T. E. Ouldridge were selected as practice [cite].

* What was done so far chronologically (labelled as ex1, ex2 … etc)

**Exercise 1 – Understanding the simulation techniques used to probe the oxDNA model**

Deciding on the simulation technique is the first step to simulating any reaction. The following exercise highlighted the differences between two main simulation techniques known as Molecular Dynamics (MD) [D. Frenkel and B. Smit, Understanding Molecular Simu- lation: From Algorithms to Applications, 2nd ed. (Aca- demic Press, Orlando, FL, 2002).] and Virtual-move Monte Carlo (VMMC) [106, 107 from Adityas paper]. MD algorithms don’t sample based on probabilities like Monte Carlo (MC) algorithms [N. Metropolis and S. Ulam, Journal of the American Statistical Association 44, 335 (1949).], but rather use Newton’s laws of motion for simulation. VMMC on the other hand, follows closely the principles followed by MC (probabilistic sampling) but rather than following the movements of individual particles, follows the movement of a ‘cluster` of particles. Although both methods are time consuming, especially at larger nucleotide lengths since more particles are available for sampling, by sampling ‘clusters` rather than individual particles, MC’s potential drawback of extremely long equilibration times can be reduced.

The following figures(a-c) show the equilibration of MD vs VMMC for nucleotide sequences of 20, 100, 1000.

[insert figures]

From this empirical experiment, it was observed that the fraction of time VMMC took to equilibrate, as compared to the overall runtime, was longer than the fraction of time required for the MD equilibration, however, VMMC was still approximately 15 times faster for a nucleotide strand of size 20 and 4 times faster with a nucleotide of size 1000. This highlights the suitability of VMMC for shorter nucleotide strands and longer simulation times and MD for longer nucleotide strands and shorter simulation times.

**Exercise 2 – The application of force to DNA strands in oxDNA**

Force can be applied to DNA strands to observe their response under tension. The role of stacking in a DNA strand’s stress response was also observed. A nucleotide sequence is stacked when the stacking force is non-zero and thus the stacking sites are aligned and not anti-parallel [make a figure illustration and include]. It was found that when nucleotides in a DNA sequence are stacked, the force applied will initially go into breaking the stacked configurations and then into ‘extending` the nucleotide sequence. Naturally, nucleotides stack with different forces when in a DNA strand. In the older version of oxDNA [do I need a citation?] all the nucleotides were stacked with the same force regardless of which base pair they represented with the force being an average of all the forces between all the combinations of base pairs. In the newer versions of oxDNA [do i need a citation?] it is possible to define each specific stacking force just as it naturally occurs, and it happens that the A-A nucleotide stacking force is higher than the average. That is why in Figure (shown below), the real sequence of a poly(dA) (red) required more force to extend the same amount as the poly(dA) sequence simulated using an older version of oxDNA (blue).

Two other sequences with different stacking forces were also simulated to observe their response under tension: one being a poly(dA) sequence with zero stacking force (black) and another with an exaggerated stacking force (yellow). It was observed that when the nucleotides were not staked, all the force exerted went into the extension of the poly(dA) sequence, hence a smooth increase in length is observed. On the other hand, when the stacking forces were exaggerated to a value of 4.709, from 1.709, almost all the exerted force has gone into breaking the stacked configurations and in the given time, the poly(dA) has not extended as much as the strands with less stacking forces.

**Exercise 3 – The thermodynamics of duplex formation and umbrella sampling**

Most biological reactions of interest will involve a process like duplex formation, and it is important to understand the underlying mechanism of the reaction as well as the sampling techniques involved. In a duplex formation reaction, the strands will spend most of their time in separated states, therefore, sampling them in bound states requires many long hours of simulation. To speed up, and increase the quantity of, sampling bound states, a technique called umbrella sampling [G. M. Torrie and J. P. Valleau, J. Comp. Phys. 23, 187 (1977).] can be used. Umbrella sampling is biasing a reaction around a state where sampling is desired so that more of that state is naturally sampled. After collecting the desired number of samples of states, the simulation can then be un-biased, and the relative real results can be found. This technique saves time in sampling as an engineer does not need to wait until a nucleotide attaches onto another nucleotide after long periods of time spent unbound.

Umbrella sampling was performed on a duplex of 8 nucleotides and the free energy, F(Q), is shown in figure (below).

[figure]

The F(Q) was calculated using the following formula:

[formula]

Where C is a Q-independent constant, and peq(Q) is the probability of observing Q base pairs in equilibrium [Adityas paper]. The shape of the graph tells us that there exists a large energy gap from being in a state where no bases are formed to going to a state where one base is formed. The consequent base pairs become more and more favourable except for the final pair which is slightly less favourable due to a phenomenon called fraying [T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, Journal

of Chemical Physics 134, 085101 (2011).][ J. SantaLucia and D. Hicks, Annual Review of Bio-

physics and Biomolecular Structure 33, 415 (2004).].

An extension of this exercise was fine tuning the weights of the sampling so that the biased sampling of all the base pairs formed, led to an equal number of times each number of pairs was sampled.

**Introduction to replication by Handhold Mediated Strand Displacement**

Handhold Mediated Strand Displacement (HMSD), whether in copying (no template created as product) or replication (template for more HMSD reaction created as product), uses the Toehold Mediated Strand Displacement (TMSD) mechanism (figure of TMSD).

Copying by HMSD uses the TMSD (Toehold binding and branch migration) (figure of copying by HMSD) in combination to handhold binding and handhold detachment. Replication by HMSD also follows a similar pattern in which TMSD is followed by the sequence of copying by HMSD (figure of replication by hmsd). To simulate replication by HMSD, the steps involved had to be broken down and each step simulated individually.

The step involved along with relevant exercises are shown in table 1. (A table of [step number, description, relevant exercise, how the exercise will be different])

**Implementation Plan**

* Steps already done.
* Plan in gant chart?
* Superscript in gant chart with number and paragraphs explaining a more detailed plan of the implementation of each step.

**Evaluation – how to know project has been a success**

**References**