**2 Background**

**2.1 Structure of the Chromosome**

To fully appreciate the discussion on epigenetic modifications and the simulation work performed in this project, it is important to understand how the DNA is stored within a (eukaryotic) cell. The DNA is located within the cell nucleus and is organised into highly compact structures called chromosomes. At the elementary level, the DNA is packaged into nucleosomes similar to a bead-on-a-string structure (see Figure XX). The core of each nucleosome is the histone octamer, which contains two copies of four kinds of histone proteins (H2A, H2B, H3, H4). The octamer is wrapped around by roughly 150 base pairs (bp) of DNA sequence. There are ~50 bp of DNA between two nucleosomes and the H1 histone protein binds to this region. The nucleosomes organise themselves into higher order structures by associating with each other to form the chromatin fibre. The fibre further compacts itself by proposed mechanisms such as forming loops and rosettes (cite). It is only during cell division (metaphase specifically) that the fibre condenses itself into the well-known X-shape structure to prepare for the separation of the genetic materials.

**2.2 Epigenetic Modifications**

As mentioned in the introduction, epigenetic modifications are heritable changes associated with gene expression without modifying the underlying sequence of the DNA. These modifications tend to be reversible, meaning that they can be added and removed from the genes throughout the lifetime of the cell. Two well-known types of epigenetic modifications are DNA methylation and histone modifications. Both have significant influence on the biological functions of a cell by regulating its genetic expression – whether a particular gene should be active or inactive.

DNA Methylation

DNA methylation refers to the addition of the methyl group (CH3) to a particular nucleotide, the basic unit of the genetic code, of the DNA. Most commonly, methylation occurs to the cytosine within the cytosine-guanine (CG or CpG) pair in the genetic sequence. Studies have shown that CpG methylation is associated with various biological processes including transcriptional repression, X-chromosome suppression in female mammals,

Histone Modifications

Histone modifications, which is the type of epigenetic modification of interest in this project, refer to the biochemical modifications that are applied to the histone proteins in the nucleosomes. As with other proteins, histone proteins are composed of amino acids, which are the basic building blocks any protein molecules. Each histone molecule has an extended tail which is composed of a thread of amino acids. Different functional groups can be bind to the amino acids of the tail to modify the histone protein. The three common types of modifications observed are acetylation (addition of CH3CO), methylation (addition of CH3), and phosphorylation (addition of PO43-). These modifications allow information to be stored in addition to the genetic sequence and have been thought as forming a “histone-code” which

The addition and removal of functional groups on histone tails are carried out by different classes of enzymes. Specifically, the addition and removal of acetyl group are done by histone acetyltransferases (HATs) and histone deacetylases (HDACs), while the addition and removal of methyl group are done by histone methyltransferases (HMTs) and histone demethylases (HMDs).

**2.3 Establishment and Maintenance of Epigenetic Marks**

The fundamental question which remains to be understood is how certain epigenetic patterns are established during development and how these patterns are faithfully inherited from one generation of cells to another. The establishment and maintenance of epigenetic patterns are particularly important for multi-cellular organisms (eukaryotes), where all cells have identical genome but have different functionalities.

Although a lot of the models

**3 Methodology**

**3.1 Simulation Model**

In the project, we model a chromatin fibre as a semi-flexible “bead-and-spring” polymer of N beads [cite]. In line with common mappings employed in modelling chromatin dynamics [cite], each bead has a size σ = 30 nm and represents roughly 3 kbp, which corresponds to around 15 nucleosomes. Each bead is also assigned a “colour” q to represent a particular epigenetic modification. Unless otherwise stated, we assume there are three colours (q = 1, 2, 3), corresponding to the three common modifications: acetylated, unmarked, and methylated.

**3.1.1 Modelling Dynamics of Chromatin Fibre**

We simulate the dynamics of the chromatin fibre using Brownian dynamics simulation. The interactions among the beads are governed the several potentials.

The time evolution of each bead (say the i-th bead) in the fibre is governed by the following Langevin equation

where m is the mass, γ is the friction coefficient, kB is the Boltzmann constant, T is the temperature, and η is a noise vector with the following statistical properties:

where the Latin indices represent particle indices and the Greek indices represent Cartesian components. For simplicity, we set m = 1 for all beads. We also set γ, kB, and T all equal to 1.

**3.1.2 Modelling Epigenetic Modifications**

We simulate the epigenetic modifications on the chromatin fibre (i.e. a recolouring step) base on the 1D model proposed by Dodd et. al. [dodd2007]. This model, in turn, is based on the observations that there are modifying and de-modifying enzymes (HATs, HMTs, HDACs, HDMs) which change the epigenetic marks on the histone tails (see section 2). In each recolouring step, we conduct N attempts of colour conversion such that each bead, on average, receives a single conversion attempt. The procedure of a specific conversion attempt is as follows:

1. A bead to be modified is first selected from the fibre. It has a probability to undergo a recruited conversion attempt (Step 2A) and a probability to undergo a noisy conversion (Step 2B).

2A. In the recruited conversion attempt, another bead is selected at random from the beads that are within the cut-off distance R­ from . The colour of is then changed one step towards that of . More precisely, the rules are as follows:

* if , is changed or
* if , is changed or
* if (unmarked) or , remains the same

2B. In the noisy conversion attempt, is changed one step towards another state such that when there is no recruited conversion (, there is, on average, an equal number of beads in each of the three states.

The recolouring rules do not allow the direct conversion between q = 1 and 3. This is to model that any existing modifications has to be de-modified before another modification can be applied, in line with the observations that there are de-modifying enzymes (HDACs and HDMs) and modifying enzymes (HATs and HMTs).

The recruited conversion models the positive feedback mechanism suggested in section 2, where a reader enzyme for a particular epigenetic mark has the ability to recruit a writer enzyme for the same mark and modify the neighbour fed . On the other hand, the noisy conversion models the activity of free modifying and demodifying enzymes that can change the modification

As detailed in the paper by Dodd et al, the key parameters which govern the epigenetic landscape in this model is ratio of the probability of a recruited attempt to that of a noisy attempt, which they referred it as the feedback-to-noise ratio . T

**3.2 Initialisation Procedure**

The initialisation procedure is as follows unless otherwise stated. T

**3.2 Programme Structure and Language Used**

To ensure that