Using Differential Adhesion to Control Self-Assembly and Self-Repair of Collections of Modular Mobile Robots

Peter Ottery

Doctor of Philosophy
Institute of Perception, Action and Behaviour
School of Informatics
University of Edinburgh
2005

Abstract

This thesis presents a novel distributed control method which allows a collection of independently mobile robotic units, with two or three dimensional movement, to self-assemble into self-repairing hierarchical structures. The proposed method utilises a simple model of the cellular adhesion mechanisms observed in biological cells, allowing the robotic units to form virtually bonded aggregates which behave as predicted by Steinberg's differential adhesion hypothesis.

Simulated robotic units based on the design of the subaquatic HYDRON module are introduced as a possible platform on which the model can be implemented. The units are used to carry out a detailed investigation of the model behaviour and parameter space focusing on the two main tasks of rounding and sorting in both two and three dimensions. These tasks assess the model's ability to reach a thermodynamically stable configuration when the aggregates consist of either a single population of units or multiple populations of units with differing adhesive properties. The results are analysed in detail with particular attention given to the role of random movements in determining the overall performance, and demonstrate that this model provides a very robust solution to these complex tasks.

Finally, a possible extension of this work is presented in which the original model is combined with a genetic regulatory network controller. The performance of this composite is evaluated, and the benefits of this hybrid approach, in which a powerful control system manipulates a robust self-organising behaviour, are discussed.

Acknowledgements

I would like to thank my supervisor, John Hallam. I first met John during my undergraduate degree at Edinburgh and was delighted to accept a PhD position with him when I graduated. John's vast knowledge in many areas has been a great resource which I have made good use of during my studies. He has always been ready to offer me guidance but has also allowed me the freedom to explore my own ideas. Thanks John, I really appreciate all your help and patience.

I also owe a great deal of thanks to Tim Taylor, who worked alongside me in the lab throughout my PhD. Tim was always on hand to discuss ideas and has given me plenty of helpful advice and suggestions. Tim also made a significant contribution to this thesis which includes the joint work we carried out toward the end of our research. Thanks Tim, I am very grateful for all the time you gave me over the years.

During my studies I also benefited from Edinburgh University's involvement with the EU funded HYDRA project. I was lucky enough to attend a number of interesting and stimulating meetings at each of the institutes involved in the project and I would like to express my appreciation to all those that I met and who made me so welcome during my visits. I would like to express a special thank you to the Edinburgh members – Ali Yener Boztas and George Konidaris (and Tim Taylor again!) who were always happy to chat and share ideas.

Of course I must also thank all the staff and students of the IPAB group at Edinburgh University where I have been based during my studies. There are too many people for me to name individually but I express my thanks to you all. I have made many good friends who I will stay in contact with for years to come.

Finally, I would like to thank those people who have made this thesis possible by supporting me in so many ways. I am extremely grateful to my parents for all the encouragement they gave me during my school years. In particular my thoughts are with my father who was so keen to see me succeed but who is sadly not here to see me complete this work. I also owe a great deal of thanks to my mum and her partner for their more recent and continued support.

Most of all I would like to thank my partner Vicky, who has patiently supported me since this thesis began and who has helped (and encouraged) me to bring it to comple-

tion. She has kept me sane and happy and I am extremely grateful for everything she has done. Of course I must also thank her for her dedicated proof reading of this work which was definitely above and beyond the call of duty!

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Peter Ottery*)

To my dad
David Frank Ottery

Table of Contents

Chapter 1

Introduction

In recent years there has been increasing interest in multi-robot systems, largely due to the higher level of robustness and adaptability that they offer in comparison to traditional monolithic architectures. Most commonly, the envisaged systems are composed of many simple homogeneous robots. Therefore, if any single robot fails any other robot may take its place and the overall performance of the system remains unaffected. However, the clear challenge in these systems is coordinating the whole population to accurately assemble into the formation or structure required for a given task.

If the number of robots in a system is small, it is possible that some form of hierarchical or central control may be suitable. However, as the number of robots increases, the huge number of interactions that must be considered quickly make such an approach infeasible. In addition, this type of approach reduces overall reliability by introducing a single point of failure. Instead, to produce more scalable solutions some method of distributed control is required.

In distributed solutions the individual robots all execute the same simple behaviour. However, through the resulting complex series of dynamic interactions some global pattern emerges. The problem is that designing this type of system is nontrivial as it is often hard to determine the local interactions which will produce the desired global behaviour. Therefore, the vast majority of current approaches take their inspiration from natural systems where distributed control is common (e.g. organisation of atoms, social insects etc.).

The approach we present in this thesis, which is also inspired by nature, is to at-

tempt to model some of the organisational properties of biological cells. Cells demonstrate an ability to rearrange themselves in extremely complex ways to produce remarkably complicated organisms. Furthermore, they achieve this through interactions with only their local environment[1]. Therefore, the behaviours that cells employ should provide a good basis for self-assembly in multi-robot systems.

The most basic organisational behaviour exhibited by cells is generated by the adhesion forces which exist between them. This behaviour can adequately be explained by the 'differential adhesion hypothesis' (DAH) [54] which shows how simple hierarchical patterns can be formed simply due to the differences in the adhesion forces between different cell types (see section 1.3). This thesis presents a novel model of the cell structure and adhesive mechanisms which provide the basis for this robust self-organising behaviour. The model is designed to operate on physically realistic robotic units with only local sensing and communication abilities. In itself, this model provides a mechanism for generating simple self-repairing configurations. However, a biologically inspired control mechanism is also introduced to allow dynamic behaviour where the population must reorganise in response to some environmental stimuli. Finally, it is proposed that this combined model provides a suitable framework to which other cell behaviours may be added, thus increasing the level of organisation that can be achieved.

1.1 Thesis Overview

This thesis concerns itself with the following four objectives:

- Developing a robust distributed control method which utilises an abstracted model of the cell adhesion mechanisms observed in biological cells.
- Investigating the model's ability to generate the full range of behaviours the DAH predicts.

[1] Diffusing morphogens do allow a cell's local environment to be influenced by other distant cells. However, this type of long range signalling can also be approximated by signal propagation which requires only local communication (see section 1.4.1.2).

- Investigating how the model's performance relates to specific system parameters and hardware configurations.

- Integrating some higher level of control with the model to allow dynamic behaviour and provide a framework to which more complex cell behaviours may be added.

The remainder of this chapter provides the necessary background to the problem. The following two sections (see sections 1.2 & 1.3) introduce the properties of biological cells and present the DAH in more more detail. The final section presents a review of the literature which relates to the thesis, both in the field of self-organising robotic systems and simulated cellular models.

Chapter 2 begins by identifying the main constraints of the problem. This is followed by a detailed description of the proposed model. Each component of the model is introduced individually and related back to the corresponding biological equivalent where appropriate.

By implementing the model on simulated robotic modules (A-Cells) it has been possible to carry out a wide variety of experiments. The results of these experiments are analysed and discussed in chapters 3, 4 and 5. The chapters have been ordered with respect to the complexity of the configurations that the A-Cells generate. Chapter 3 investigates 2D rounding where the A-Cells form approximate disc configurations. Chapter 4 investigates the more complex 2D sorting case where the A-Cells are divided into two distinct subsets with differing adhesive properties. More specifically this chapter compares the performance of the model with the behaviour predicted by the DAH. The final chapter in this experimental portion of the thesis (see section 5) investigates the model's performance when the A-Cells have full 3D movement.

An identified failing of the basic cell adhesion model is the inability of the virtual aggregates to change function over time or in response to some environmental stimuli. In chapter 6 this issue is addressed by integrating the model with a flexible genetic regulatory network controller (GRN). The proposed approach allows the GRN to dynamically adjust the adhesive properties of the individual A-Cells. This chapter introduces the GRN controller and describes the details of integration process. The

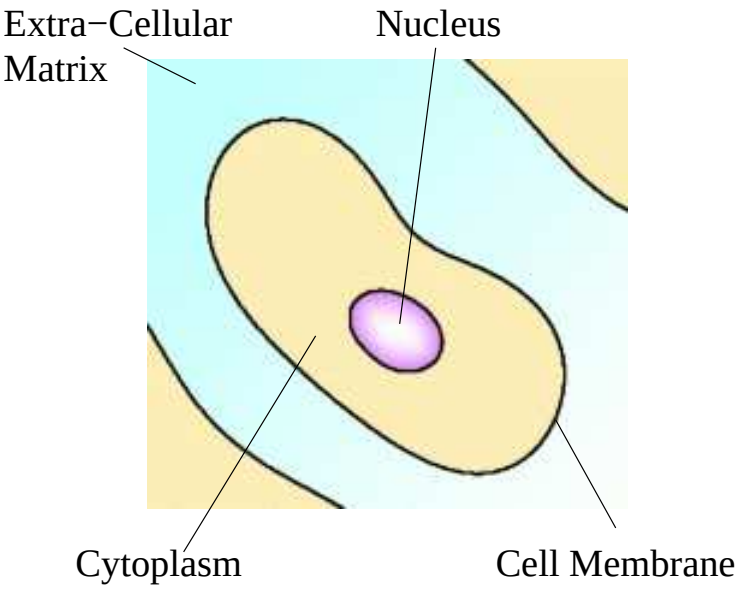


Figure 1.1: *The basic structure of a eukaryotic cell.*

resulting hybrid model is then evaluated on a number of tasks and its performance is compared with that of a GRN controller without the added cell adhesion model.

A summary of the main achievements of this thesis is presented in chapter 7 along with a brief discussion of future work which could be considered.

Details of a number of algorithms which have been developed to aid the initialisation and processing of the experiments carried out in this thesis are included in appendix A. Appendix B presents a number results for experiments which are summarised in the main body of the text. The final appendix includes publications which accompany this work.

1.2 Introduction to Cells

Biological cells form the basic building blocks of every living thing[2]. Therefore, unsurprisingly they are extremely diverse and vary greatly in both shape and size (examples range from human fibroblast cells at only $15\mu m$ to ostrich eggs [38, Chap. 1.]). However, despite this variation cells do share some similarities in structure and function.

[2]This excludes viruses which arguably may or may not be classed as living things.

All cells are essentially watery sacs which are protected from the environment by a semi-permeable membrane. This cell membrane[3] allows the inside of the cell to remain chemically distinct from the outside (extracellular matrix) while still allowing certain molecules to pass between them. Therefore, it is possible for the cells to both influence and be influenced by their environment.

The structure inside the membrane classifies cells as one of two types; prokaryotic or eukaryotic. Prokaryotic cells, which include single celled bacteria, are the oldest and simplest form of cell. These cells have little internal structure with all the internal components mixing freely. However, despite this, there is sufficient organisation to support all the bio-chemical reactions of life. Eukaryotic cells (see fig. 1.1) comprise all members of the plant and animal kingdom (including fungi) [46, Chap. 4.], existing as unicellular yeasts or 'glued' together by cell adhesion molecules (CAMs) to form the tissues of complex multi-cellular organisms. These cells are thought to be more recent in origin [60, Chap 4.] and are both larger and more complex than the prokaryotes. Their internal space contains numerous compartments called organelles which create micro-climates within the semi-fluid matter (cytoplasm) contained within the cell's membrane. The organelles allow specialised chemical reactions to take place and the nucleus, which is the largest organelle in eukaryotic cells, contains the cell's genetic material (DNA). In addition to these, the cytoplasm of eukaryotic cells also contains a cytoskeleton. This is a network of three-dimensional protein fibres which criss-cross their way through the cell's internal space. Some of these fibres simply provide reinforcement to the cell's rather weak structure, while other more dynamic fibres actively push or pull the cell membrane into particular shapes and also assist in moving the contents of the cytoplasm.

Clearly, for either of these cell types to survive they must be able to perform various key functions. For example, all cells must be able to respond to changes in their environment. Most cells must be able to divide and some cells must be able to move. However, most importantly, cells must be able to produce energy from ingested foods (respiration and fermentation) and recycle or expel the waste products. Without this energy cells would not be able to combine simple amino acid building blocks into the

[3] Also known as the plasma membrane.

complex proteins on which most of the structural and functional properties of the cell depend.

1.3 Differential Cellular Adhesion

It has been observed that when aggregates of certain cell types are mixed together the cells do not remain distributed at random and instead a degree of sorting takes place, causing the cells to form hierarchical patterns of coherent homeotypic cells.

Steinberg [54] first demonstrated this sorting behaviour with three separate experiments involving tissues obtained from a chick embryo.

- Exp. 1: limb-bud cells intermixed with heart cells.
  - Result: the limb-bud cells migrated centrally and became completely engulfed by the heart cells.
- Exp. 2: heart cells intermixed with liver cells.
  - Result: the heart cells migrated centrally and became completely engulfed by the liver cells.
- Exp. 3: limb-bud cells intermixed with liver cells.
  - Result: the limb-bud cells migrated centrally and became completely engulfed by the liver cells.

These experiments showed a clear transitivity in the central migration of the cells. However, the true significance of the result was that the hierarchical relationship which had been demonstrated was the same type of relationship known to occur in non-living immiscible liquids [17]. When such liquids are intermixed, the demixing that occurs is driven by surface tension forces arising from the difference in adhesion between the separate liquid molecules. This therefore led Steinberg to postulate that a similar mechanism was driving the cell sorting process and thus he proposed the “differential adhesion hypothesis” ([54], reviewed in [16, 55]).

Steinberg's hypothesis states that cells will rearrange to form the most thermodynamically stable configuration. This means the cells will attempt to minimise their free energy. To illustrate this, consider a single cell. The surface has a fixed size and thus a fixed area available for bonding to other cells. As adhesions are formed, the free bonding area and thus the free energy that remains (unused energy in the form of adhesions which haven't been formed, but could be) will diminish.

With an aggregate of cells, those at the boundary of the aggregate will have the highest unsatisfied bonding potential as they have neighbours only on one side. This gives rise to the surface free energy of the aggregate or, alternatively, the surface tension, which is a measure of this energy per unit area or equally the reversible work required to expand the surface by unit area. To minimise this surface free energy, the aggregate should round to form a shape with the lowest surface area to volume ratio – a sphere.

This same notion of energy can also be used to explain the more interesting case where two tissues are intermixed. If we name the cell types 'a' and 'b' then the possible adhesions are a-a, a-b, b-b (discounting adhesions to any external medium) and the work of adhesion, which is defined as the work done in the formation of an adhesion over unit area[4], for each of these can be expressed as W_{aa}, W_{ab} and W_{bb} respectively. The four equilibrium states that can be achieved are shown in figure 1.2.

- (a) $W_{ab} = 0$. In this case there is no adhesion between the different cell types and thus each cell type will simply round independently of the other, forming two separate spheres.

- (b) $W_{ab} \geq (W_{aa} + W_{bb})/2$. In this case the heterotypic (a-b) adhesions are stronger than or as strong as the average of the homotypic (a-a or b-b) adhesions. Therefore, the most thermodynamically stable pattern will be achieved when the 'a' and 'b' cells are arranged alternately.

- (c) $W_{bb} \leq W_{ab} < (W_{aa} + W_{bb})/2$. This is the typical cell sorting case often referred to as the onion configuration. The average of homotypic bonds is greater

[4]Throughout this thesis work of adhesion is used to refer to both homotypic and heterotypic adhesions. Some early sources use the additional term 'work of cohesion' to refer to the former case.

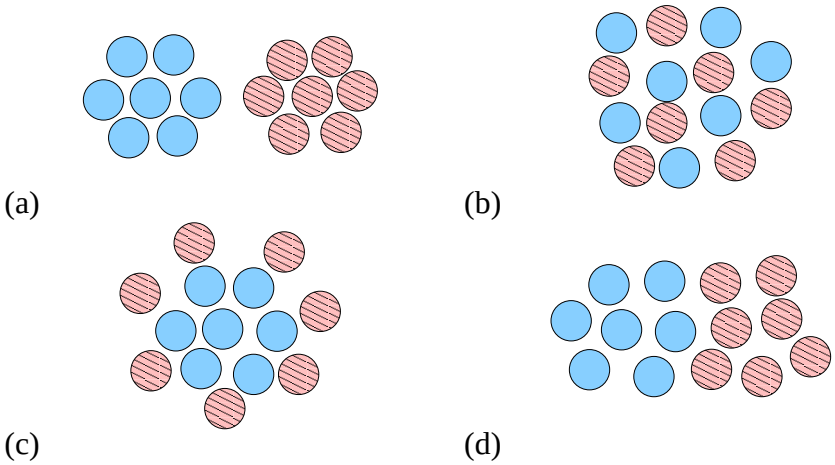


Figure 1.2: *Four equilibrium states formed when aggregates of differently adhering cells are combined. Blue = type a; pink = type b.*

than or equal to the heterotypic bond strength and thus the cells segregate with the more adhesive ‘a’ cells becoming centrally located.

As the centrally located ‘a’ cells round to maximise a-a adhesions and the ‘b’ cells form a shell around them in order to both maximise a-b adhesions and reduce surface free energy, the resulting structure resembles one large sphere with a smaller sphere located within.

- (d) $0 < W_{ab} < W_{bb} \leq W_{aa}$. This case is similar to case (a). As the heterotypic bond strength is the weakest, the cells will attempt to round separately. However, as the heterotypic bond strength is greater than zero there will be some attraction between both of the cell populations. The greater the heterotypic bond strength, the greater the surface area of the ‘a’ cell aggregate that will be engulfed by type ‘b’ cells.

Simple aggregate rounding is clearly easy to demonstrate and has been shown experimentally many times [12, 16, 54]. However, until more recently [12, 17] it was not possible to accurately determine the surface tension properties of real cell aggregates, and thus validate that these properties alone were sufficient to explain the sorting out of intermixed cells. This has now been shown to be the case [12]. In addition, more recently still [16], it has been demonstrated that the surface tensions which are observed

are specified almost entirely by the number of CAMs as predicted in Steinberg's original work [54]. Foty and Steinberg [16] state "We regard this as the final verification of the explanation of 'tissue affinities' proposed in the DAH.".

1.4 Related Work

1.4.1 Robotic Models of Self-Organisation

Current work on self-organising robotic systems can easily be divided into three main categories:

- Metamorphic robots
- Collective robotics
- Swarm robotics

Metamorphic robots are robots composed of many individual units which connect together to form rigid structures. This simplifies some aspects of self-organisation as the physical connections provide a reliable mechanism for inter-unit communication and also ensure the exact positioning of units with respect to each other. However, as the units must remain physically connected, repositioning units becomes more difficult. In most cases, the units are relatively simple with few actuated parts making them cheap to replace. Therefore, any movement often involves cooperation between neighbours [32, 34, 35, 42, 50, 62]. However, work by Kotay et al. [37] shows that units capable of moving without assistance from the other modules (except as a surface to connect to) can be constructed if unit complexity is not an issue.

Collective robotics is the term used to describe systems of independently mobile agents which interact to achieve some coordinated behaviour. Unlike metamorphic robots, in these systems the issue of mobility is not as central to the problem. Instead, the more interesting issue is that of communication and sensing. Some systems assume an infinite communication range and the ability to detect the exact distance and direction of every signal [58, 63]. However, as the number of agents and the distance they span increases, implementing this type of system becomes problematic. To overcome

this a number of systems assume only local sensing and communication [19, 39, 52]. However, coordinating behaviours across the population of agents then becomes more difficult.

The final category of swarm robotics is a broadly used term assigned to systems which mimic behaviours observed in populations of social insects such as ants and termites. Therefore, as in collective robotic systems, the agents are independently mobile. However, the term is increasingly being used to describe systems where, like the insects they model, the agents have the ability to connect and disconnect from each other as they interact. This has clear advantages as the unconnected population can explore areas efficiently and pass through small gaps. However, when faced with tasks and obstacles a single agent cannot overcome, they can physically connect and increase their size and strength. With these tasks in mind, in this area, work on self-organisation has focused on simple behaviours of chain formation and aggregation [10, 27].

The control strategies employed in each of these areas are normally very specific to the abilities of the agents involved. Therefore, as the work presented in this thesis falls into the collective robotics category, the following sections focus on some of the most significant work in this area.

1.4.1.1 Local Force Based Approaches

In all physical systems a remarkable degree of organisation takes place at an atomic level, with atoms and molecules arranging themselves in response to short range attractive and repulsive forces. The models presented in this section aim to use similar forces to create highly scalable distributed systems which require minimal communication to operate.

Balch and Hybinette [3, 4] proposed a model which is inspired by the molecular bonding involved in natural crystal formation. Each agent in the system behaves like a simple atom with n bonding sites arranged uniformly and extending some small distance r from its hull (see fig. 1.3 - (a) & (d)). To move into a formation, the agents first determine the nearest available attachment site using the position and orientation of their neighbours and their knowledge of the current bonding geometry. Once a site has been determined, the agents generate an attractive vector, pulling them toward the

Figure 1.3: *Attachment site geometries (a) Column (b) Line (c) Diamond (d) Square.*

chosen location. In this way the desired crystal formation emerges. In addition, to allow the orientation of the formation to be given with respect to the agents' direction of movement, an additional parameter θ is used to give an angular offset of the bonding geometry from the front of the agent. Therefore, formations including line and diamond can also be generated (see fig. 1.3 - (a) & (d)). A similar approach is proposed by Fredslund and Matarić [18] in which each agent has essentially a single bond with a unique offset. Therefore, instead of lattice formations, the agents form connected chains with specific angles between the links.

Balch and Hybinette demonstrate that this control method can operate successfully on a small population of agents. When the agents encounter obstacles, the formation deforms appropriately, reassembling once the obstacle has passed. However, there is no enforced global configuration. Therefore, the agents can reorganise into any formation which satisfies the bonding geometry. For example, agents with the 'square' bonding geometry can equally form a line.

Fujibayashi et al. [19] propose an alternative approach in which agents within some range of each other behave as though they are connected by virtual springs. In this way, the attractive and repulsive forces between the agents force them into a hexagonal lattice. However, unlike the previous approach, Fujibayashi et al. also propose a method of controlling the global outline of the resulting lattice. This is achieved by probabilistically breaking some of the springs[5], which in turn causes some subset of the agents to reorganise. Candidate springs are selected depending on the number of connections which have been made by the agents that they join. Combinations which exist in the desired formation have a breaking probability of zero, while the probabilities for those

[5]The connections are actually broken by replacing them with larger springs which force the agents apart until they are out of range, at which point the spring vanishes.

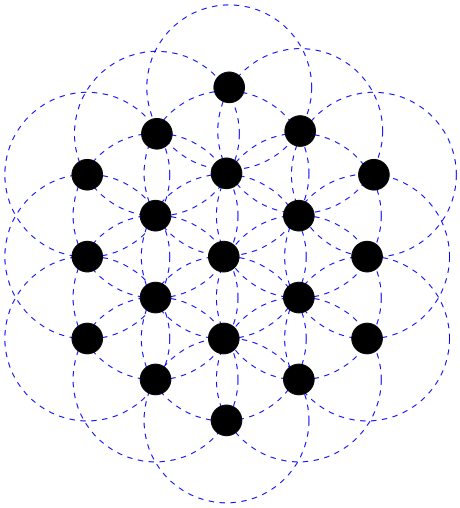


Figure 1.4: *A hexagonal lattice formed by the intersection of circular potential wells (indicated by dashed lines).*

which don't tend to be high. Therefore, some agents will continue to move randomly around the structure until the final stable configuration is achieved. The main disadvantage with this approach is the difficulty of choosing the most appropriate parameters for each configuration, such that the agents converge as quickly as possible while avoiding any deadlock states.

A similar, but much more adaptable approach is proposed by Spears and Gordon [52]. Spears and Gordon introduce a framework called 'Artificial Physics' (AP), more recently named 'Physicomimetics' [53], in which agents react to virtual forces which are motivated by natural physical laws. In the simplest example given, agents closer than some specified distance R repel each other, while agents further apart than R but within some visual range, are attracted closer together. Therefore, each agent has a circular potential well of radius R around itself, where neighbouring agents will be stable. When a large number of agents interact, the potential wells intersect with each other, creating nodes of low potential which form the basis of a hexagonal lattice (see fig. 1.4). Spears and Gordon demonstrate that this simple approach, which requires agents with only basic sensing abilities to detect the distance and direction of any neighbours, can achieve reasonably good lattices starting from an initial randomly organised cluster. However, the outline of the lattices remains irregular and occasionally multiple

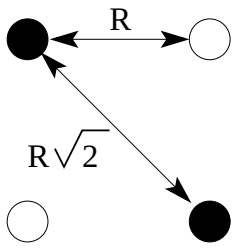


Figure 1.5: *The arrangement of agents in a square lattice. Agent colour denotes spin.*

agents get trapped together in tight clusters where only a single agent is expected.

Adding an attribute called 'spin' to the above system also allows the formation of square lattices. In this case, half the agents are initialised as spin 'up' and the other half spin 'down'. The detected distance r between two agents with like spin can then be modified to $r/\sqrt{2}$. This effectively creates a different potential well for agents of each spin causing the formation of a square lattice where the spins are intermixed (see fig. 1.5). It has been shown that this approach does produce regions where the lattice is as expected, but often results in global flaws and the same local clustering as the previous example. However, many of these can be eliminated by adding noise to the system in the form of 'spin-flipping'. With this mechanism, any agents which are within some threshold distance of their nearest neighbours (e.g. in a local cluster) probabilistically change spin and are immediately repelled. In addition to fixing initial flaws, the increased flow of agents provides a crude repair mechanism which will continue to fix flaws as they arise.

A final adaption of this approach allows the formation of perfectly symmetrical lattices by introducing a notion of sorting. Each agent is assigned an attribute (m,n) where m defines the agent's relative order to other agents in the same row and n defines the same for columns. As the ordering must have some direction, each agent requires global information indicating its current orientation. The sorting takes place by introducing a new attractive force between agents which are not ordered correctly with respect to either the m or n value. Therefore, the agents will be drawn past each other and toward their correct position. It has been shown that this technique can produce both perfect squares and hexagons. Also, if agents are initialised with both sets of

attributes, it is possible for them to switch from one formation to the other in response to some observable signal. However, the additional performance comes at the expense of the agents' simplicity.

1.4.1.2 Beacon/Gradient Models

Another approach to self-organisation, which has also been inspired by nature, is the use of virtual gradients. In biological systems cell fates are often determined by the concentration of specific soluble molecules called 'morphogens'. Morphogens diffuse from the source region, producing concentration gradients through the organism [20, Chap. 3.]. The cells then react to the different concentrations and change their behaviour accordingly. Some gradients even trigger secondary gradients, allowing more complex patterns to develop. This mechanism is particularly important in symmetry breaking and establishing some degree of polarity in developing organisms.

In the simplest case, this type of system can be modelled by allowing agents to broadcast signals over greater distances such that the strength of the signal deteriorates the further it travels. In this way, the signal behaves like a diffusing protein with some limited range and indicates the distance and direction of the source. This is the approach taken by Ünsal and Bay [63] who propose a system where certain agents take on the role of beacons around which others arrange themselves. In the simplest example, a single agent acts as a beacon and the others organise into a ring formation by moving toward a region around the beacon where the signal strength falls within some range. Once within this region, the agents are repelled by their neighbours, forcing them to distribute themselves evenly. In this way, the system is behaving similarly to the AP model [52] (see section 1.4.1.1) but utilises signals and sensors with much greater range. Ünsal and Bay [63] also demonstrate that the model adapts easily to 3D and show how a simple paraboloid structure can be formed with the addition of only one more beacon. However, the system assumes some method of initially getting the beacons themselves accurately positioned. In a related work, Kondacs [36] propose a method of accurately positioning reference (beacon) agents by using gradients from specific neighbours to effectively triangulate their position. However, the described system is aimed at networks of inaccurately positioned sensors and not mobile agents.

The main drawback with this beacon signal approach is that it may not be suitable for future technology where the size of agents is greatly reduced and issues of scalability, battery consumption, and line of sight make long range communication infeasible. To overcome these problems, Nagpal et al. [43] proposed a system where the gradient signal is simply propagated across a population of agents using local hops. The agents themselves have minimal sensor capabilities and only detect the signals and not the distance and direction of their source as in other models. To generate a gradient, a seed agent broadcasts a signal containing the value zero. Each agent that receives the signal then increments the value by one and forwards it on. However, to prevent the signal from travelling backward, recipient agents only record the lowest value they receive from all their neighbours. Therefore, a wave will spread out from the source across the entire network. In addition, if the communication range of the signal is known to be some value r and an agent i indicates the minimum signal value it has received is h_i then it is known that the agent is no more than rh_i from the signal source. In a more recent work, Mamei et al. [39] demonstrate how this elegant gradient propagation approach can be successfully applied to mobile agents exhibiting similar sensor capabilities allowing simple configurations such as filled circles, rings and lobed circles (rough approximations of polygons) to form. The formation of each configuration begins with the detection of the 'barycentre' of an initially random cloud of agents. This is achieved by each agent propagating a unique gradient such that the agent which has the highest total of the signal values for all the gradients is the one at the centre of the cloud. Once the barycentre has been found, the simple filled circle formation can be achieved by propagating a single gradient from the central point. As in the beacon model proposed by Ünsal and Bay [63] the agents simply move toward the source of the signal until they are within some predefined range. However, the agents do not have the ability to detect the direction of the gradient and can only sense whether the gradient increases or decreases as a result of their movement. Therefore, the agents randomly choose a direction to move in and if it turns out that the gradient actually increases, they invert the direction. The critical drawback of this technique is that agents at the periphery of the cluster may wander away from their neighbours and lose all communication with the rest of the cluster. Therefore, the agents would

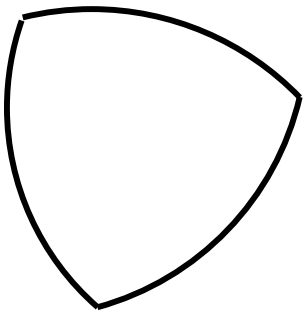


Figure 1.6: *Reulaux's triangle.*

essentially be lost.

1.4.1.3 Rule Based

Not all the control systems for collective robotics take their inspiration from nature. Instead, some novel techniques have been designed where the agents follow simple hand crafted algorithms to achieve a desired collective task. However, designing such algorithms so that they are robust and scalable is complex as the interactions between individual units which will produce a desired global behaviour are difficult to determine.

In an early work Sugihara and Suzuki [58] proposed a selection of these algorithms for producing circle (ring), filled circle, line, polygon and filled polygon configurations. In each case it is assumed that the agents can sense the distance and direction of all the other agents, but have no ability to communicate directly.

To generate the circle configuration, Sugihara and Suzuki [58] propose the following three rules for the agents to follow.

$$\begin{cases} d > D, & \text{move toward R'.} \\ d < D - \delta, & \text{move away from R'.} \\ D - \delta \leq d \leq D, & \text{move away from R".} \end{cases} \tag{1.1}$$

where R' is the furthest agent and R'' is the nearest agent, d is the distance to R', D is the diameter of the circle and δ is some small constant.

These rules are similar to those proposed by Ünsal and Bay [63]. However, by

using the distance from the furthest agent, this method eliminates the need for a central beacon. The drawback is that the final configuration can actually take the form of any oval of constant width and a circle is only one possibility, Reulaux's triangle is another (see fig. 1.6).

The algorithm for forming more complex polygon configurations also begins with formation of a circle. However, once it has formed, the agents split into two subsets. The first subset of agents are those which form the vertices of the polygon and must be moved to their desired locations manually. Although, it could also be possible that they utilise some additional navigation ability the others do not posses. The remaining agents simply execute the 'CONTRACTION' algorithm which moves them toward the midpoint of the agent on their immediate left and right. In this way, the sections between the vertices will flatten out forming linear sections.

The key drawback with this approach is that the decentralised behaviour is actually only carrying out the simplest aspect of the task and the more complex operation of positioning of the vertices is simply assumed.

1.4.2 Multicellular Models & Models of the DAH

Models of multicellular systems vary greatly in their representations of cell aggregates and the mechanisms through which these aggregates reconfigure. This section presents those categories of model which investigate differential adhesion and discusses the positive and negative aspects of each.

1.4.2.1 Cellular Automata Models

In cellular automata (CA) models the world is divided into a regular lattice of discrete sites. The particles or agents which inhabit the CA occupy some or all of these sites and their state and position evolves over time in relation to a set of specific rules. In general, these rules depend on both the current state of the agent and that of its neighbours.

In early CA models of cell sorting (Cell Lattice Models), each cell is represented by a single agent occupying a specific site on a regularly tessellated plane. Different cell types are represented by agents with static predefined states, and mobility rules relating

to the adhesions between these cell types control the cells' movements. Typically, for convenience, the medium surrounding the aggregates is simply represented as another distinct cell type with its own adhesive properties. In the intuitive model proposed by Antonelli et al. [2], the energy of each hexagonal cell is calculated from the bonds it makes with its six neighbours. These energy levels are compared and cells on neighbouring tiles swap positions if the swap results in a greater increase in energy for one cell than the corresponding decrease for the other. In practise, '0-moves' (moves where there is no change in the energy of either cell) were also allowed, with some probability, to provide some level of random exploration. This simple model resulted in the rapid segregation of the two cell types. However, there was no trend toward the formation of a central cluster as predicted by the DAH. Improved performance was achieved by Goel and Rogers who modified this approach to allow an exchange of position between any two cells on an interfacial border, not just between nearest neighbours as before [23, 49]. The justification for this is that the effect of these movements could instead result from a sequence of rearrangements in each aggregate, with cells only actually moving to neighbouring spaces. Simulations revealed that this modification allows some degree of engulfment [23], and sorting [49] to take place. However, the best performance was only achieved when cell energies included some influence from other cells up to four cell diameters away[6], not just the local interactions predicted by the DAH.

The main criticisms of this approach are that the rigid cells fail to account for any membrane or cytoplasm dynamics and the exchange of cell positions assumes the cells have knowledge of how energetically favourable any position will be prior to moving there. An alternative CA approach, first proposed by Glazier and Graner [21], overcomes these limitations by using a Q Potts model to represent each cell as a subset of agents each exhibiting a state, also known as spin, uniquely associated with the particular cell[7] (see fig. 1.7). Therefore, the total surface energy of any cell will be a product of the adhesions resulting from all the interfacial boundaries of the agents

[6] On a lattice of squares, 8 cells are within 1 cell diameter of any given cell, 16 within two diameters, 24 within 3 etc.

[7] In a typical Potts model each site has one of two spins. A Q Potts model allows each site to exhibit one of Q different spins.

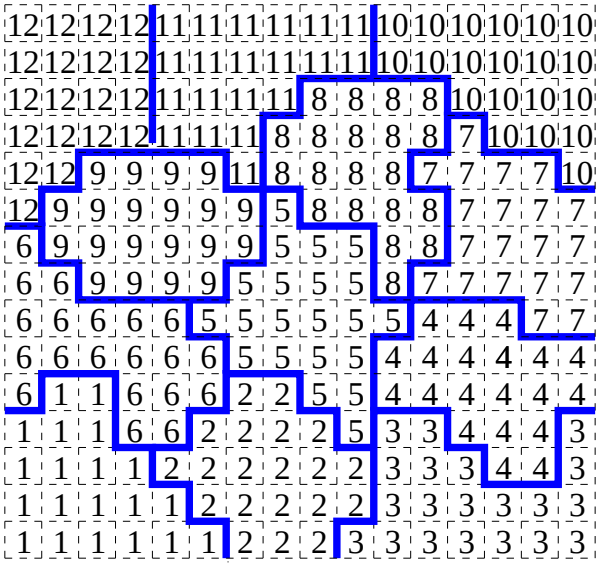


Figure 1.7: *A section of the CA for a Q Potts model. Twelve different cells are shown and notional boundaries are marked by thick blue lines.*

at its perimeter. To provide some method of constraining the cells' volumes (or areas in two dimensions), each cell's total energy (E) also includes some additional term which relates its actual volume (v) with that of some predefined target volume (V). Glazier and Graner [21] propose that this term ($\lambda(V_c - v_c)$) provides a simple method of encoding all the membrane and cytoplasm dynamics of the cell without explicitly modelling the dynamics themselves. Thus:

$$E_c = \lambda(V_c - v_c) + \sum_{e_{ij} \in \mathcal{B}} J_{ij} \tag{1.2}$$

where $\mathcal{B}$ is the set of all boundary edges (e) associated with a particular cell c, and J_{ij} the surface free energy of a contact between two membranes of types i and j. Once more, the medium surrounding the cells is also modelled as an additional cell type. However, this additional cell type has no volume constraint and can, therefore, extend indefinitely.

The update rule of the CA is dependent on the cellular energy defined above. An agent in the CA is selected at random and its spin is converted to that of one of its

eight neighbours, also selected at random, with some probability relating to the change in energy (ΔE). In addition, a thermal noise factor (T) is included to account for any random fluctuations in the cells' membranes. Thus, the probability P of flipping any site's spin from x to y is given by:

$$P(x \to y) = \begin{cases} exp(-\Delta E/T), & \Delta E > 0 \\ 1, & \Delta E \leq 0 \end{cases} \tag{1.3}$$

In this way, barring the random effects caused by noise, the cells tend to move in such a way that the previous energy term is minimised (i.e. adhesions are maximised and the cells' volumes are at the target level).

It has been shown that these more flexible models are capable of recreating the observed biological cell sorting dynamics in addition to some more subtle cell behaviours, in both two and three dimensions [21, 22, 41]. However, cells in the model have the ability to fractionate and recombine over the course of a simulation. Therefore, although they provide a useful approximation, the dynamics of the individual cells don't truly match those observed in the real biology. Proponents of the non-CA approaches also argue that cell boundaries in these models will always be limited by the structure of the lattice and thus smooth angled boundaries, which their models allow, are not possible.

1.4.2.2 Cell Centric Models

Cell centric models make no attempt to explicitly define the geometric structure of the cell membranes. Instead, the location of each cell is marked by only a single centre point and the cell membranes are determined by calculating the resulting Voronoi tessellation (also referred to as the creation of Dirichlet domains in the literature). In the 2D case, this subdivides the plane into convex polygons where the interior of each polygon represents the set of all points closer to a particular centre point than any other (see fig. 1.8). Honda [29] demonstrated that this type of tessellation very closely approximates the actual cell boundaries observed in a number of 2D aggregates of cells. In addition, he showed that by replacing a specific point by two new ones and slowly moving them apart it was possible to create behaviour resembling cell division. This

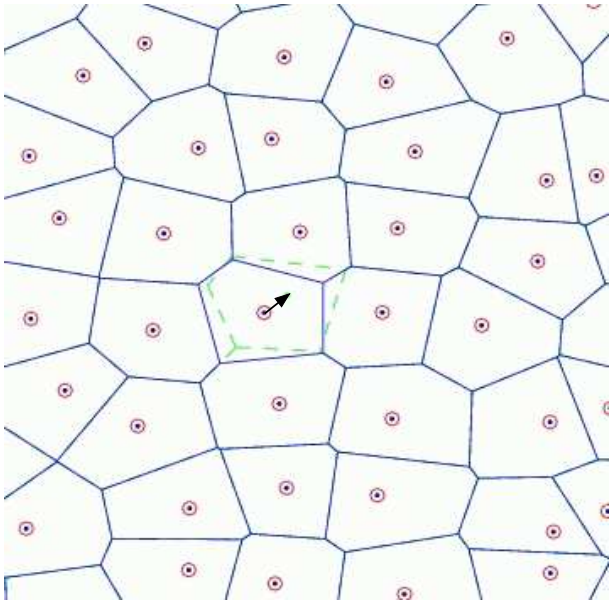


Figure 1.8: *Example of a Voronoi tessellation.*

works because each centre point affects the geometry of any adjacent cells. Therefore, as a centre point is moved, the length and number of cell-cell boundaries will change appropriately (see fig. 1.8). These smooth changes in the cell-cell boundaries provide the basis for the cell centric models presented by Sulsky et al. [59] and Graner and Sawada [25].

The movements of the cells in these models depend on a calculated energy term. This term combines the total adhesive energy produced by the various cell-cell boundaries and an additional component representing membrane elasticity or compressibility which effectively constrains the area of the cells. At each time step, complex calculations are carried out to determine the force which should be applied to each centre point to move the overall pattern along the steepest path of the energy landscape. These forces can then be combined with an appropriate viscous drag force and any additional forces which have been incorporated into the model (e.g. random fluctuations).

In the model presented by Sulsky et al. [59] it was again necessary to represent the medium by cells of a specified type thus allowing the cells at the boundary of the aggregate to take on a realistic shape. However, Graner and Sawada [25] enhanced their model by placing a maximum radius around each centre point which the cell

boundary cannot extend beyond. Therefore, there is no need to explicitly represent the medium as membranes of cells at the boundary of the aggregate round up appropriately. It can also be seen that using this limit, a single cell in isolation will realistically display a circular boundary.

Cell centric models provide a simple representation of cell aggregates in which the cell geometries can change smoothly over time and it has been shown that they are capable of recreating many of the observed cell sorting behaviours [25, 59]. However, the main limitation of these models is that the cells are restricted to polygonal shapes which can be appropriately represented by a Voronoi tessellation. Thus, cells with anisotropic structures cannot be modelled.

1.4.2.3 Boundary Shortening Models

The boundary shortening model was initially developed to provide a method of estimating the level of tension which was thought to exist in microfilaments running around the periphery of cells forming a particular tissue [30]. However, it is included here for completeness as the later adaption, the weighted boundary shortening model [31], provided a novel method of simulating differential adhesion between heterotypic cells.

In complete contrast to the cell centric models, the boundary shortening models explicitly define the membrane of each cell as a polygon with known vertices. Starting from an initial configuration which is constructed by tracing an actual cell aggregate, the algorithm proceeds by iteratively adjusting the membrane edges, shortening them where possible, while conserving each cell's area. This is achieved by exploiting some simple geometry. At each time step a single edge (PQ) is selected at random and the vertices P and Q are moved to reduce the total edge length, $AP+BP+PQ+QC+QD$ (see fig. 1.9). However, P is only moved along the line PP' parallel to AB, thus conserving the area of the cell which includes vertices A, B, and P. Similarly Q is only moved along the line QQ' which is parallel to CD. The final vertices P' and Q' are selected by sequentially moving P' further from P by small distances (positive or negative), and in each case moving Q a suitable amount in the opposite direction to conserve the area of the remaining two cells, until $AP'+BP'+P'Q'+Q'C+Q'D$ is minimised. Once

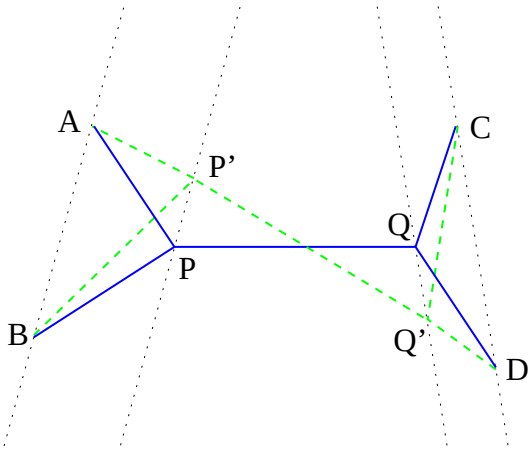


Figure 1.9: *The iterative step in the boundary shortening model. Figure adapted from [30].*

the cells reach an equilibrium state, the algorithm is terminated and the total boundary length is compared with the boundary length of the initial configuration to determine the degree of boundary shortening that has taken place. The assumption is that if this value is small it suggests that a contractile system is in operation and that the cell sheet is likely to be in tension.

The modified version of this model allows differential adhesions to be simulated by applying different weightings to edges between different cell types. Therefore, at each step in the simulation, the term being minimised becomes $k_{AP'}AP' + k_{BP'}BP' + k_{P'Q'}P'Q' + k_{Q'C}Q'C + k_{Q'D}Q'D$ where the k values are determined by the cell types meeting at the particular edge. If the k value is small, the edges will tend to elongate thus approximating the effects of greater adhesion. In addition to this, to realise smoother movements, a maximum value was placed on the distance that any vertex could be moved in a given time step.

Clearly, in some instances this boundary shortening procedure will create edges which have very short or almost null length. In such cases the 'neighbour change' mechanism introduced in [30] is applied. This technique removes these short edges and replaces them with a new interface between cells that were previously isolated from each other (see fig. 1.10). Therefore, it is possible for much larger rearrangements to take place within the aggregate.

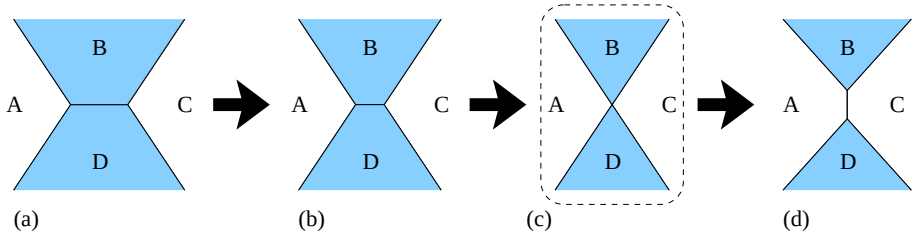


Figure 1.10: *This figure demonstrates the 'neighbour change' mechanism. (a) and (b) show the hypothetical situation where the edge separating cells B and D is reduced in length. Once the edge length becomes smaller than some critical value, the edge is removed and replaced by a new edge separating cells A and C; (d). In this way, it is as though the rearrangement has passed through the intermediate stage (c).*

Honda et al. [31] have demonstrated that, when starting from specific initial configurations, this simple approach can successfully generate the 'checkerboard' pattern which results from strong heterotypic adhesions. However, the model has a number of important limitations:

- Unlike the lattice and cell-centric models, in the boundary shortening model, the cell membranes are strictly limited to their initial area. Thus, there is no concept of elasticity or compressibility.
- Rearrangements can only take place via very specific membrane changes.
- The edge weightings cannot easily be translated to corresponding work of adhesion or surface tension measures.

1.4.2.4 Finite Element Models

Finite element models are a relatively recent development and represent a new approach to modelling cells. Other models account for all the internal mechanics of the cell using some general parameter which restricts the cell's volume or area. However, finite element models attempt to break each cell into its individual elements which can then be described by relatively simple mathematical expressions (see fig. 1.11).

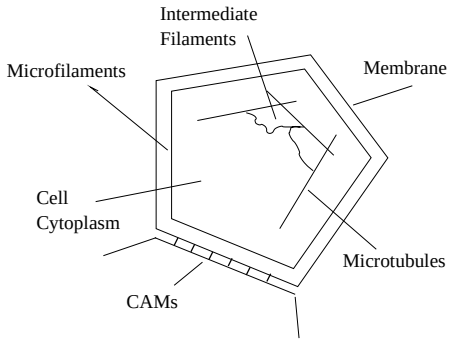


Figure 1.11: *The individual cell elements incorporated into the finite element model proposed by Chen and Brodland [8]. Figure adapted from [7].*

In the finite element model proposed by Chen and Brodland [6, 7, 8] cells are represented by n sided polygons which are broken into n triangular sections (see fig. 1.12). Each of these sections contains an appropriate volume of cytoplasm with some predefined viscosity value μ, such that the cytoplasm will resist any deformation caused by the other forces in the system. However, the cell description does allow cytoplasm to flow between neighbouring sections to prevent the mesh from simply becoming locked in a particular configuration [6]. The model proceeds by using the mathematical expressions identified for each of the cell's mechanical elements to determine the forces acting on the nodes associated with each triangular section. The resulting displacements and velocities can then be calculated for each node in the model. Additionally, as in the boundary shortening procedure, a 'neighbour change' mechanism is applied to edges which fall below some minimum length, thus allowing the quadruple junction stage to be skipped (see fig. 1.10 - c).

The main advantage of these models is that they can accommodate any number of mechanical elements and material properties, each of which can be assigned realistic values. Therefore, it is possible to directly compare the simulation time scale with that of the real system. Additionally, it is possible to modify these properties to allow specific cell shapes to be modelled unlike other models which assume an isotropic structure. However, including all these details makes these models complex and they require a great deal more computing power than the other approaches presented in this chapter. It is also assumed that it is possible to accurately obtain appropriate values for

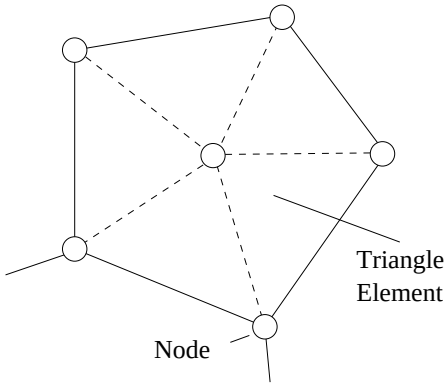


Figure 1.12: *The finite element model representation of a cell.*

each of the cellular elements which are identified. However, this may not always be the case.

In the simulations presented in [7] and [6], it can be seen that this model does produce intermixing and sorting behaviours when the model's cells are assigned surface tension values obtained from real aggregates of cells. However, the sorting which is observed is only partial and no central clustering is observed. Brodland and Chen [7] suggest that this points to failings in the DAH. However, their model does not appear to include any form of random membrane fluctuations. Therefore, it is possible that this could also be an explanation for the same behaviour.

1.4.2.5 Force Driven Model

In another recent work, Palsson [45] proposes that the main flaw with many of the existing models is that they are limited to two dimensions and thus cannot accurately describe the movements of cells in real multicellular environments. The exceptions to this are the Q Potts models. However, Palsson [45] claims that these models do not contain enough information about the forces between the cells or the individual cell characteristics.

The model presented by Palsson [45] attempts to overcome these limitations by representing each cell as a deformable 3D object with characteristics of cell stiffness, cell adhesion, and active locomotion. To allow simulations of large cell populations, the actual cell structure is simplified to that of an ellipsoid with all deformations rep-

resented by a change in the length of the semi-axes (e.g. the three axes which describe the ellipsoid). The actual changes are controlled by a simple equation which mimics the viscoelastic properties of single cells while conserving the total cell volume. The cell movement is determined by the interaction of the passive, active, and drag forces acting on each of the cells in the system.

The passive forces are those generated by the adhesive and elastic interactions between the neighbouring cells. To mimic real adhesion, as the distance between the cells (d) is reduced the adhesion increases (see fig. 1.13). However, when the cells get too close, their resistance to deformation creates a repulsive force acting in the opposite direction. Thus, the overall attractive force between two adhering cells will be zero at some limiting distance, increase to some maximum and then reduce until it becomes negative.

The drag forces are generated by the viscosity of the extra-cellular matrix and viscosity due to cell-surface interactions. Therefore the total drag force is determined from both the speed of the cell relative to its neighbours and its speed relative to the global frame.

Finally, the active forces in the system are generated by the cells themselves. These forces are included to allow the model to simulate chemotaxis in addition to general cell sorting. The forces are based on the pseudopod formation exhibited by amoeboid cells[8]. This involves extending part of the cell body which attaches to the surroundings, allowing the cell to pull itself forward. In the model, the cells achieve this by orienting the main axis of the cell body toward the direction of a signal if the signal's gradient is above some minimum threshold. Once the cell has reoriented it generates an active force in the chosen direction.

Palsson [45] gives the details of a number of sorting experiments in which many characteristics of the model, including cell stiffness and random motion, are examined. The results show that when two tissues are initially intermixed, the DAH successfully predicts the resulting engulfment behaviour. In addition, the random motion results appear to agree with other models and indicate that increasing random motion does have a beneficial effect on the sorting behaviour.

[8] In vertebrate cells, this type of movement is generated by the formation of similar structures called lamellipodium [38, Chap. 19.].

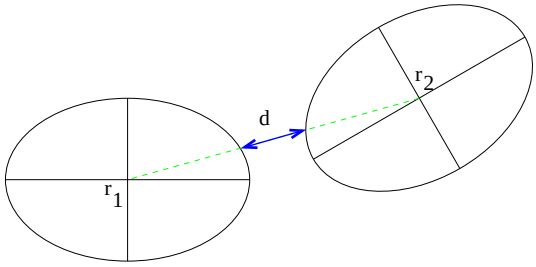


Figure 1.13: *Two adhering cells. The distance d is used to determine the strength of the force between the cells and any resulting force is applied along the line joining r1 and r2.*

1.4.3 Summary

The work presented in this thesis spans two distinct research areas; collective robotics and multicellular models. This section has presented a review of the related literature in each of these fields.

The review of the collective robotics literature identifies a number of distributed control systems which allow populations of agents to organise into a variety of different configurations. However, the vast majority of these systems fail to consider how the dynamics of the model will be affected by the noise (in both sensing and movement) that would be inherent in any physical implementation. In addition, those systems which are both scalable and robust are limited to a few simple configurations and it is unclear how additional behaviours could be incorporated to increase their abilities.

The review of the literature relating to multicellular models, highlights the variety of different cell representations which have been explored. However, many of the models which are presented require realistic but complex interactions between neighbouring cells and abilities to rearrange which are only plausible in simulation. The most significant approach, with respect to this work, is the force driven model proposed by Palsson [45]. Palsson, greatly simplifies the structure of the individual cells and models the interactions between them as a combination of forces. In this way Palsson's approach is similar to that taken in this work. However, his virtual system still incorporates dynamics which which would be very difficult to model with a physical robotic system.

Chapter 2

Problem and Model

This chapter discusses the differential cellular adhesion problem in more detail and presents a simulated robotic model which, by adhering to all the identified constraints, attempts to reproduce the observed behaviour.

2.1 Problem Constraints

There are three sets of constraints that must be taken into consideration when constructing a physically realisable model of some biological phenomenon.

1. The intrinsic constraints of the phenomenon, without which the phenomenon will not take place.

2. The physical and technical constraints of a physical embodiment (e.g. communication mechanism, size, etc.).

3. The biological constraints. The actual methods employed in the biological phenomenon (e.g. cell membrane fluctuations).

Obviously, if the model is to perform at all it must conform to the first set of constraints. Additionally, the second set of constraints must be adhered to, otherwise it will not be possible to construct a physical realisation. Finally, if the model is to accurately resemble the observed biology the third set of constraints must be realised.

However, it is likely that the underlying biology will have to be greatly abstracted to meet the limitations of the technical and physical constraints.

2.1.1 Intrinsic Constraints

In this work, as outlined earlier, we are interested in developing a physical realisation of the cell sorting behaviour observed in aggregates of living cells (see section 1.3). Steinberg [54] states that the only requirements necessary for sorting to occur are that the units are both mobile and mutually adhesive. Any differences in adhesion therefore drive the system to some equilibrium state. However, this only explains cell sorting at a macroscopic level and does not presuppose the nature of the microscopic dynamics that drive the cells. Mombach and Glazier [40] demonstrate that fluctuations in the cell membrane appear to play a key role in the process and that blocking such fluctuations prevents full sorting from taking place. This is also corroborated by *Q Potts model* simulations [41] which demonstrate that only partial sorting occurs when random boundary fluctuations are suspended. Furthermore, observing the movement of individual cells in real aggregates, it can be shown that when these fluctuations are present, the cells appear to perform an approximate random walk[1] [40, 48]. It is proposed that the reason this random motion is so vital is that it allows cells to interact with their neighbours and explore their local area [40, 41]. As a consequence, the aggregates are less likely to become trapped in metastable states and can reach the true energy minimum.

It is therefore possible to summarise the intrinsic constraints as follows:

1. The units must be mobile.

2. The units must be mutually adhesive. For sorting to take place, there must exist some difference in this adhesion.

3. For global sorting to occur, the units must perform some local exploration.

[1] The random walk is biased by the surface energy minimisation.

2.1.2 Physical and Technical Constraints

In this work the 'HYDRON' module, developed by the HYDRA European Consortium, was used as a rough guide to achievable structure and performance. However, the model itself remains flexible enough to adapt to a multitude of possible hardware solutions. The HYDRON modules were chosen as they replicate the most basic capabilities of real biological cells, providing active movement in three dimensions and local inter-unit communication.

The physical module consists of a near spherical semi-transparent hull, in which a limited number of optical transmitter-receiver pairs can be fixed. The original design allowed for eight such sensors, located on the corners of a cube inscribed inside the hull, to provide full sensor coverage. As well as enabling data transfer, this system provides approximate distance information through the intensity of the light signals being received.

The locomotion apparatus, which constitutes the majority of the hull's internal space, consists of a syringe actuator and four jet-actuators. The syringe actuator controls the module's buoyancy (vertical motion), and the jet-actuators control the horizontal motion. Such a system should therefore be capable of producing a force within some range in any direction.

Finally, to allow the modules to operate autonomously, they carry their own power supply and processor board.

2.1.3 Biological Constraints

Mobility and mutual adhesion were highlighted as intrinsic constraints of the DAH. Therefore, the biological constraints will include the methods employed by the cells to achieve these properties. In addition, the actual structure of the cells must be considered as this governs how cells interact with their neighbours.

Cell Structure:

Section 1.2 discusses the basic structure of both prokaryotic and eukaryotic cells. However, in this work, the eukaryotic cells which make up multicellular organisms are of the greatest interest. At a high level of abstraction, these cells consist of three

distinct sections; nucleus, cytoplasm and cell membrane. The membrane behaves like a deformable physical barrier which surrounds the viscous cytoplasm. As there is no free flow of material through the membrane [60, Chap 4.], it must be strong enough to withstand the fluid pressures exhibited by the contained cytoplasm. Clearly, all the interaction between neighbouring cells must occur at this cell membrane, allowing the central nucleus to remain protected within.

Adhesion Mechanisms

Steinberg [54] postulated that the simplest explanation for type specific differences in the adhesion between cells was that a number of adhesion sites were abundantly scattered across the surface of a cell's membrane. Thus, the work of adhesion would be proportional to the number of adhesion sites that are able to bond per unit of surface area. It is now known that such sites do exist in the form of CAMs (CAM = Cell Adhesion Molecule)[20, Chap. 3]. Additionally, in their most recent work, Foty and Steinberg [16] have validated this original hypothesis by demonstrating that alone, the numbers of surface cadherins (the major group of CAMs) almost entirely specify the surface tensions of cell aggregates.

There are many different forms of CAM. However, there are only three ways they can interact, as summarised in [1, Chap. 19. Section: Cell-Cell Adhesion]:

1. Homophilic binding where molecules on a cell bond to the same kind of molecules on another cell.

2. Heterophilic binding where molecules on a cell bond with molecules of a different kind on another cell.

3. Secreted multivalent linker binding where cells bind through an extracellular linker molecule.

Cell Motility:

There are two types of cell motion; long range active motion of the sort involved in chemotaxis, where the cell extends and contracts in a very directed manner [38, Chap. 19] and passive motion, where random fluctuations of the cell membrane cause it to randomly explore [40, 48]. However, although it has been demonstrated that passive

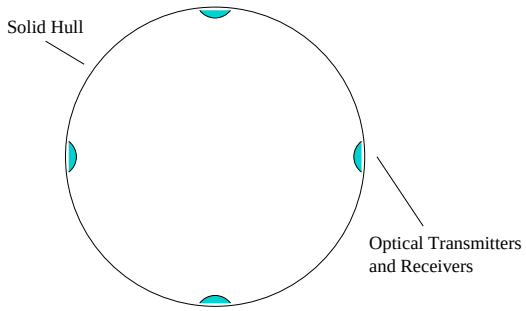


Figure 2.1: *An A-Cell cross-section.*

motion may play a key role (see section 2.1.1), the DAH explanation of cell sorting requires no directed cell migration.

2.2 Cellular Adhesion Model

This section presents the proposed cellular adhesion model. The model is implemented on abstracted HYDRON modules (see section 2.1.2), which were simulated using a realistic physics engine[2] to model their interactions with each other and the environment. An additional software component has also been added to provide a mechanism for modelling the optical communication at a rather simpler level. By demonstrating the model on this platform it can be shown that, at least to some extent, the physical and technical constraints (see section 2.1) of such a system have been addressed.

The model has been separated into individual components which relate back to the biological constraints raised in section 2.1.3. In the following sections each of these components is discussed and the details of the A-Cell implementation are presented.

2.2.1 A-Cell Structure

As mentioned previously (see 2.1.2), the HYDRON module has been used as a guide to a possible physical realisation of this model. Therefore, the following structure bears many similarities to its design.

[2]The commercial 'Vortex' physics engine V2.0 produced by CM Labs was used for this purpose.

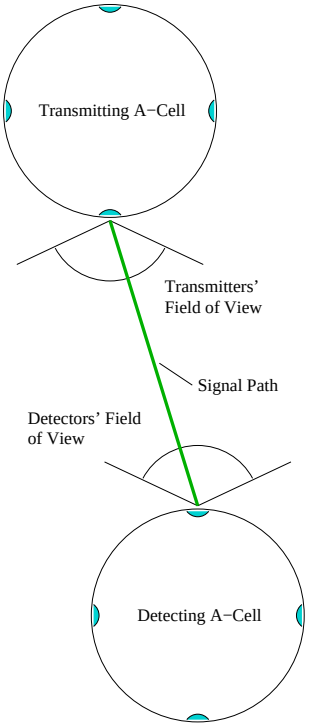


Figure 2.2: *Line of transmission between two A-Cells*

Each A-Cell consists of a spherically shaped hull which encloses all the internal components. It is assumed that the hull is weighted to provide a self-righting torque keeping the A-Cell upright. Therefore, only free rotation around the vertical axis is possible. A number of communication sites consisting of a local optical transmitter-receiver pair are located on the hull's surface (see fig. 2.1). However, although the implementation presented here keeps such sites to a minimum, the model actually allows complete flexibility in both their number and position. The optical signals are modelled with a limited predefined range and the sensors themselves with limited field of view. Therefore, a signal will only be detected if it is possible to draw an unocluded line between the transmitter and detector, such that the line falls into their respective fields of view and its magnitude does not exceed the signal range (see fig. 2.2). The full complexities of signal reflection and interference have not been included in the model as these would be specific to the sensor specification. However, it should be possible to account for this type of disruption by careful selection of the communication protocol.

Finally, the A-Cells do not presume any specific means of motility. Instead, for generality, they only require that it is possible to generate a piecewise constant force below some predefined maximum in a given direction. Such a force vector could be further processed on real hardware to produce the mechanism-specific sequence of actions that are required.

2.2.2 A-Cell Communication

The communication system used by the A-Cells is relatively simple. Each A-Cell continually broadcasts a 'transmission signal', containing some uniquely identifying attribute, which may be detected by any of the neighbouring units.

The transmission signal consists of two parts. The first part of the signal contains information relating directly to the A-Cell from which it was broadcast. The second part of the signal is a list containing information aimed at individual A-Cells in the system.

In addition to this data transfer, the transmission signal also provides essential positional information. As discussed in section 2.1.2 it is possible to crudely estimate the distance a signal has travelled by measuring its relative detected strength [3]. Therefore, by comparing the distance values calculated at different detectors, it is possible to determine an approximate direction vector and distance for the signal source. There are a multitude of techniques, varying in complexity, which can be used for this type of source estimation. However, in the implementation present in this work, the following simple but fast approximation is used. Given a subset of communication sites which detect the signal (C), a direction vector of the source (V^s) is estimated by averaging the direction vectors of the sites (V^d)[4]. Likewise, the distance estimate (D^s) is calculated from the average of the signal strengths (S) detected at each site.

$$V^s = \sum_{c \in C} V_c^d \qquad (2.1) \qquad\qquad D^s = \frac{1}{|C|} \sum_{c \in C} Dist(S_c) \qquad (2.2)$$

[3] The current simulated model allows the accuracy of this measurement to be predefined (see section 3.5.2)

[4] The direction vector of a communication site is the vector from the centre of mass of the A-Cell to the communication site.

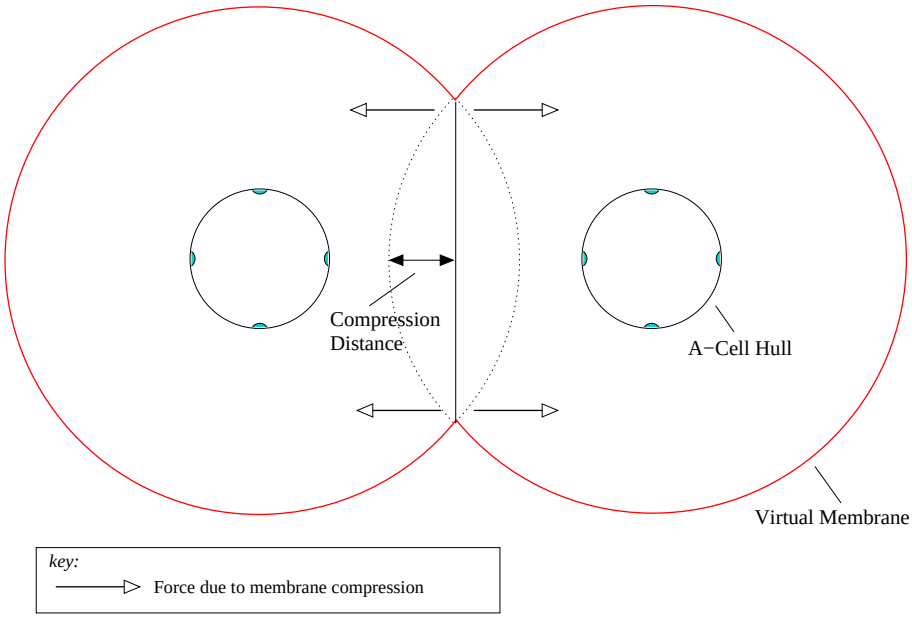


Figure 2.3: *Schematic showing the virtual membranes of two colliding A-Cells.*

where $Dist(S)$ determines the distance over which a signal with strength S has travelled.

This type of approximation has the added benefit of scalability, as any number of sensors may be added. In addition, section 3.5.3 demonstrates that not only is this approximation sufficient but also that the errors in the estimates can actually be beneficial by indirectly causing random movements.

2.2.3 Membrane Model

Section 1.2 identifies the cell-membrane as one of the most important structural components of the cell. For the purposes of this work, one of the key properties is its ability to deform, allowing cells to move fluidly around one another. Clearly without this property, the A-Cells would simply become locked in rigid configurations and sorting would cease.

The model provides a very simple approximation of a membrane by using the A-Cells' communication system to create a virtual force field which extends some dis-

tance from the actual A-Cell hull[5]. Like a fluid membrane, this force field behaves like a viscous spring, resisting compression from similar force fields generated by neighbouring A-Cells with a force proportional to the speed of the compression (see fig. 2.3). In practice, this viscous component prevents A-Cells from oscillating when an attraction force exists between them and, instead, allows them to settle quickly at some equilibrium.

To model this membrane (virtual force field) using the simple communication system, each A-cell transmits to its neighbours the radius of its physical hull and the distance its membrane extends from this hull. Therefore, by comparing the approximate distance of neighbouring A-Cells with their relative sizes, an A-Cell can detect membrane collisions and estimate the area of membrane that has been compressed. The amount each membrane compresses will be a function of the individual membranes' properties (imagine pressing a rubber ball against a steel one). However, in the current model, it has been sufficient to keep the membrane properties constant for all the A-Cells[6]. This means that the membrane of each A-Cell involved in a collision will compress by an equal amount (see fig. 2.3).

If we make the simplifying assumption that the membrane is under constant hydrostatic pressure, the repulsive force developed when a membrane is compressed will be proportional to the area of membrane under which the pressure acts. Therefore, it is possible to define the repulsive force induced by the compression as:

$$R = a * m \tag{2.3}$$

where R is the repulsion force acting along the line joining the two A-Cell centres (Kgms^{-2}), a the area of membrane (m^2) that has been compressed and m the repulsion factor ($\mathrm{Kgm}^{-1}\mathrm{s}^{-2}$, a system parameter to represent the pressure acting against the membrane). The area of membrane that is compressed is proportional to to the square of the compression distance (c). Therefore, as an optimisation this simple approximation is used.

[5]This provides maximum flexibility as it should allow any of the membrane properties to be changed dynamically throughout the course of a simulation.

[6]For an expansion of this work this may not be the case, and additional computation will be required.

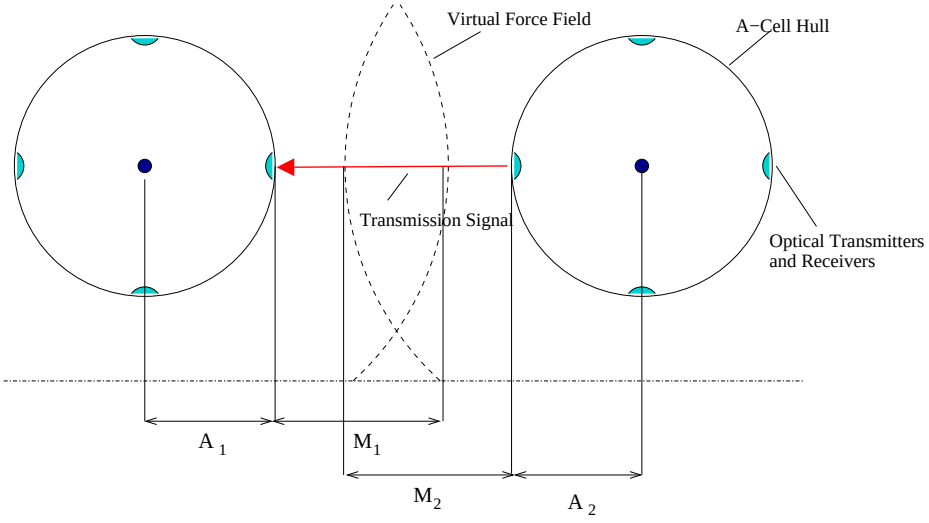


Figure 2.4: *Overlap of two A-Cell membranes. A – the actual radius of the A-Cell hull, M – the extra distance the membrane extends from the hull.*

$$R = c^2 * m \tag{2.4}$$

To add a viscous component to this equation, it is necessary to estimate the relative velocity with which any two A-Cells are travelling toward each other. As the A-Cells are constantly broadcasting their transmission signal, from which an estimated distance is determined, the calculation of this relative velocity is reasonably trivial and simply requires some comparison of the change in relative distance over some short time period.

Using this average relative velocity, the total repulsive force exerted on another A-Cell by the compression of an A-Cell's membrane is given by the following equation:

$$R = (c^2 * m) + (d * v) \tag{2.5}$$

where d is the damping coefficient (Kgs^{-1}, another system parameter to scale the velocity term appropriately) and v is the estimated relative velocity (ms^{-1}) concerned along the line joining their centres.

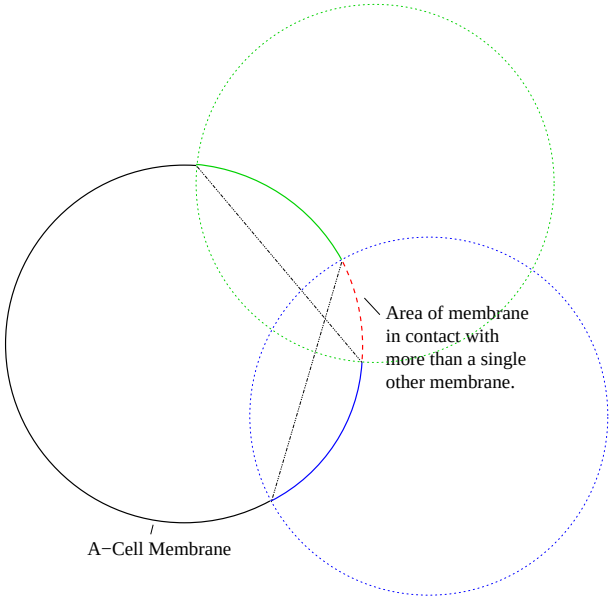


Figure 2.5: *An A-Cell membrane in contact with multiple other membranes.*

As with any collision between two bodies, the A-Cells should experience forces which are equal in magnitude and opposite in direction (Newton's third law of motion). However, as each A-Cell is simply using its own estimate of the direction and distance of the other, they might compute slightly different forces. To avoid this imbalance, the cells communicate the value they have calculated to each other. They do this by placing this value in the second part of their transmission signal with an identifier stating which other A-Cell the information is for (see 2.2.2). The A-Cells can then simply average their own calculated value with the value communicated to them, thus ensuring both A-Cells experience an equivalent force.

One flaw with this simple model is that each membrane collision is considered separately. Therefore, it is possible for the same area of membrane to collide with more than a single other membrane simultaneously. However, the principal repulsive behaviours of the membrane remain consistent.

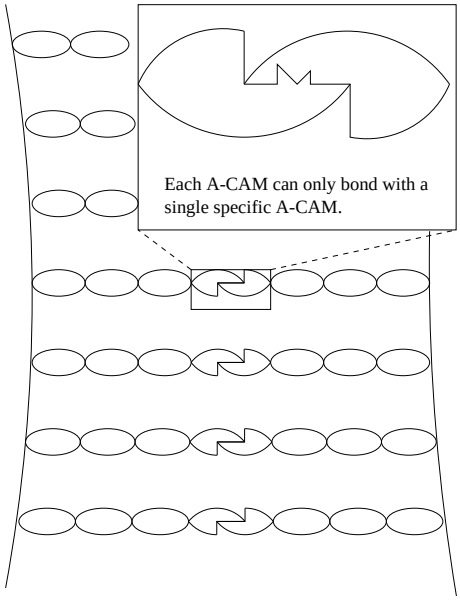


Figure 2.6: *A-CAM bonding.*

2.2.4 A-Cell Adhesion Molecules (A-CAMs)

The membranes of real cells use Cellular Adhesion Molecules (CAMs) to form bonds which are both reversible and selective (see section 2.1.3). We attempt to model these key properties through the use of A-CAMs. Like CAMs, the A-CAMs are capable of both homophilic or heterophilic binding. However, unlike some real CAMs each A-CAM may only bond with a single type (see fig. 2.6). This greatly simplifies the implementation, and as it should be possible to approximate the multiple bonding some CAMs exhibit by simply using several types of A-CAM, the power of the model is not affected.

When the A-CAMs form bonds they produce a small attractive force pulling the two areas expressing the A-CAMs closer together. Thus bonded A-Cell membranes are pulled together until the attractive force is balanced by the repulsive force generated by the compression (see fig. 2.7).

When two areas of real cell membrane come into contact, only some subset of the CAMs on each membrane will be close enough to CAMs on the other membrane to allow them to bond. The size of this subset is dependent on both the number of

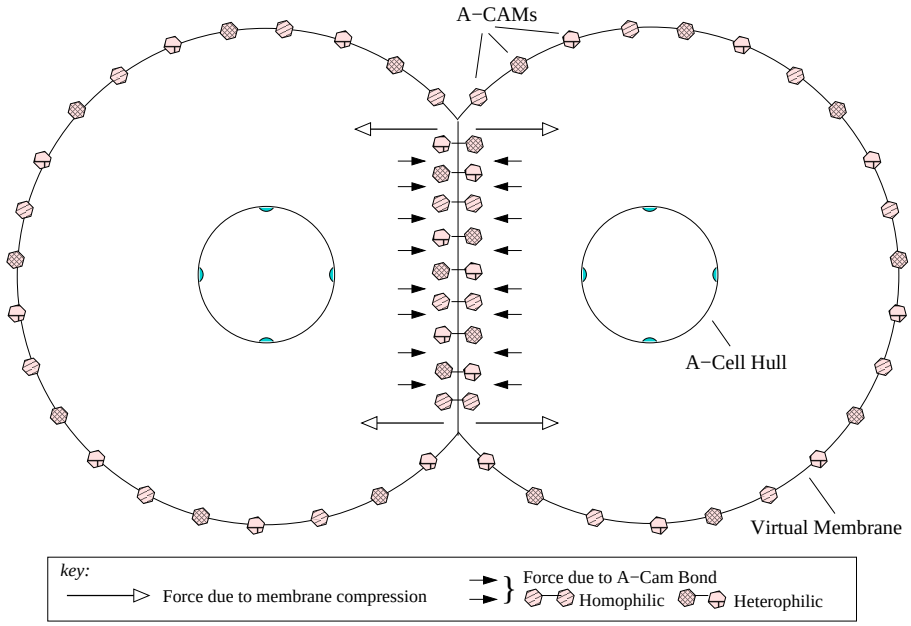


Figure 2.7: *Schematic of two bonding A-Cells.*

CAMs on each area of membrane and the strength of the bond between the CAMs. Greater surface densities of CAMs (number of CAMs per unit area) will increase the chance that any two are facing each other and a greater bond strength will allow the CAMs to seek each other out more assiduously. To express these relationships with the A-CAMs we constructed the following equation which satisfies all the observed constraints when relating numbers of available A-CAMs to numbers that bond (N).

$$N = \min(A_{\alpha 1}, A_{\beta 2}) \text{ sigmoid} \left(m(\alpha, \beta)k + \frac{(A_{\alpha 1} - A_{\beta 2})^2}{A_{\alpha 1} * A_{\beta 2}} \right) \tag{2.6}$$

where $A_{\alpha 1}$ is the number of A-CAMs of type α on the contact area of membrane 1, $A_{\beta 2}$ is the corresponding number of type β A-CAMs on contact area of membrane 2, $m(x,y)$ is the attractive force between A-CAMs of type x and y [7] and k is a system parameter that controls the influence of the bonding strength on bonding numbers.

For simplicity the model does not restrict the number of A-CAMs to an integer value.

[7] Each A-Cell is initialised with a table of the attraction values at the start of a simulation.

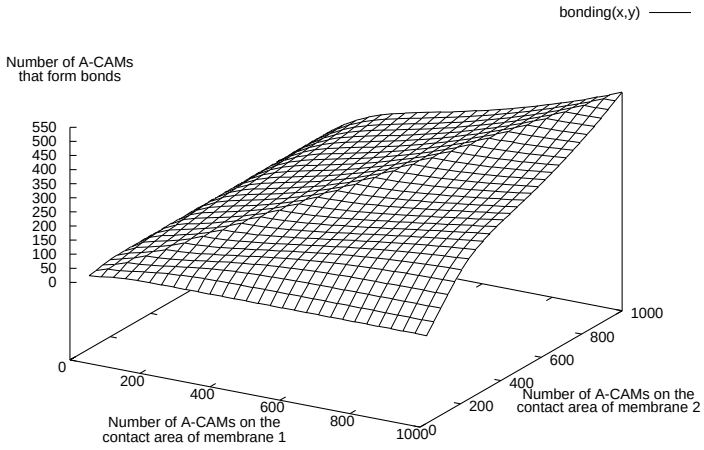


Figure 2.8: *Graph of Eqn. 2.6 with the scaling constant k set to 0.*

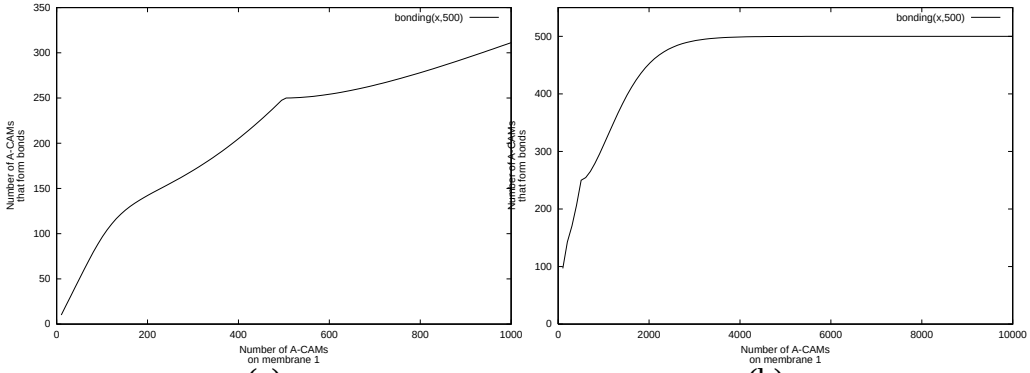


Figure 2.9: *(a) Graph of Eqn. 2.6 with the scaling constant k set to 0 and the number of A-CAMs on membrane 2 fixed to a value of 500. (b) Previous graph with a larger x axis range.*

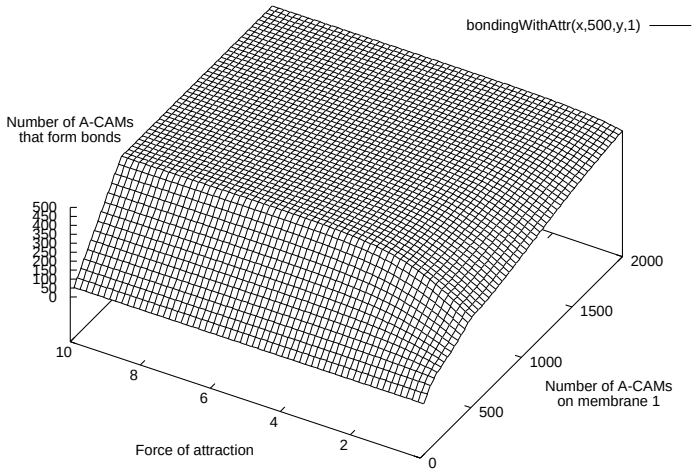


Figure 2.10: *Graph of Eqn. 2.6 with the number of A-CAMs on membrane 2 fixed to a value of 500 and the scaling constant k set to 1*

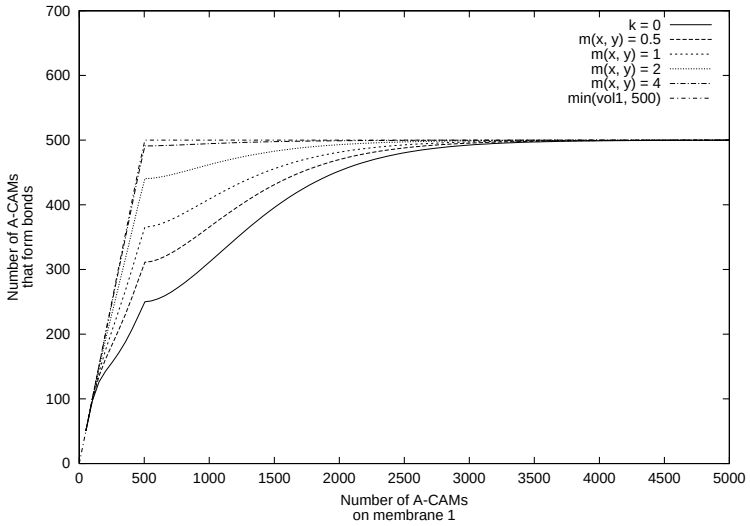


Figure 2.11: *Graph of Eqn. 2.6 with the number of A-CAMs on membrane 2 fixed to a value of 500 and a range of attraction values.*

Figures 2.8, 2.9, 2.10 and 2.11 illustrate the equation's behaviour. The graphs in figures 2.8 and 2.9 show the result of setting the scaling constant k to 0. This would mean that the attraction between A-CAM bonds has no effect and the total number of bonds formed depends only on the surface density of A-CAMs. In this case we should expect the equation to exhibit two behaviours. Firstly, if one membrane ($\mathcal{A}$) has a fixed A-CAM surface density and the A-CAM surface density of the other membrane ($\mathcal{B}$) increases, then the number of A-CAMs that bond should increase. This behaviour is shown in figure 2.9 – a (the slight bump in the function is introduced by the dynamics of the sigmoid term). Clearly, as the A-CAM surface density of membrane $\mathcal{B}$ approaches infinity, the number of A-CAM bonds formed should tend to the minimum of the numbers of A-CAMs present on each area of membrane (see fig. 2.9 - b).

The second behaviour we would expect is that increasing the minimum of the numbers of A-CAMs present on the membranes should cause the number of A-CAMs that form bonds to increase. This behaviour is visible in the graph shown in figure 2.8.

The graphs in figures 2.10 and 2.11 illustrate the equation's behaviour when the scaling constant k is set to a positive value greater than zero[8]. As stated earlier, if the force of attraction between two types of A-CAM increases, the number of A-CAMs that form bonds should also increase. This behaviour is clearly depicted by both graphs. Additionally, it can be seen that as the attraction value tends to infinity, the sigmoid term tends to one. In this case, the number of A-CAMs that form bonds is equivalent to the minimum of the numbers of A-CAMs present on each area of membrane.

2.2.5 Adhesion Sites

The A-Cell membrane model allows the same area of membrane to be in contact with more than a single other membrane simultaneously (see fig. 2.5). However, to achieve more realistic selective bonding, the bonding of each area should be restricted to only one of the membranes in contact.

This is achieved by dividing the surface area of the membrane into a number of

[8] Actually the scaling constant k is set to exactly 1 in both figure 2.10 and figure 2.11, but they illustrate the general case.

equally sized patches or adhesion sites, such that each adhesion site can only bond with a single membrane. Therefore, it can be seen that in order to determine the attractive force which exists between two membranes, it is necessary to calculate firstly the subset of adhesion sites that decide to bond with each membrane, and secondly the number of A-CAMs this represents.

The second of these calculations is relatively straightforward as the A-CAMs are considered to be evenly distributed across the membrane surface[9]. Therefore, the number of A-CAMs of type t present on the surface of each adhesion site (ACAM_t^S) is given by:

$$\mathrm{ACAM}_t^S = \frac{\mathrm{ACAM}_t^T}{n} \tag{2.7}$$

where ACAM_t^T is the total number of A-CAMs of type t which are expressed over the whole membrane and n is the total number of adhesion sites. A simple multiplication of ACAM_t^S by the number of bonding adhesion sites gives the total number of A-CAMs of type t that are able to bond.

Calculating which adhesion sites should bond to which other membranes is a more complex problem. Figure 2.12 outlines the algorithm used. The remainder of this section describes some of the algorithm steps in more detail.

In order for the adhesion sites to decide which membrane they should bond to, there needs to be some measure of the possible attraction they would have toward each of them. In the algorithm this is referred to as the 'estimated attraction value' (Att^E), and is equivalent to the total expected attraction force per unit area of bonded membrane. To calculate this value, an A-Cell must determine the surface density of each type of A-CAM (α) present on its membrane[10] ($\alpha \in \mathcal{A}$) and each type of A-CAM (β) on the membrane in contact with it ($\beta \in \mathcal{B}$). It is then possible to calculate the number of bonds formed between each bonding pair of A-CAM types (N see equation 2.6). Finally, each of these calculated values is multiplied by the attraction value for the respective pair of A-CAM types (m), and the calculated forces are summed to give

[9] This is a simplification and may not necessarily be true for the CAMs in real cells.

[10] This is simply the total number of each type of A-CAM divided by the surface area of the membrane.

```
1. For each membrane in contact with an A-Cell:

    (a) Calculate the estimated attraction between the A-Cell's
        membrane and the membrane in contact.

    (b) Calculate which of the A-Cell's adhesion sites are in contact
        with the other membrane.

    (c) For each adhesion site in contact:

         i. Store the A-Cell id associated with the other membrane and
            its estimated attraction value.

2. For each of the A-Cells adhesion sites:

    (a) Probabilistically decide which membrane to bond to, based on
        their respective estimated attraction values.

3. For each membrane in contact with an A-Cell:

    (a) Calculate the number of the A-Cell's adhesion sites that have
        decided to bond with this membrane.

4. END.
```

Figure 2.12: *Adhesion Site decision algorithm*

the final estimate.

$$Att^E = \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{B}} m(\alpha, \beta) N(\alpha, \beta) \tag{2.8}$$

It can be seen that for the A-Cells to perform this calculation, they require knowledge of the total number of each type of A-CAM present on each membrane making contact. Therefore, these values are included in the information which is broadcast in the first half of each A-Cell's transmission signal and are, therefore, available to any A-Cell in range.

After all the estimated attraction values have been calculated and stored with the relevant adhesion sites, the adhesion sites must decide which A-Cell membrane to bond to. There are two possible options:

- Greedy – The membrane with the highest estimated attraction value is selected.
- Probabilistic – A fitness proportional selection is made with the fitness determined by the estimated attraction value. Therefore, the probability that adhesion site *s* attempts to bond with membrane *m* is given by:

$$P_{bond}(s, m) = \frac{Att^E_m}{\sum_{\alpha \in \mathcal{M}} Att^E_\alpha} \tag{2.9}$$

where $\mathcal{M}$ is the set of all membranes *s* must choose between.

If a greedy approach is applied, the bonding would be completely deterministic. This is unlikely to be a good model given the indeterministic nature of real biological processes. In addition, simply by considering the dynamics of an A-Cell aggregate, it can be seen that if this approach were to be used in circumstances where the heterotypic adhesion was only slightly less than the highest homotypic adhesion, there would still be a clear bias toward the latter bonds forming. This would lead to partial or possibly complete sorting even though the DAH suggests that the A-Cells should remain evenly mixed. To overcome this, the more realistic probabilistic approach is

used (see fig. 2.12 - step 2a). Therefore, when any two membranes have an equivalent estimated attraction value each will have a fifty percent chance of bonding. As the bonding between A-Cells is reevaluated at every cycle in the simulation, over an extended period of time and excluding the other dynamics of the model, each membrane should bond for an equivalent time period.

Finally, the adhesion sites bonding to each membrane are counted, and the values are used to help calculate the overall force of attraction the A-Cell should experience toward each of the A-Cells making contact (see section 2.2.6).

2.2.6 A-Cell Bonding

The general algorithm for calculating the bonding between two A-Cells (that is the force of attraction between them – see 2.2.4) is given in figure 2.13. This section describes the whole process in more detail.

The algorithm cycles through each of the membranes that are making contact with the A-Cell and determines the force of attraction experienced toward each of them[11]. For each membrane, the A-Cell first has to determine what number of each of its A-CAM types will be available to bond to the membrane. This step is discussed in full in section 2.2.5. It is then necessary to determine the number of A-CAMs on the membrane in contact which are available to bond. In our implementation this is achieved by placing the available A-CAM numbers and the ID of the cell they can bond with in the second half of the 'transmission signal'. Therefore, when any two membranes collide, the A-Cells will each have access to the relevant information.

Once the numbers of each type of A-CAM on the respective membranes have been calculated, it is possible to calculate a total force of attraction. This involves cycling through each pair of A-CAM types and first checking if there is a non-zero attraction force between them[12]. If there is a non-zero attraction between two A-CAM types, the algorithm proceeds by first establishing the number of each type of A-CAM that will actually form bonds (see section 2.2.4), and then multiplying this value by the attraction value of a single bond. This gives a force of attraction that will be generated

[11] The details of how these forces are applied are covered in section 2.2.7.

[12] A zero attraction force indicates that the two A-CAM types do not bond, and therefore there is no point in carrying out any further calculations relating to them.

1. Determine which of an A-Cell's adhesion sites are bonding to each of the membranes in contact with the A-Cell (see fig 2.12).

2. For each membrane in contact with an A-Cell:

   (a) Calculate the number of each type of A-CAM that is available to bond to the other membrane.

   (b) Communicate each of these calculated values to the A-Cell that the membrane belongs to.

   (c) Wait until the numbers of each type of A-CAM the other A-Cell has available to bond are received.

   (d) For each pair of A-CAM types:

      i. If there is a non-zero attraction force between the two types:

         A. Calculate the number of bonds that will be formed (see eqn. 2.6).

         B. Multiply the calculated value by the attraction value.

         C. Store the computed force.

3. END.

Figure 2.13: *A-Cell Bonding Algorithm*

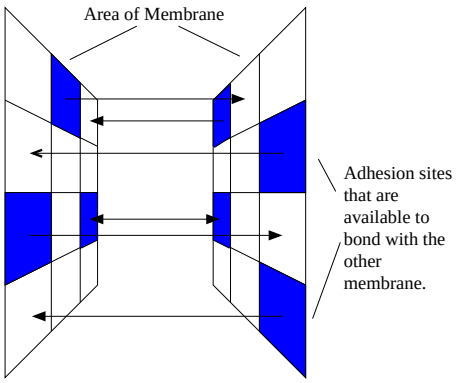


Figure 2.14: *Figure showing a positional miss-match between bonding adhesion sites.*

by each bonding pair of A-CAM types and thus, by summing all these values, the total force of attraction that the A-Cells should experience toward each other.

At this point it is important to highlight two big simplifications the algorithm makes: The first of these is that it is assumed that the adhesion sites on each membrane that wish to bond will be facing areas on the other membrane that also wish to bond. However, it is clear that this might not be the case (see fig. 2.14). One reason for this simplification is that each membrane in contact is considered independently of the others. Therefore, the position that the adhesion sites should actually be in is not clear as the same area of membrane may be compressed by different amounts for each of the membranes it is in contact with. Also, calculating whether specific adhesion sites did face areas of membrane that are also available to bond would be computationally expensive, and would require far more inter-A-Cell communication.

The second simplification is that A-CAMs that an A-Cell allocates to bond with another membrane are not reallocated if they are unused in the bonding process. For example, if certain numbers of each type of A-CAM on an A-Cell's membrane are allocated to bond with another A-Cell and that A-Cell allocates nothing, then those A-CAMs will not take part in any bond for that time step.

These simplifications have been made as they do not break any of the intrinsic constraints of the problem (see 2.1.1), but they do prevent the model from becoming overly complex.

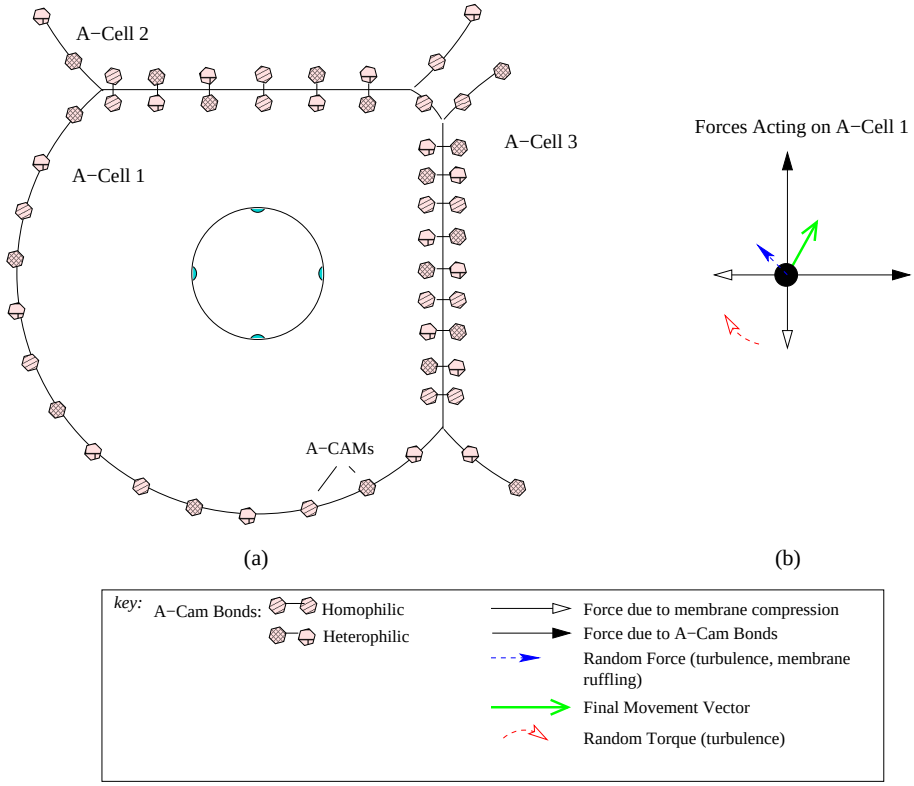


Figure 2.15: *Example of the forces acting on an A-Cell. (a) Schematic showing the membrane collisions made by an A-Cell (A-Cell 1). (b) A representation of all the forces acting on the A-Cell (A-Cell 1). The direction of forces due to membrane compression and the forces due to A-CAM bonds will depend on the estimated position of the neighbouring A-Cells not their actual position.*

2.2.7 A-Cell Movement

In this model the movement of the A-Cells is completely passive and is simply governed by the forces they experience. At each time step the membrane contacts are determined and any attractive forces generated by bonding (see section 2.2.6) or repulsive forces generated by membrane compression (see section 2.2.3), are calculated. These forces are then combined to produce a single force vector using the algorithm outlined in figure 2.16 and this is applied until the force is updated at the following time step (see fig. 2.15). Therefore, the A-Cells are driven by a piecewise constant

force with each step corresponding to a single time step. However, section 2.1.1 proposed that for full sorting to take place, this passive movement must also include some random component allowing local exploration. In real cells this random movement is a result of membrane ruffling[40, 48], localised expansions and contractions of the cell membrane (see section 2.1.3). As our simple model of the cell membrane is a uniform sphere, the exact nature of the ruffling cannot easily be reproduced. However, the result of this phenomena is simply that the cells make small random movements. Therefore, an equivalent behaviour can be achieved by adding some random component directly to the A-Cell movements. In the proposed model this component consists of an additional random movement force, the magnitude of which is scaled by a predefined system parameter. Therefore, both active and inactive membranes can be modelled.

To add accuracy to the physical simulation of the A-Cells, the properties of the fluid environment they operate in also has to be considered. This is achieved by adding realistic drag forces to the A-Cell hulls (see section 2.3.2) and allowing some level of turbulence to be introduced.

Clearly, some aspect of the turbulence present in the fluid can be accounted for by the random movement forces outlined previously. However, small random rotations of the A-Cells also have to be considered. As the A-Cells can only rotate around the vertical axis this can be achieved by adding appropriately directed random torques. Once more, in the actual model this is scaled by a predefined system parameter.

2.3 Model Implementation

This section begins with specific details of the implementation of the model used for the experiments which are discussed throughout this thesis. The final part of this section validates the implemented model, showing that it behaves in accordance with the theoretical behaviour discussed in section 2.2.

2.3.1 Application of Forces

The only forces the A-Cells produce are the movement forces determined by the virtual adhesions and repulsions they experience (see section 2.2.7). Each force is maintained

1. Create an accumulated force vector of strength 0.
2. For each membrane in contact with the A-Cell:
    (a) Normalise the direction vector associated with the A-Cell producing the other membrane.
    (b) Calculate the repulsive force generated by the membrane collision.
    (c) Multiply the normalised direction vector by the repulsion value to produce a force vector.
    (d) Subtract this force vector from the accumulated force vector.
    (e) For each A-CAM pair that produce an attractive force:
        i. Multiply the normalised direction vector by the attraction value to produce a force vector.
        ii. Add this force vector to the accumulated force vector.
3. Apply the accumulated force vector to the A-Cell (that is, apply a force such that the A-Cell experiences a force equivalent to the force vector).
4. END.

Figure 2.16: *A-Cell Movement Algorithm*

for a single time step of the simulation and is then replaced by the newly calculated force which reflects the A-Cell's new state.

This behaviour is implemented by applying the respective piecewise constant force to each A-Cell's centre of mass via function calls to the physics engine [13]. The physics engine's internal integrator then ensures that these forces are realistically modelled such that, in the absence of turbulence or drag, the A-Cells behave as predicted by the standard equations of motion.

Each step of the integrator corresponds to a single time step in the simulation.

2.3.2 Fluid Drag

To ensure the behaviour of the A-Cells is as realistic as possible, some of the additional effects of fluid dynamics have been included in the simulation. This section gives the full details of the drag calculations. However, a simple model of turbulence is also included as discussed in section 2.2.7.

The implementation of the calculations presented in this section was carried out by Timothy Taylor as part of his work on the E.U. funded Hydra project [61].

All the calculations are based on a fluid environment with the properties of water as this is the likely operating environment of any physical implementation of the model. More specifically, the values are based on room temperature conditions ($20^{\circ}C$).

- Water Density: $1000kg/m^3$
- Water Viscosity: $0.001kg/ms$

The calculation of the main drag force (F_d) acting on an object is based on the following accepted equation:

$$F_d = 0.5\rho v^2 A C_d \tag{2.10}$$

where an object with a projected (cross-sectional) area of A (m^2) moves through a fluid of density ρ (kg/m^3) with a velocity of v (ms^{-1}). Therefore, for the spherical A-Cells with radius r the drag is given by:

[13] The commercial 'Vortex' physics engine V2.0 produced by CM Labs

$$F_d = 0.5\rho v^2 \pi r^2 C_d \tag{2.11}$$

The parameter Cd is the dimensionless drag coefficient. However, as it accounts for all the complexities of the flow around an object, it does not remain constant. Instead, another dimensionless number called the Reynolds number (R_e) is used to characterise the C_d of an object based on its relative velocity:

$$R_e = \frac{L\rho v}{\eta} \tag{2.12}$$

where an object with length L (m) along the direction of flow, moves through a fluid of density ρ (kg/m^3) and viscosity η (kg/ms) with a velocity of v (ms^{-1}) (relative to the fluid).

For the calculation of the C_d of a sphere the following three regions of distinct behaviour have been identified [57]:

$$R_e \begin{cases} < 1, & C_d = 24/R_e \text{ (laminar flow)} \\ > 1 \text{ and } < 1000, & C_d = 18R_e^{-0.6} \text{ (transition)} \\ > 1000 \text{ and } < 2x10^5, & C_d = 0.44 \text{ (turbulent flow)} \end{cases} \tag{2.13}$$

These regions were used to dynamically determine the appropriate C_d and thus the appropriate drag force acting on the A-Cells during each simulation.

An additional drag force which must also be considered is *added mass*. This is the additional volume of fluid that must be accelerated with a body when it experiences non laminar flow ($R_e > 1$, see 2.13). For a sphere with volume V, the additional mass (m_a) is equal to half that of the displaced volume [11]. Therefore,

$$m_a = 0.5V\rho \tag{2.14}$$

where ρ is the density of the fluid.

This additional mass is included in the simulation by adjusting the actual mass of the A-Cells (m_r) accordingly:

$$m_t = m_r + m_a \tag{2.15}$$

where m_t is the total mass associated with each A-Cell in the simulation.

Both the drag force and the added mass are applied to the A-Cells through the physics engine's callback mechanism. This ensures they are updated at each stage of the integration carried out by the physics engine between any two time steps[9].

2.3.3 Validation

The previous sections in the chapter have presented the cell adhesion model in some detail. In this section, the behaviour predicted by this model is compared with the behaviour observed in the actual software implementation used throughout the thesis. This comparison aims to show that the implementation of the model is valid and thus that the following experiments demonstrate genuine behaviour.

The comparison is based on a simple two A-Cell simulation which was carried out in the absence of turbulence. As there are only two A-Cells, resulting in a single membrane overlap, the adhesion sites are not required to make any probabilistic decisions. Therefore, the simulation is completely deterministic and should converge to the stable state where the A-Cells are the equilibrium distance apart.

As in the following chapter, the A-Cell movements were restricted to the horizontal plane. However, this restriction only affects the placement of the sensors and adhesion sites, which can then be distributed around the respective equators of the hull and membrane. The implementation of the model is unaffected (see Chapter 3 for details).

For consistency, the default simulation parameters discussed in Chapter 3 (see section 3.1.3) were also used. Of these, the values which are significant to the calculations presented in this section are:

- A-Cell Physical Hull Radius (R_h) = $0.05m$
- Membrane Size (M_s) = $0.5m$
- Number of Adhesion Sites (A_t) = 64 distributed around the equator of each A-Cell's membrane.

- Membrane Repulsion (m) = $2Kgm^{-1}s^{-2}$
- Bonding Equation Scaling Constant (k_s) = 1

The A-Cells in the simulation were initially placed 1m apart, such that their membranes were only just in contact. In addition, their sensors were aligned to position two sensors on each A-Cell along the line formed by the A-Cell centres. The simulation was then executed until the A-Cell positions converged to a stable state.

The following calculations demonstrate how the theoretical number of A-CAMs that will result in a required equilibrium distance can be determined. This calculated value is then used in the actual simulation and the resulting behaviour is presented. For this comparison a required equilibrium distance (D_e) of $0.55m$ is used.

The first step in the calculation is to determine the total repulsive force (F_r) that will result from the membrane compression when the two A-Cells are the equilibrium distance apart. As the forces are balanced this value should equal the attractive force generated by bonds between the A-CAMs on the areas of membrane which are in contact.

The repulsive force is calculated from the compression distance (D_c) using equation 2.5. However, by assuming that the two A-Cells will be stationary at the equilibrium distance, the viscous component can be ignored.

$$D_c = \frac{2R_v - D_e}{2} \quad (2.16)$$
$$= (2 * 0.55 - 0.55)/2 \quad (2.17)$$
$$= 0.275m \quad (2.18)$$

$$F_r = (D_c^2 * m) \quad (2.19)$$
$$= 0.275^2 * 2 \quad (2.20)$$
$$= 0.15125N \quad (2.21)$$

The number of A-CAM bonds required to balance this force (B_r) can easily be calculated from the A-CAM types that are expressed by each A-Cell. For simplicity, in this simple simulation each A-Cell expresses only a single type of A-CAM and the attractive force formed by a single bond (F_b) is set to $0.0001N$. Therefore:

$$B_r = \frac{F_r}{F_b} \quad (2.22)$$
$$= \frac{0.15125}{0.0001} \quad (2.23)$$

$$= \quad 1512.5 \tag{2.24}$$

To calculate the actual number of A-CAMs that must be in contact (ACAM_c for the required number of bonds to be formed, equation 2.6 is reorganised as follows:

$$\min(A_{\alpha 1}, A_{\beta 2}) = \frac{B_r}{\text{sigmoid}\left(m(\alpha,\beta)k + \frac{(A_{\alpha 1} - A_{\beta 2})^2}{A_{\alpha 1} * A_{\beta 2}}\right)} \tag{2.25}$$

where $m(\alpha,\beta) = F_b$.

If each A-Cell simply expresses an equal amount of the A-CAM this can then be simplified to give:

$$\begin{aligned} \text{ACAM}_c &= \frac{B_r}{\text{sigmoid}(kF_b)} && (2.26)\\ &= 1512.5/0.500025 && (2.27)\\ &= 3024.84876 && (2.28) \end{aligned}$$

Finally, it is necessary to calculate the number of A-CAMs that must be expressed across each whole membrane (ACAM_t), such that the required number will be present on the areas that are in contact. As the bonding between A-Cells is approximated by using adhesion sites, this value must be obtained by first calculating the number of A-CAMs that will be associated with each site. This value can be calculated by using a simple geometric test to determine how many adhesion sites will be on the contact area (A_c)[14]. In this simple simulation at the equilibrium distance of $0.55m$, twenty one of the sixty four adhesion sites are present on the contact area. Therefore:

$$\begin{aligned} \text{ACAM}_t &= \frac{\text{ACAM}_c}{A_c} * A_t && (2.29)\\ &= (3024.84876/21) * 64 && (2.30)\\ &= 9218.5867 && (2.31) \end{aligned}$$

[14] In this implementation, the plane perpendicular to the line through the A-Cell centres which passes through the A-Cell membrane at the compression distance is calculated. Each adhesion site is then tested to determine which side of the plane it lies on.

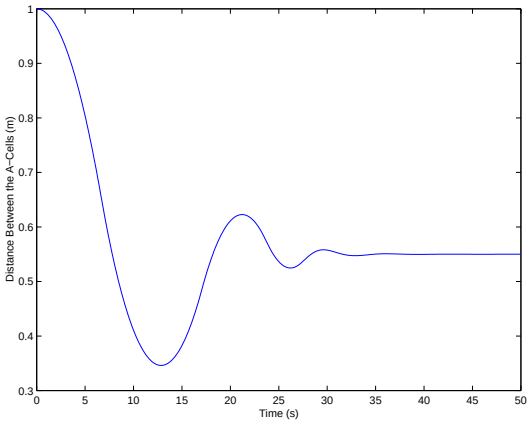


Figure 2.17: *The calculated distance between two adhering A-Cells.*

The results of the simple simulation using this value are shown in figure 2.17. It can be seen that the A-Cells initially oscillate. However, the viscous component of the membrane repulsion in conjunction with the fluid friction, damps the oscillation allowing the A-Cells to achieve a steady state at the required equilibrium distance of $(0.55m)$.

These results demonstrate that the implementation behaves as the theoretical model predicts.

2.4 Summary

This chapter has presented a robust distributed control method which is based on the cell adhesion mechanisms employed by biological cells. To ensure the model behaves as predicted by the DAH, the key constraints have been identified and incorporated into the model's design. Furthermore, to demonstrate the feasibility of the model as a control mechanism for populations of mobile robots and to provide a testbed for experimentation, the model has been implemented on simulated robotic units (A-Cells) with physically realistic abilities.

The critical component of the model is the membrane which surrounds each of the A-Cells. The membrane takes the form of a virtual force field which extends some distance from the actual A-Cell hull and interactions between neighbouring A-Cells

only occur when their membranes collide. These collisions result in two distinct forces:

1. A repulsive force which models the restorative force that is generated by the compression of a real cell membrane.

2. An attractive force which models the adhesions that occur when two cell membranes come into contact.

The repulsive force is simply proportional to the area of membrane which would have been compressed by the collision. However, the attractive force depends on the number of adhesive elements (A-CAMs) which appose each other on the areas of membrane that are in contact. The combination of these forces means that any two A-Cells will be attracted closer together until they reach some equilibrium distance, where the repulsive force balances the attraction.

It was highlighted that one drawback of this simple membrane model is that the same area of membrane on one A-Cell may be in contact with multiple other membranes simultaneously. However, a novel approach was introduced where the membrane is divided into a number of individual adhesion sites. These sites then record which membranes they are in contact with and probabilistically choose which one to bond to. In this way, the bonding remains realistic.

Chapter 3

Exploration of 2D Rounding

In section 1.3 we discussed how the behaviour of an aggregate of cells can be described in terms of the minimisation of free energy. In the simplest case this results in the rounding of an aggregate as it attempts to form the most compact structure.

In this chapter, the rounding behaviour of the cell adhesion model is investigated for the simplified case where A-Cell movements are restricted to the horizontal plane. The basic measures and techniques are presented, followed by a detailed investigation of some key model parameters.

3.1 Methodology

3.1.1 Measures

3.1.1.1 Measure of Rounding

Rounding is the process whereby the structure formed by an aggregate of cells approaches that of a sphere (disc in two dimensions), thus maximising the adhesions between the cells (see section 1.3). Throughout this chapter and later in section 5.2 we will examine our model's ability to replicate this behaviour and determine the effects of modifying its various parameters. For this purpose we therefore require some measure of 'roundness', and more specifically one which satisfies the following three constraints:

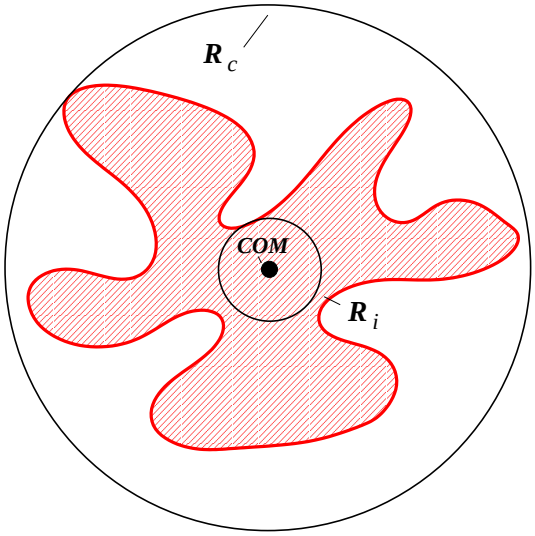


Figure 3.1: *Sphericity measure - inscribing and circumscribing circles.*

1. **The measure should represent the compactness of an A-Cell aggregate.**

2. **The measure should be applicable to both two and three dimensional objects.**

   This will allow the same measure to be used for both the initial two dimensional experiments and the later three-dimensional experiments (see section 5.1.1), thus enabling comparisons between them.

3. **The measure should be scale invariant.**

   As both the size of individual A-Cells and the number of A-Cells in a given aggregate are variable, it is essential that the rounding measure is independent of the overall aggregate size.

Two measures which meet these constraints are outlined below. Both measures produce an intuitive index which varies between zero and one, such that a perfect sphere or disc gives the maximum value.

- **Sphericity:**

$$S = \frac{R_i}{R_c} \tag{3.1}$$

where R_i and R_c give the radius of the inscribing and circumscribing circles or spheres which are concentric at the object's centre of mass COM (see fig. 3.1).

- **Compactness:**

$$C = \frac{4 * \pi * a}{p^2} \tag{3.2}$$

where a is the area of the aggregate (m^2) and p its perimeter (m).

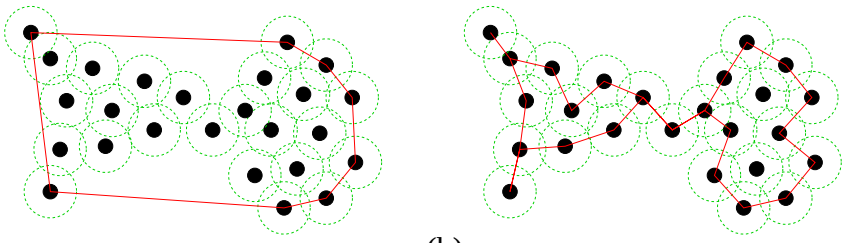
It should be noted that one failing of both these measures is that they ignore any empty space present in an object. Thus, a ring configuration will receive the same compactness index as a disk. However, the dynamics of our model mean that it is extremely unlikely that aggregates containing empty space will arise except through introduction in the initial configuration. Additionally, any empty space which is introduced will be rapidly eroded. Therefore, it was thought that the extra complexity involved in identifying this space in these rare instances would be of little benefit.

Each of the proposed measures require properties which involve first calculating an A-Cell aggregate's boundary. One approach to this is to consider such an aggregate as a collection of discrete vertices with each vertex located at the centre of an A-Cell. It is then possible to simply calculate the convex hull [1] (see fig. 3.2 - a). However, as the convex hull smooths out many of the irregularities present in the actual aggregate boundary, this would produce very biased results. An alternative possibility is to use additional information about the contacts between the A-Cells to calculate a connected graph of the aggregate. Then, using the algorithm described in appendix A.1, the more accurate concave hull[2] (see fig. 3.2 - b) of the graph can be calculated.

The graphs in figure 3.3 show how the indices calculated from the two roundness measures and the two boundary approaches vary over a small section of a complete

[1] The smallest convex shape that contains all the vertices.

[2] We define the concave hull of a connected graph as the smallest shape containing all the graph nodes that can be constructed from the edges of the graph



(a) (b)

Figure 3.2: *(a) Convex hull of an A-Cell aggregate; (b) Concave hull of an A-Cell aggregate*

simulation. As expected, the convex hull boundary produces a higher estimate of the aggregate roundness than that indicated by the concave hull boundary, most notably during the early part of the simulation when the aggregate's boundary is most irregular. The sphericity graph, however, produces a lower estimate with more extreme fluctuations. This behaviour is most likely to be caused by the sphericity measure's reliances on the position of only two of an aggregate's A-Cells relative to its centre of mass. Therefore, any small fluctuation in only these positions will have a large effect on the calculated index.

By comparing the graphs with the actual simulation data it is possible to observe further failings with the sphericity measure. Firstly, as the initial state of the simulation is a perfect line (see fig. 3.4) we would expect the initial index to be zero. However, as our implementation of this measure only considers inscribing and circumscribing spheres which pass through the A-Cell centres, when the aggregate consists of an even number of A-Cells the initial centre of mass will fall between two of them. Therefore, the radius of the inscribing sphere and thus the calculated index will be non-zero. Secondly, the simulation clearly shows that the aggregate rounds rapidly during the first 1000 seconds (see fig. 3.4). However, although this behaviour is shown clearly in the compactness graph, there is no observable increase in the sphericity graph during the same time period. Finally, the sphericity measure shows several large peaks which do not represent any significant events in the simulation (e.g. at approximately 1800 seconds).

Contrary to this, the compactness measure based on the concave hull boundary appears to closely follow observed changes in the aggregate's structure and shows the

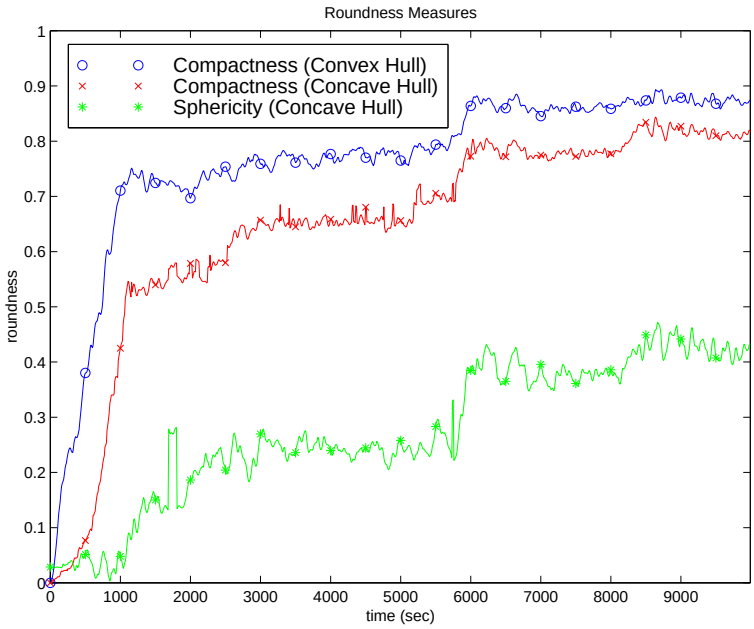


Figure 3.3: *A graph of different roundness measures.*

steady increase in the aggregate's roundness as time progresses. Therefore, it is this more representative measure which is used throughout this work.

3.1.1.2 Compacting Rate, Converged Compactness, and Compactness Variance

Simply comparing all the compactness versus time graphs discussed in 3.1.1.1 would make it very difficult to observe how changes in parameter values affect the model's rounding ability. Instead, it is necessary to decompose these results into their key values which can then be plotted in a single graph, thus revealing any general trends.

It can be seen from longer simulations that the A-Cell aggregates tend to round very rapidly at first before converging at some maximum value (see fig. 3.5). Therefore, it is possible to summarise this behaviour with the following two measures.

1. The rate at which the A-Cell aggregates round.

2. The convergence value achieved.

When calculating the rate at which the A-Cells round, we are only interested in the rate until the roundness value converges. An efficient method for estimating this

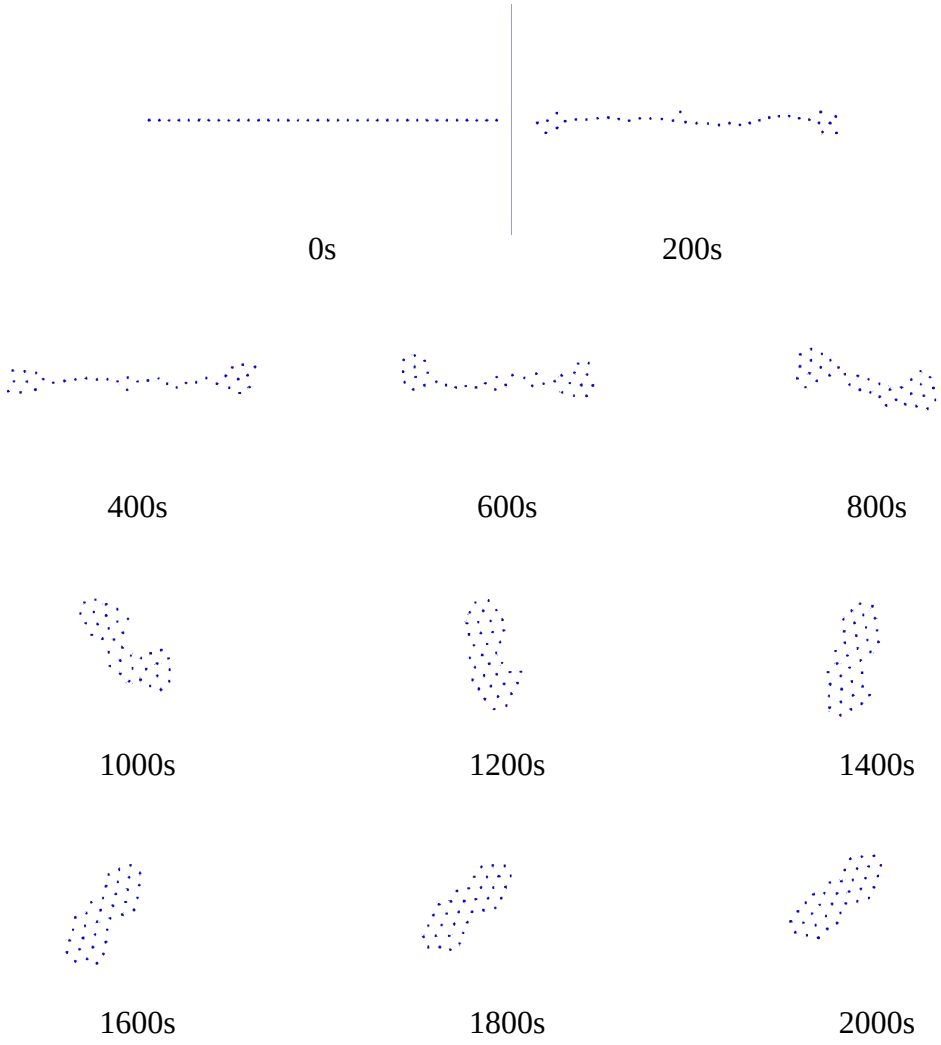


Figure 3.4: *Simulation of the A-Cell Model*

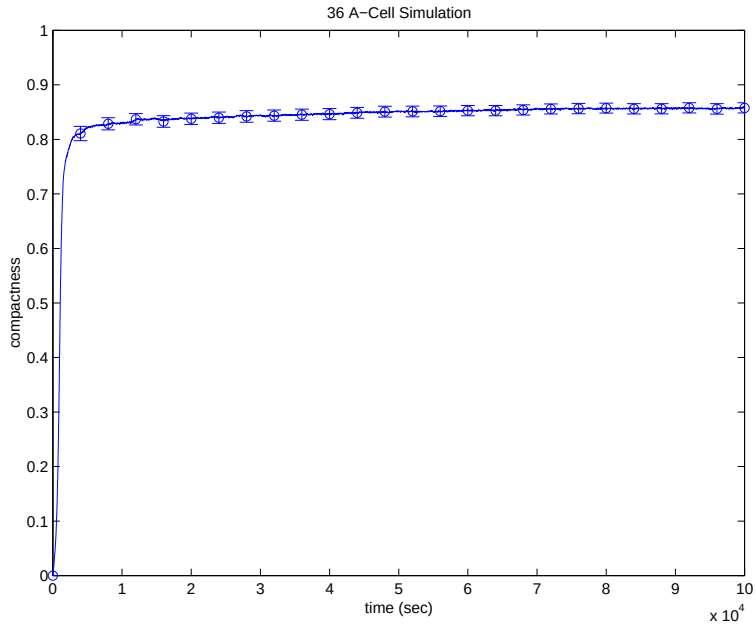


Figure 3.5: *A-Cell rounding with an aggregate of 36 A-Cells initially configured in a straight line: sample size = 100.*

rate is to use a variation of the 'rise time' which estimates the time it takes signals to change from a specified low value to a specified high value (typically 10% and 90% of the final value) (see fig. 3.6). This can be modified to give the 'convergence time' by taking the full time it takes until the value has reached 90% of the convergence value (see fig. 3.6). As this is a conservative estimate of the full convergence time, this period should only include the initial rounding stage of the simulation. The average compacting rate (CR) until convergence can then be calculated by dividing the 90% limit by the convergence time.

The method we employ to calculate the converged compactness (CC) is to simply take the mean of the compactness indices for the last *n* seconds. By selecting a sufficiently large value for *n*, any minor fluctuations in compactness will be ignored. Obviously, if the simulation terminates before the compactness index has converged this value will only give some lower bound on the true convergence value. However, the average rate of rounding until convergence should remain reasonably accurate. In addition, as a relatively long simulation time will be used (see section 3.1.3), these

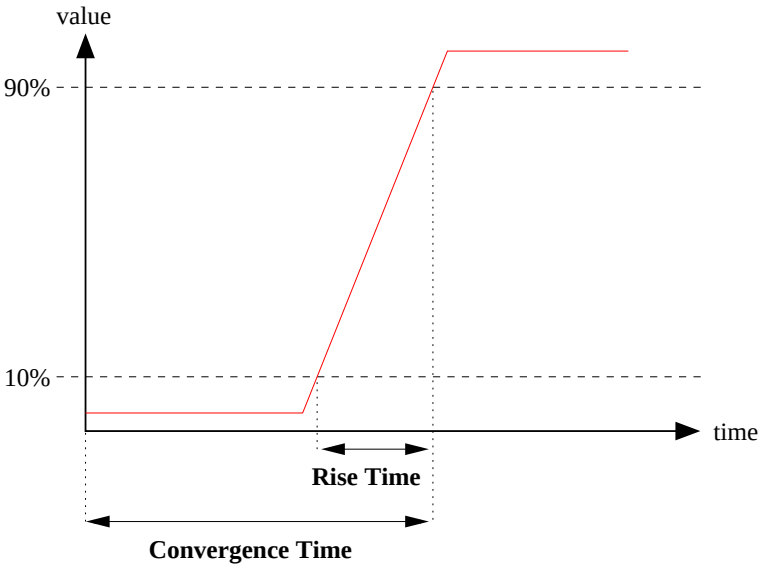


Figure 3.6: *The rise time and convergence time of a graph.*

cases should be rare and will at least be indicative that the chosen parameter set is very poor.

An additional measure which can also be calculated from the original data is the variance in compactness values. This variance gives an impression of the stability of an aggregate's shape. Therefore, calculating the variance of those compactness values from which the convergence value is derived (i.e. when the major rearrangements in the aggregate's structure have ceased) indicates how vigorously the individual A-Cells are moving. Figure 3.16 calculated from the random force results show how this value increases as the A-Cells make increasingly large random movements.

3.1.2 Initial Configurations

To give a lower bound on the model's performance, we use the worst case starting configuration in each of the simulations. As the experiments are designed to investigate the performance of the model with respect to rounding, the worst case starting configuration is a perfect line which has a compactness index of 0 (see section 3.1.1.1).

To generate this configuration each of the A-Cells is placed in a line such that the

distance between any pair of A-Cells is equal to the equilibrium distance[3]. In addition, each A-Cell is positioned with a similar initial orientation.

In each set of experiments three different aggregate sizes are investigated. The sizes that were selected are 16, 36 and 64. These values provide good coverage of the range of aggregate sizes which can be simulated in a reasonable period of time. Also, each of these values has an integer square root to allow similar aggregate sizes to be used for simulations where the initial configuration is a grid.

To save resources the 16 and 64 A-Cell aggregates are investigated over less detailed ranges. However, their inclusion should still indicate if any trends are specific to a particular aggregate size.

3.1.3 Base Parameter Set

As it is not possible to carry out an exhaustive exploration of the model's complete parameter space, it is necessary to choose some sensible base parameter set from which small sets of parameters can be investigated at a time.

The values that have been selected are a combination of estimates from the HYDRON prototype design and reasonable values which have been determined by a 'pepper pot' exploration of the parameter space.

- Simulation Time = 20000s

  The measures that will be used to evaluate the A-Cell aggregates assume that the aggregate has achieved its maximum compactness by the end of the simulation (see section 3.1.1.2). Therefore, to ensure that this is true even for poor parameter sets, a relatively long simulation time has been selected. This has the added benefit of indicating which parameter sets produce the most stable aggregates (i.e. aggregates which remain intact for this whole time period).

- Simulator Time Interval = 0.1s

- A-Cell Mass = 1kg

[3]The point where adhesive and repulsive forces are balanced (see section 2.2.3).

The mass of the HYDRONs was not known at the time the following experiments were carried out. Therefore, this parameter was set to a default value of $1kg$. Thus, all forces which act on the A-Cells are simply converted to accelerations.

- A-Cell Physical Hull Radius = $0.05m$
- Maximum Signal Range = $1.1m$
- Membrane Size = $0.5m$
- Equilibrium Distance = $0.55m$
- Number of Sensors = 4

  As the A-Cells are limited to movement in only two dimensions it is sufficient to evenly distribute the sensors around the equator of the A-Cell hull. Four sensors provide near complete coverage around the entire hull.

- Number of Adhesion Sites = 64

  Again, as these simulations are limited to two-dimensional movements the adhesion sites are simply evenly distributed around the equator of each A-Cell's membrane. Therefore, each adhesion site represents a small segment of the membrane.

- Membrane Damping = $0.4Kgs^{-1}$
- Membrane Repulsion = $2Kgm^{-1}s^{-2}$
- Attraction at Equilibrium Distance = 1.525

  This specific value is derived from the other parameters.

- Detector and Transmitter Field of View = $\frac{\pi}{2}$
- Bonding Equation Scaling Constant (see equation 2.6) = 1
- Maximum Random Force = $0N$

Number of A-Cells	16	36	64
Mean Time Per Simulation	1h 45m	5h 30m	17h 30m

Figure 3.7: *Mean time per simulation for the different aggregate sizes.*

- Maximum Random Torque = $0.001N$

  This low torque is sufficient to break the symmetry of sensor alignments in the A-Cell aggregates. Without this the model's performance would be greatly affected in all the simulations (see section 3.3).

- Maximum A-Cell Force = $0.01N$

  This parameter limits the maximum force that any of the A-Cells is able to produce. The value shown was determined in the initial aggregate stability experiments (see section 3.2).

3.1.4 Frequency Distribution and Sample Size

The A-Cell simulator is a very processor intensive application (see table 3.7). Therefore, as the resources available for this work are limited, there is a trade off between the sample size (the number of runs per experiment) and the number of experiments that may be carried out. In this case to allow adequate exploration of the model we have limited the sample size to 15.

To determine how best to estimate the variance of the sample it is necessary to examine the distribution of the value being calculated. To this effect, 100 simulations were carried out for each of the 16, 36 and 64 A-Cell aggregates. Each simulation started with the base parameter set and straight line initial configuration outlined previously in this section. The CC and CR were then calculated for each simulation.

The resulting histograms (see fig. 3.8) revealed that for each measure the distribution is slightly negatively skewed and in some cases appears to be multi-modal in nature.

As the distribution of the two measures is non-normal and the sample size is small, the best estimate of the variance of the sample mean can be achieved by *bootstrapping*

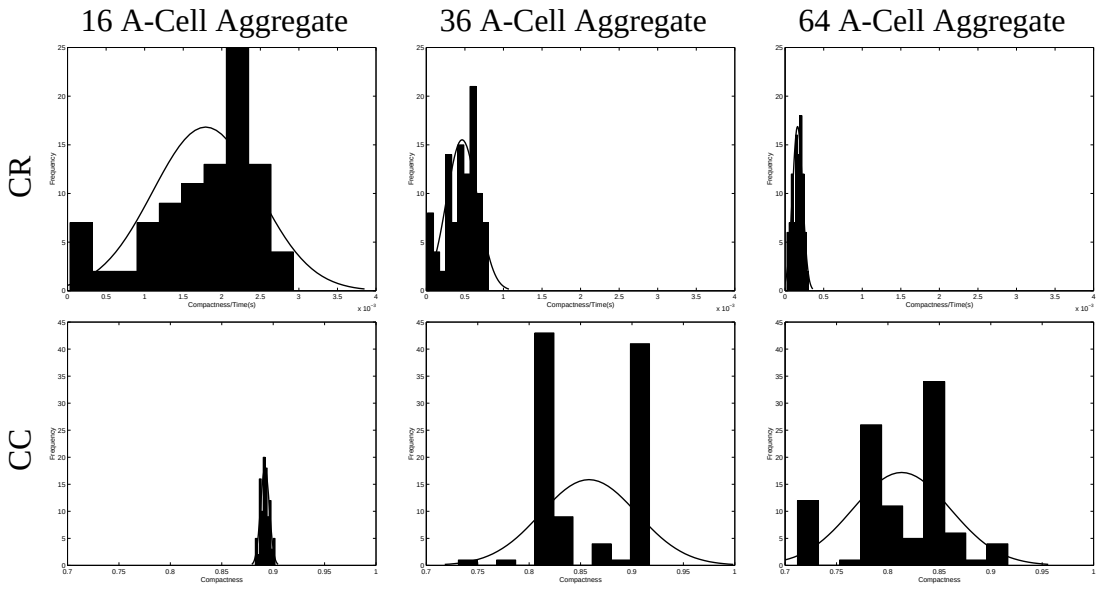


Figure 3.8: *Frequency distributions for both the* CR *measure and the* CC *measure. The best fit normal curve has been superimposed on each plot.*

(see Efron and Gong [14]).

Bootstrapping is a method of empirically assessing the uncertainty of a small sample without making any assumptions about its distribution. The process works by applying some statistic, in this case the mean, to a large number of pseudo-replicate samples created from the original data set. These subsamples are generated by randomly resampling the initial data, with replacement, to produce new samples the same size as the original. Therefore, each original piece of data may appear in a new subsample multiple times or not at all.

The values calculated for the statistic form an estimate of its sampling distribution and the standard deviation of this distribution then provides an approximation of the real standard error.

To calculate the standard error of samples used in this chapter, each sample is resampled 5000 times[4]. This standard error estimate is then used to determine the confidence error which provides the range of the error bars.

[4]This number should be large enough to produce stable values.

3.2 Aggregate Stability

It has been observed from initial experiments that if a number of A-Cells experience a force in a similar direction, the whole aggregate will begin to rotate. However, instead of stabilising, occasionally this rotational movement accelerates until the aggregate pulls itself apart.

Once the aggregate begins to rotate, the A-Cells nudge each other around in the same direction as they attempt to maintain an equilibrium distance apart. A build up develops when the forces generated by the A-Cells accumulate faster than they are cancelled out by the resistance of the fluid. The result is that the rotation becomes self-perpetuating and in fact self-accelerating until the forces involved reach the point where they begin to break the adhesive bonds, causing the aggregate to split.

There are three possible solutions which will prevent this build up of forces:

1. The resistance of the fluid can be increased.

2. The magnitude of the forces generated by the A-Cells can be sufficiently limited.

3. The A-Cells can coast for a period of time producing no movement force.

Clearly, increasing the resistance of the fluid assumes a level of control over the environment. As this may not always be the case, this is not considered. Of the remaining options, the latter coasting approach has the additional advantage of conserving the A-Cells' energy. However, it should also be noted that it would be a further abstraction from the real biology where the forces between cells are continuous.

In the experiments which were carried out, the rotational stability of the aggregate is measured by calculating each A-Cell's speed of rotation around the aggregate's centre of mass [5]. The mean of these values gives an estimate of the speed of rotation of the entire aggregate, which can then be used to estimate the mean rotational speed of the aggregate for the entire simulation.

Figure 3.9 shows the combined results of both increasing the coasting period and placing a maximum limit on the forces produced by the A-Cells. In each simulation,

[5] The magnitude of the rotational velocity.

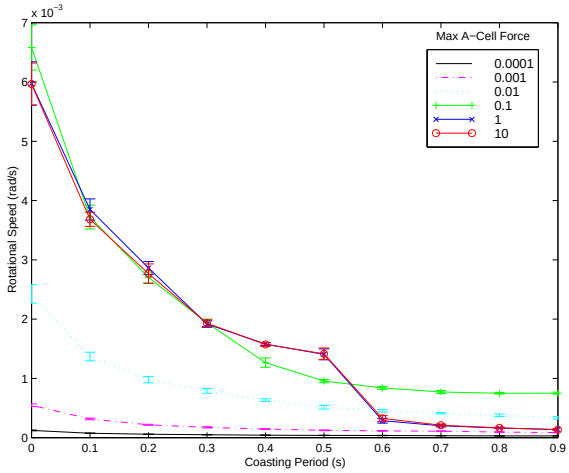
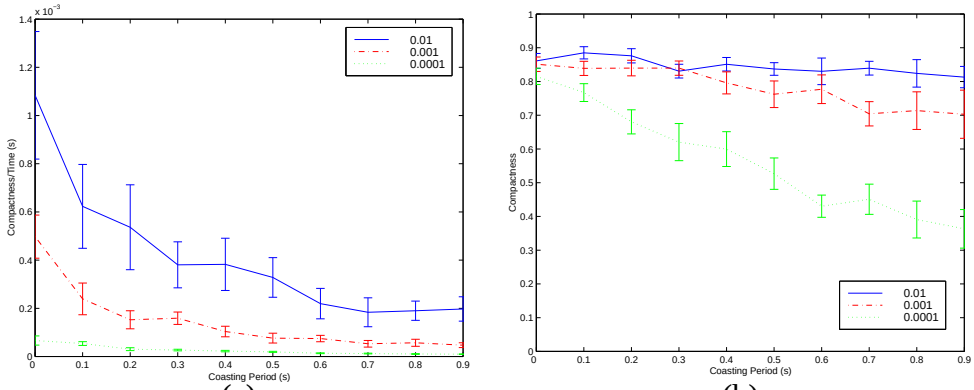


Figure 3.9: *Mean rotational speed of A-Cell aggregates for both when the maximum A-Cell force is limited and the coasting period is increased.*

the base parameter set was used and the A-Cells were also exposed to small random movement forces ($0.0001N$).

These results show that maximum force limits which exceed $0.1N$ result in almost identical behaviour, suggesting that these values are too high to impose any sufficiently large limiting effect on the forces which are being produced. There is a notable difference for these results when the coasting period increases to $0.6s$. However, the sudden drop in rotational speed which is exhibited by the two highest force limits is simply an artifact of the aggregates failing to remain intact. This rather dramatic effect is a result of the very slightly higher forces and a longer coasting period allowing the A-Cells to drift apart. The remainder of the results show the expected behaviour with both approaches causing a noticeable reduction in the rotational speed. However, figure 3.10 demonstrates that as either value becomes too extreme, the performance of the model deteriorates with both the CC and CR values dropping.

From these results a maximum force limit of 0.01N and 0s coasting period were selected for the base parameters of the following experiments. These values provide a suitably large decrease in mean rotational speed without adversely affecting the performance of the model. In addition, the lack of coasting period provides a better match to the observed biology.



(a) (b)

Figure 3.10: *The* CR *and* CC *plot for A-Cell aggregates for both when the maximum A-Cell force is limited and the delay between forces is increased. (a)* CR *(b)* CC

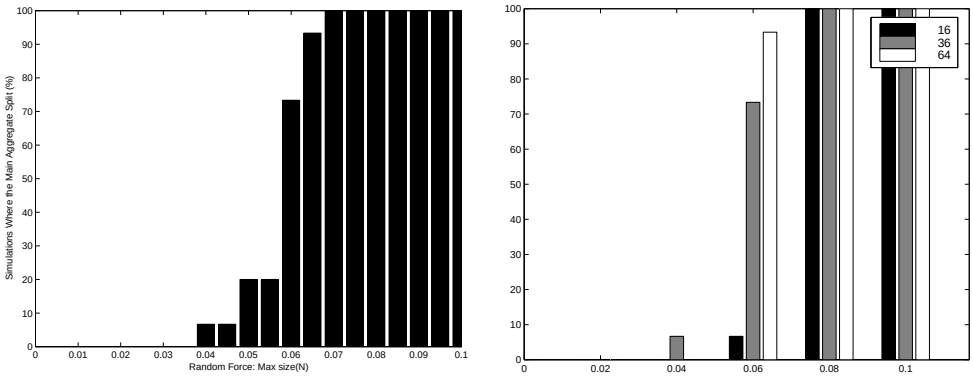
3.3 Random Forces

The purpose of this section is to investigate how robust the rounding behaviour exhibited by our model is to small random movements and torques. Such forces could either be the result of environmental perturbations or, in the case of the random movements, could be generated by the A-Cells themselves (see section 2.2.7). These results are summarised in Ottery and Hallam [44].

The first two sets of experiments investigate the effects of random movements and random torques independently of each other. Therefore, in each case, when one parameter is varied, the other is set to zero.

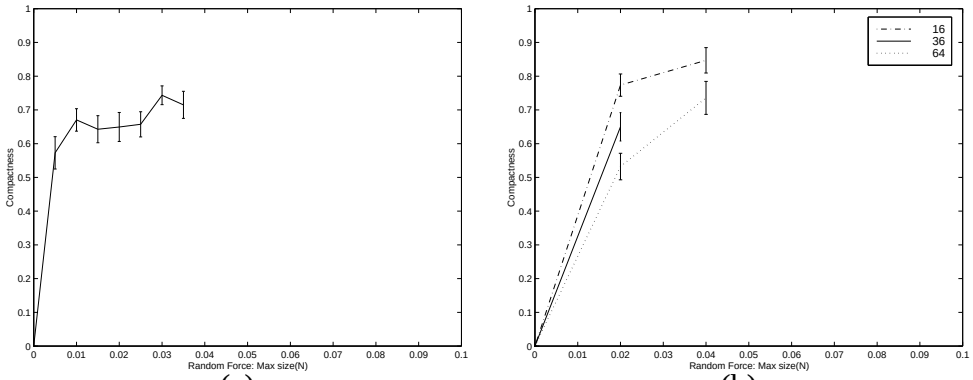
The random movements are generated by adding a small force in a random direction to each A-Cell, such that the magnitude of the force is a random value between zero and some predefined maximum limit. It is this maximum limit that is varied in these experiments. Figure 3.11 shows the histograms of the number of simulations where the main A-Cell aggregate did not remain intact[6]. There is nothing surprising about these results; they simply indicate that as the magnitude of the random movements increases, so does the chance that this random force will outweigh the attractive forces between the A-Cells, causing the aggregate to split.

[6] At least one A-Cell lost communication with the others.



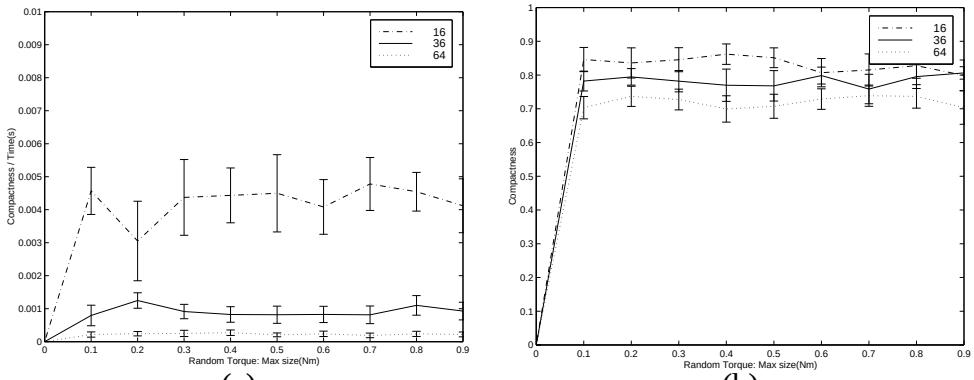
(a) (b)

Figure 3.11: *Histograms showing the number of aggregate that did not remain intact over a range of random movement force limits. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*



(a) (b)

Figure 3.12: CC *values of A-Cell aggregates over a range of random movement force limits. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*



(a) (b)

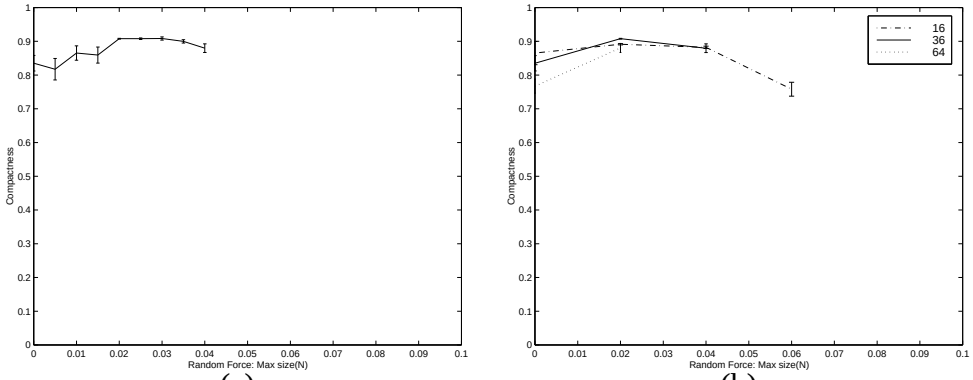
Figure 3.13: CR *and* CC *values for A-Cell aggregates over a range of random torque limits. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a)* CR *(b)* CC

Figure 3.12 shows the CC plots of the results. The more detailed plot for the 36 A-Cell aggregate shows a slight increase in the CC measure as the magnitude of the random movement forces increases, a result which is also reflected in the combined plot. This is expected, as larger random movements should increase the exploration performed by the A-Cells, thus increasing the chance that they will encounter A-Cells they can bond with more strongly.

The second of this first set of experiments investigates the effect of random torques. In this case, a small torque was applied to each A-Cell, generating a spin around the A-Cells' vertical axes[7]. The direction of the spin is selected at random and the magnitude is a random value between zero and some predefined maximum limit. Again, it is this maximum limit that is varied in this set of experiments.

Unlike the forces present in previous experiments, random torques do not directly cause the A-Cells to make random movements. However, as the A-Cells rotate, their sensors become misaligned, making the aggregate less rigid. Figure 3.13 shows the combined plot for both the CR measure and the CC measure. These graphs show that the level of torque has little effect on the performance of the model and that in fact, regardless of the level of torque, the aggregate remains intact. However, by adding

[7]This is the only axis around which the A-Cell can freely rotate (see section 2.2.1).



(a) (b)

Figure 3.14: CC *values of A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random movement force limits. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

some torque the more flexible aggregates have achieved a significant increase in the compactness index that is obtained.

The second set of experiments investigate the combined effect of these parameters. As the previous results show that the level of torque is not significant, in this set it remains fixed whilst the movement force is varied. The torque was set to a low value as this is likely to be more representative of the slight torques which would be caused by any environmental perturbations. However, the choice of the specific value of $0.01Nm$ is relatively arbitrary.

The combination of these two parameters produced the best results. Figure 3.14 shows that each aggregate size achieved similarly high CC compactness measures of just below 0.9. It is also clear from the error bars that they achieved this high compactness far more consistently than in the previous set of experiments. In addition, the more detailed plot for the 36 A-Cell aggregate also shows an initial slight improvement as the magnitude of the random movement forces increases. This slight increase in performance is also present in the CR plot (see fig. 3.15).

Overall, it appears that by applying the random movement forces to a less rigid aggregate, the rounding process can be significantly improved. There is also an indi-

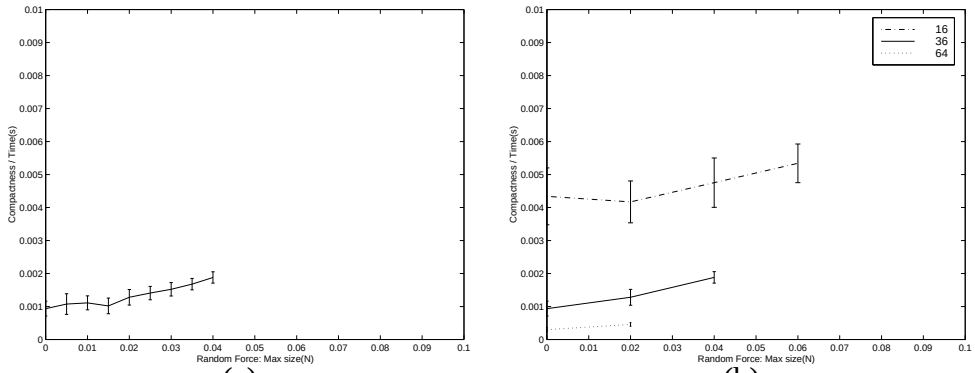


Figure 3.15: CR *values of A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random movement force limits. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

cation that larger magnitudes of random movement produce better results for both the CR and CC up to the point where the aggregate begins to split.

The exception to this is the 16 A-Cell aggregate which, although intact for the highest magnitude of random movement force, shows a sudden drop in the CC measure (see fig. 3.14). The most likely hypothesis that explains this is that the magnitude of the random movements becomes so large that the aggregate cannot maintain any compact configuration. Figure 3.16 supports this interpretation by demonstrating a clear link between the size of the random movement force limit and the level of compactness variance that the aggregates exhibit.

In summary, these results show that the model is reasonably robust to both types of random force. More specifically, they indicate that the presence of each type improves the overall performance of the model. In particular, higher magnitudes of random movement increase performance up to the point were the random movements outweigh the attractive forces between the A-Cells, causing the aggregate to become unstable and possibly break apart.

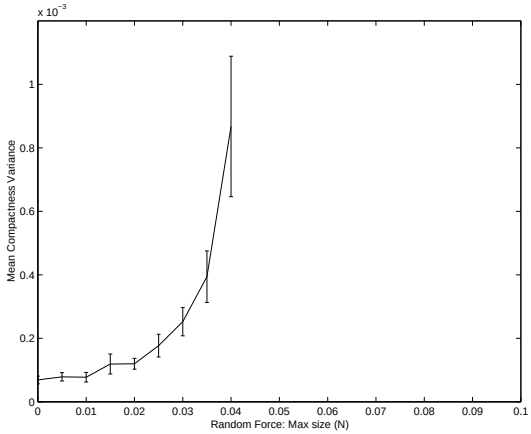


Figure 3.16: *Compactness variance for a 36 A-Cell aggregate over a range of random movement force limits.*

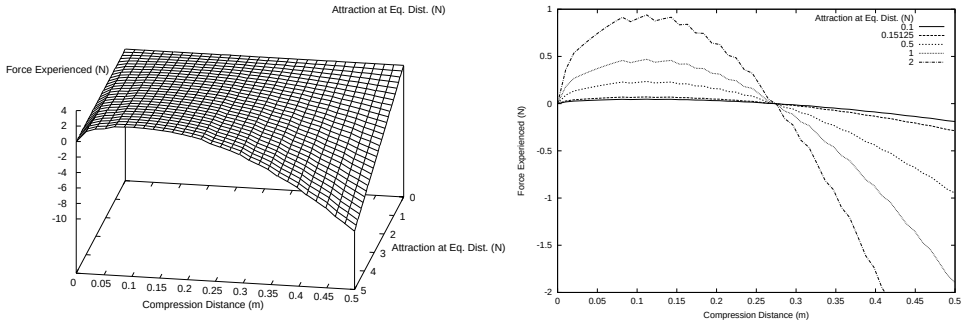
3.4 Membrane Properties

The following sections investigate how the performance of our model relates to the choice of specific membrane properties.

3.4.1 Adhesion Force

The experiments presented in this section investigate how the model's performance is affected by the level of the adhesion and thus attraction between the A-Cells. As the attraction between the A-Cells increases, the equilibrium distance will be reduced. Therefore, to control for this, the repulsion factor of the A-Cells' membranes (see equation 2.5) is adjusted such that the equilibrium distance will remain constant. Additionally, the attraction the A-Cells experience for any amount of membrane compression depends on many parameters. Therefore, for clarity, the attraction values presented here are given in terms of the actual attraction experienced when two A-Cells are the equilibrium distance apart.

The actual force that an A-Cell generates is dependent on both the attractive and repulsive forces it experiences. Figure 3.17 shows how the resulting generated force changes as two A-Cells move increasingly close together. It is clear from the two dimensional plot (see fig. 3.17 - b) that the magnitude of the forces generated by the

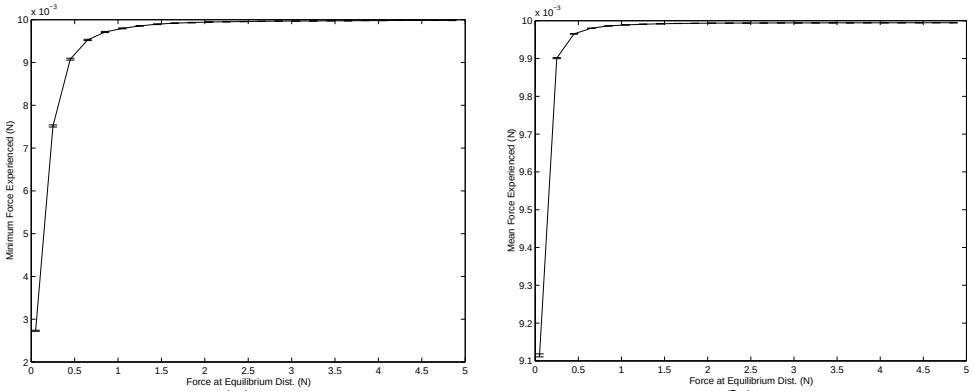


(a) (b)

Figure 3.17: *Calculation of the force experienced by an A-Cell adhering to a single other A-Cell as both the force at the equilibrium distance and its membrane compression distance are increased.*

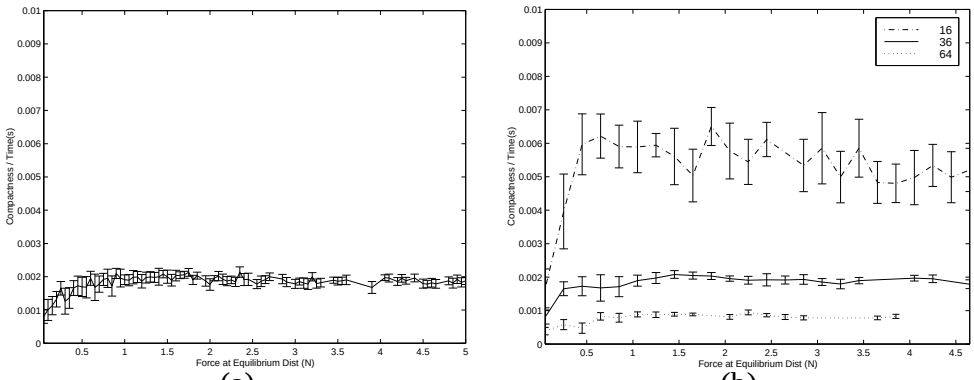
A-Cells, for any amount of membrane compression, increases as the force of attraction at the equilibrium distance rises. Also, even small forces of attraction at the equilibrium distance result in generated forces which are greater then the maximum force limit of $0.01N$ (see section 3.2) for the majority of situations. However, any single A-Cell in an aggregate will bond with many others and thus experience forces in many directions which will, to some extent, cancel each other out. Additionally, once the aggregate reaches a reasonably stable configuration with the A-Cells in some equilibrium state, the forces being generated will also be reduced.

Figure 3.18 shows how the forces generated by the A-Cells in the actual simulations change as the force at the equilibrium distance rises. These results show that the forces being generated rapidly converge toward the maximum limit. This trend is reflected in the CC and CR graphs of the same simulations (see fig. 3.19 & 3.20) which show a rapid initial change, after which the values stabilise. However, the CC graph then shows a gentle decrease as the force at the equilibrium distance continues to rise, and some small percentage of the aggregates also fail to remain intact (see fig. 3.22). The compactness variance graph (see fig. 3.21) provides an explanation for this behaviour by showing a steady increase in the instability of the aggregates as the forces generated by the A-Cells become increasingly limited.



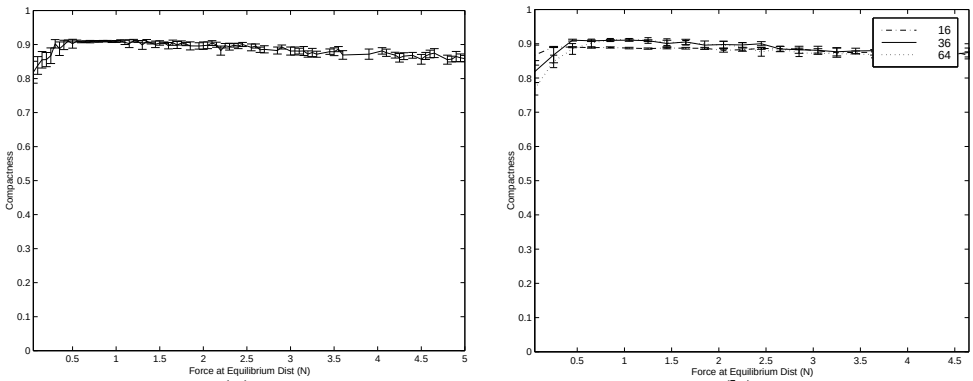
(a) (b)

Figure 3.18: *The mean forces generated by the A-Cells of a 36 A-Cell aggregate as the force at the equilibrium distance is increased. (a) Mean of the minimum force generated at each cycle in the simulation. (b) Mean of all the forces generated in the simulation.*



(a) (b)

Figure 3.19: CR *values of A-Cell aggregates over a range of forces at equilibrium values. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*



(a) (b)

Figure 3.20: CC *values of A-Cell aggregates over a range of force at equilibrium distance values. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

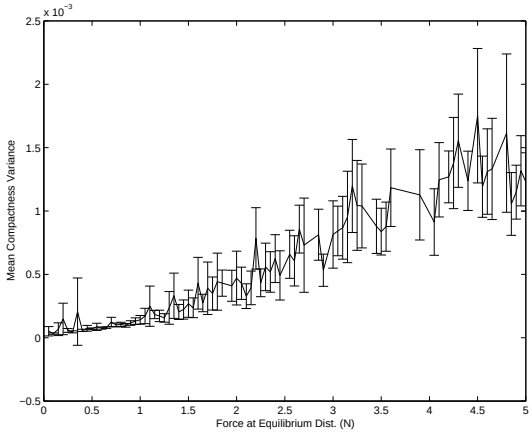


Figure 3.21: *Compactness variance for a 36 A-Cell aggregate over a range of 'force at equilibrium distance' values.*

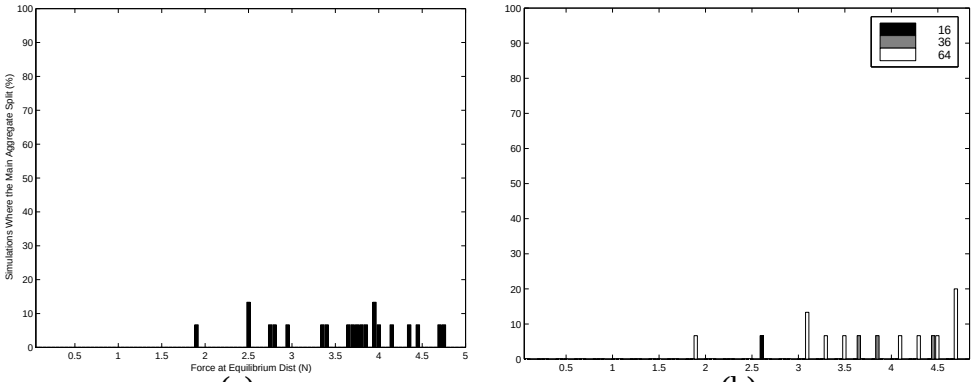


Figure 3.22: *Histograms showing the number of aggregate that did not remain intact over a range of 'force at equilibrium distance' values. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

3.4.2 Number of Adhesion Sites

When two membranes are in contact, the number of A-CAMs that may form bonds will be some multiple of the total number of A-CAMs divided by the total number of adhesion sites (see section 2.2.5). Therefore, the number of adhesion sites determine to what level of detail the number of A-CAMs, and consequently the number of bonds, can be calculated.

In this set of experiments, aggregates of A-Cells with numbers of adhesion sites ranging from eight to one hundred and sixty were simulated. The results show that there is no noticeable effect on the CR . However, although this is also largely true for the CC values, there is a slight peak in performance for all the aggregate sizes when only eight adhesion sites are used (see fig. 3.23). To investigate this low end of the range in more detail, a further series of simulations were carried out with only the 36 A-Cell aggregate (see fig. 3.24). The results of these simulations show that a minimum of five adhesion sites are required to prevent A-Cells from breaking away from the aggregate. Also, although there is still no effect on the rounding rate, the results show a rise in CC as the number of adhesion sites increases to eight. At this point the CC plateaus before dropping slightly when the number of adhesion sites increases to twelve. The CC then appears to remain relatively constant, with the exception of an

isolated peak when fourteen adhesion sites are used.

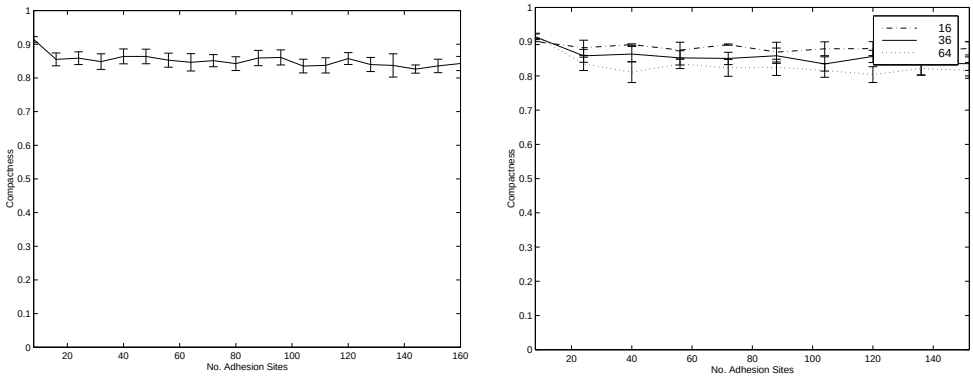
Calculating the compactness variance of the latter set of simulations reveals that as the number of adhesion sites drops below approximately ten, the compactness variance increases rapidly (see fig. 3.25). When the number of adhesion sites is higher than this, the compactness variance, like the CC , remains relatively constant. More specifically, comparing these results with the CC values suggests that there is a window of compactness variance between 1×10^{-4} and 2×10^{-4} where the best performance is achieved.

This change in compactness variance can be explained by considering how the number of adhesion sites on an adhering area of membrane can change as an A-Cell rotates. For example, figure 3.26 shows that in one orientation only a single adhesion site will be considered to be on the adhering area of membrane. However, if the A-Cell rotates, this can increase to two. The figures shown in 3.27 were generated by determining this maximum and minimum number of adhesion sites and calculating the percentage of the total membrane that they represent. The graphs were calculated for an A-Cell which is $0.6943m$ from a single neighbour it is adhering to. This is the mean distance between neighbouring A-Cells calculated from a sample of 15 simulations of a 36 A-Cell aggregate with the base configuration[8]. The graphs show that the two values differ most for small numbers of adhesion sites, with the difference dropping very quickly at first, then asymptotically approaching zero as the number of sites continues to increase. These results match the corresponding drop in the compactness variance, indicating that this fluctuation in the number of A-CAMs which are available to bond causes the resulting fluctuations in the A-Cells' positions, and thus, the aggregate's shape.

3.4.3 Equilibrium Distance

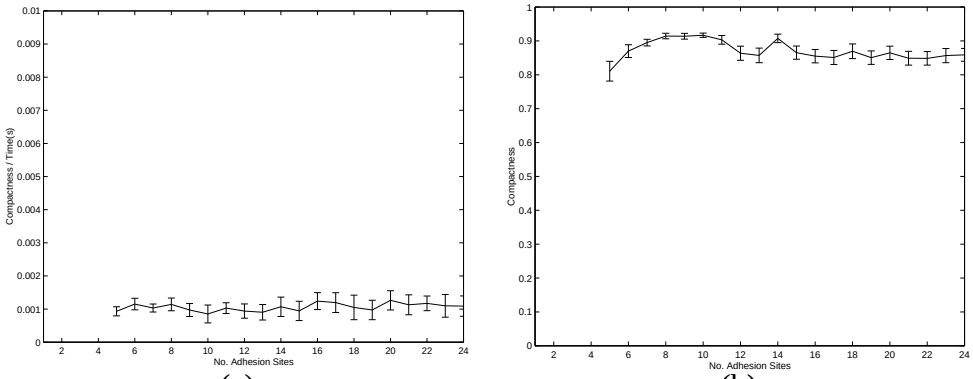
The equilibrium distance is the distance at which the forces of attraction and repulsion generated by two adhering A-Cells are balanced. In this set of experiments this distance is varied by adjusting both the membrane repulsion multiple (see equation 2.5)

[8]This value is greater than the calculated equilibrium distance which is estimated from only two A-Cells as the A-Cells in an aggregate have many neighbours. Therefore, some areas of an A-Cells membrane will be shared.



(a) (b)

Figure 3.23: CC *values of A-Cell aggregates with a range of adhesion site numbers. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*



(a) (b)

Figure 3.24: CR *and* CC *values for a 36 A-Cell aggregate over a range of adhesion site numbers. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a)* CR *(b)* CC

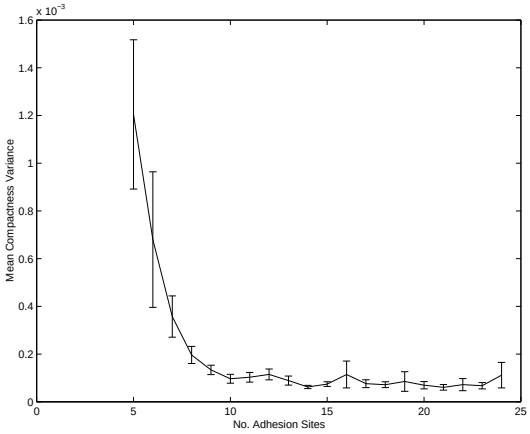


Figure 3.25: *Compactness variance for a 36 A-Cell aggregate over a range of adhesion site numbers.*

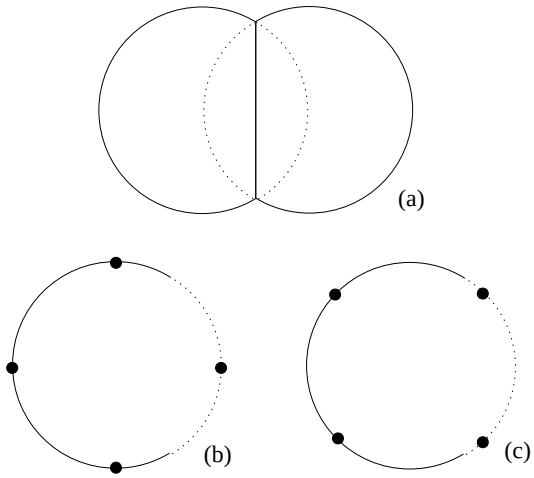
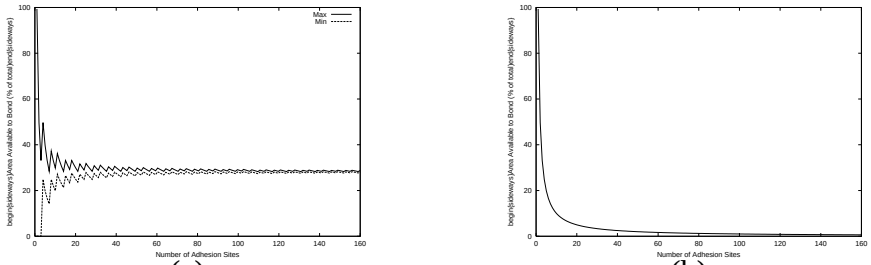


Figure 3.26: *Orientation effects on number of adhesion sites in contact.*



(a) (b)

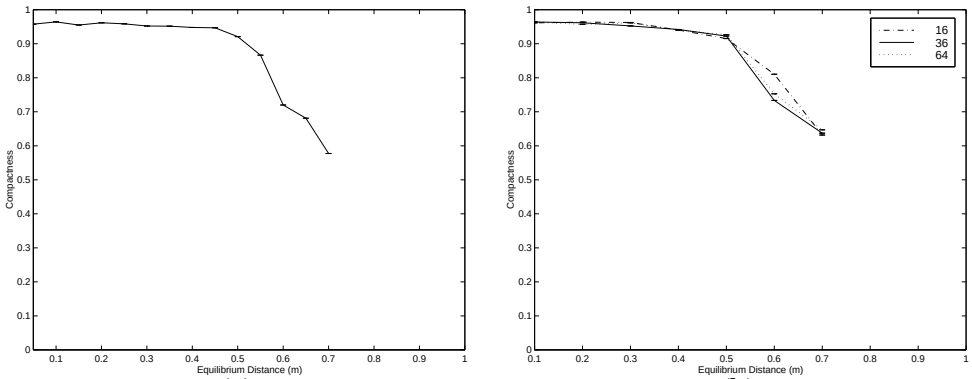
Figure 3.27: *Values calculated for A-Cells with a range of numbers of adhesion sites. (a) The maximum and minimum percentage of the total bonding area that is available to bond, calculated from the maximum and minimum number of adhesion sites on the area of contact between two A-Cells. (b) The difference between the maximum and minimum percentages of the total area available to bond.*

and the number of A-CAMs produced by the A-Cells. In this manner, it is possible to change the equilibrium distance while keeping the adhesion at this distance constant.

The equilibrium distances that were considered cover the entire possible range from the point where the A-Cells are just in contact to the point where their membranes almost fully overlap. The results from these simulations (see fig. 3.28 & 3.29) show that very high CR and CC values are achieved when the equilibrium distance is small. However, as the equilibrium distance increases, both these values drop and once the equilibrium distance exceeds $0.7m$, the aggregates no longer remain intact.

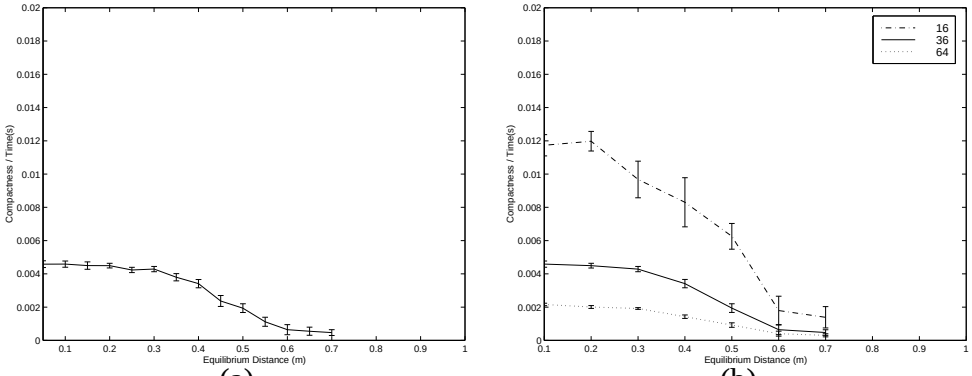
It is also clear from the results that a significant change in the model's performance occurs when the equilibrium distance drops below the virtual radius of the A-Cells (e.g. hull radius plus membrane size) which is, in these simulations, $0.55m$. When the equilibrium distance is below this threshold, both the CC and rounding rate values that are achieved are higher than those obtained during all the other experiments presented in this chapter. However, as soon as the equilibrium distance exceeds this threshold, the CC drops very rapidly.

Both these features are linked with the probability of A-Cells in an aggregate discovering new neighbours to bond with. If the A-Cells have a high equilibrium distance, they will be much further apart. Therefore, this probability is low, and the aggregate



(a) (b)

Figure 3.28: CC *values of A-Cell aggregates with a range of equilibrium distances. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*



(a) (b)

Figure 3.29: CR *values of A-Cell aggregates with a range of equilibrium distances. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

which is formed will be much more rigid. However, as the equilibrium distance is decreased this probability will rise, making the aggregate much more fluid and allowing it to break out of locally optimal configurations. The dramatic change in performance occurs when the equilibrium distance becomes small enough for the A-Cells to form bonds through another A-Cell's membrane, effectively allowing them to see past their immediate neighbours. Although this causes a significant improvement, it does deviate further from the biological model in which the membrane forms a physical barrier.

3.5 Hardware Specificity

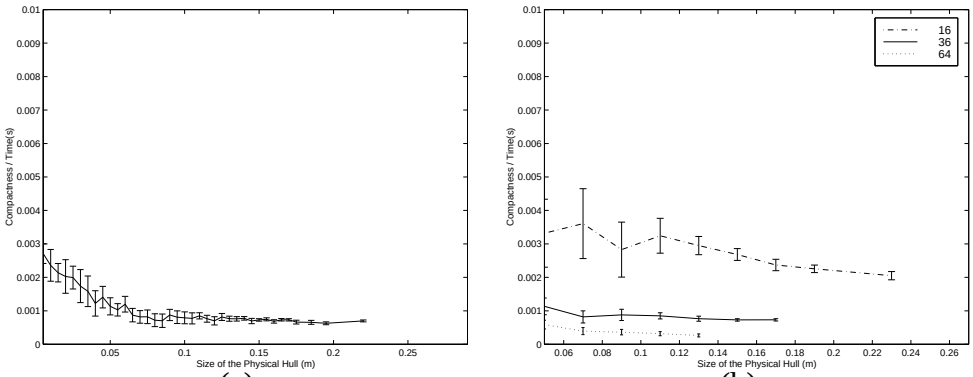
The following sections investigate how the performance of our model relates to the specification of the different hardware components.

3.5.1 Hull Size

The virtual nature of the A-Cell membrane means that it is possible to change the size of the physical hull while keeping the virtual size of the A-Cell constant. Therefore, such a change should have no direct impact on the performance of the model. However, it is still possible for any change to have an indirect effect through the following:

1. The level of fluid drag acting on the A-Cells.
2. The probability of physical collisions between the A-Cells' hulls.
3. The level of error in the distance and direction estimates.

To investigate these more subtle effects, simulations were carried out covering a broad range of possible hull sizes. As discussed above, the membrane size of the A-Cells in these simulations was adjusted to maintain a constant virtual size. The results of these simulations show several clear features. Firstly, the CR results (see fig. 3.30) show an initial drop as the hull size increases, most predominantly in the more detailed graph for the 36 A-Cell aggregate. This is most likely to be caused by the increase in drag which results from the larger hull sizes. Secondly, the CC results (see fig. 3.31) show a slight improvement as the hull's radius becomes larger. However, once



(a) (b)

Figure 3.30: CR *values of A-Cell aggregates over a range of hull sizes. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

the hull's radius reaches approximately $0.13m$ the CC drops and becomes increasingly erratic. The compactness variance (see fig. 3.32) suggests that this trend is caused by the aggregates becoming increasingly volatile. As in section 3.4.2, initially as the compactness variance increases, the CC value also increases. However, after a certain threshold is reached, the CC values drop as the aggregate becomes increasingly unstable. Figure 3.33 demonstrates that once the hull radius reaches approximately $0.2m$ this then results in a reasonably steady increase in the number of aggregates which fail to remain intact.

The increase in compactness variance will be partly due to an increase in the number of collisions between the A-Cell hulls as the hull size approaches half the equilibrium distance ($0.275m$). However, it is unlikely that this is solely responsible particularly when the hull radius is reasonably low.

To determine what part the error in the distance and direction estimates might be playing in this increasing compactness variance, an artificial two A-Cell model was created. In this simple model the position of the A-Cells was fixed while they were rotated anti-clockwise around their vertical axis. The error in the distance and direction estimates were then calculated for a large number of orientation pairs (see fig. 3.34). As the A-Cells are rotationally symmetrical, only rotations of less than $\frac{\pi}{2}$ were con-

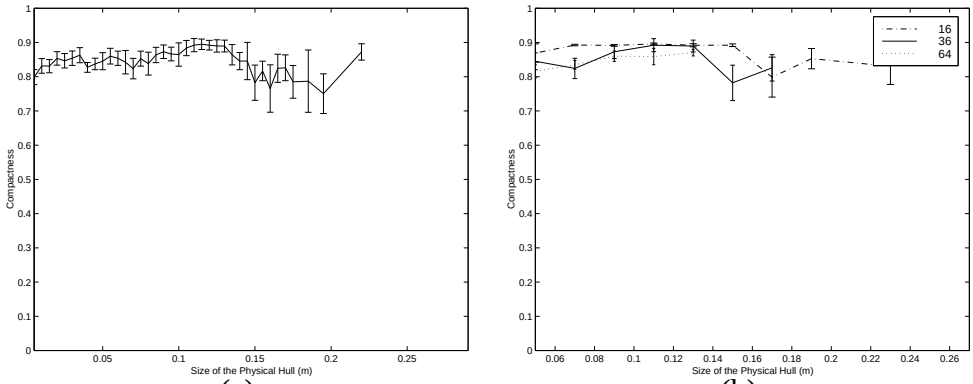


Figure 3.31: CC *values of A-Cell aggregates over a range of hull sizes. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

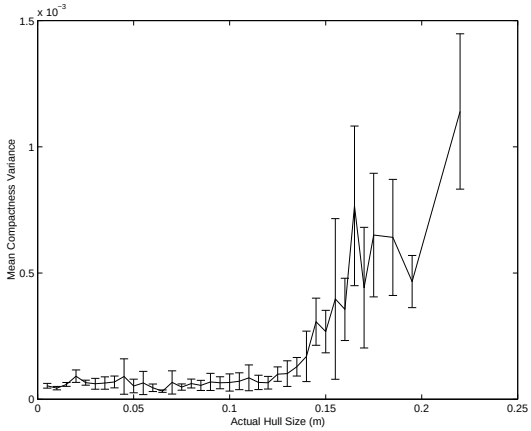


Figure 3.32: *Compactness variance for a 36 A-Cell aggregate over a range of hull sizes.*

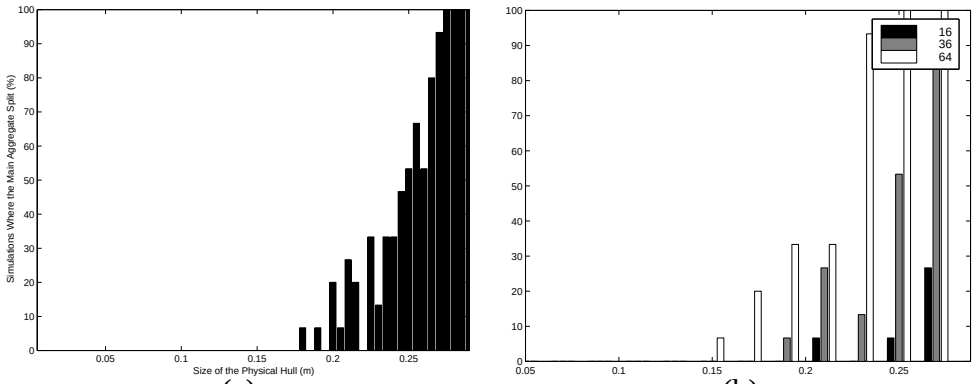


Figure 3.33: *Histograms showing the number of aggregate that did not remain intact over a range of hull sizes. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

sidered. Also, the distance between the two A-Cells in the model was set to $0.6943m$ which is more representative of the distance between neighbouring A-Cells in an actual aggregate[9].

The results obtained by running this simple model for the same range of A-Cell hull sizes as were used in the main simulations, show that the mean error in the direction estimate changes very little (see fig. 3.35). However, there is a clearly linear increase in both the mean and maximum distance error estimates. Figure 3.36 shows how this increase relates to the simple distance estimate algorithm. The estimate is calculated as the mean of all the distance estimates for the same signal, each of which consists of the estimate of how far the signal has travelled between the transmitter and detector plus the additional distance between the sensors and the A-Cell centres (r). In the worst case situation shown in figure 3.36, when the two A-Cells rotate very slightly toward each other, the distance estimate of signal path (b) will be approximately correct. However, signal path (a) will be out by approximately $2r$, where r is the radius of the A-Cell hull. Therefore, the average and thus the final distance estimate will be r greater than the actual distance.

[9]This distance is the mean distance between neighbouring A-Cells calculated from a sample of fifteen 36 A-Cell aggregates with the base configuration.

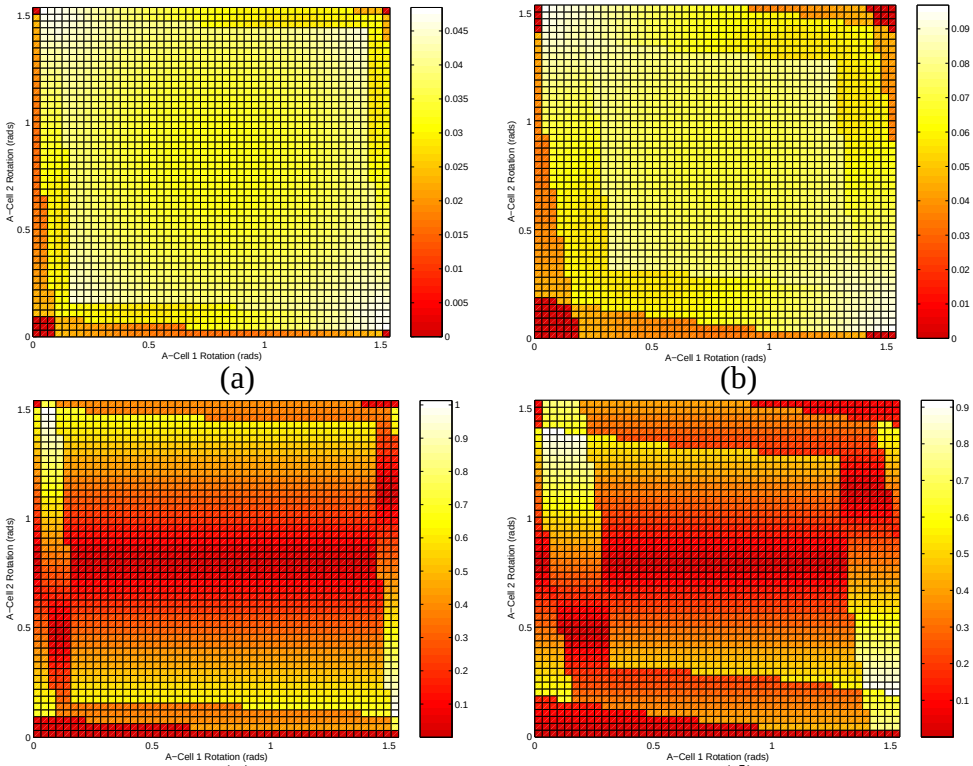


Figure 3.34: *Example estimate error grids calculated from a simple two A-Cell model. (a) Distance estimate error for a hull radius of* $0.05m$*. (b) Distance estimate error for a hull radius of* $0.1m$*. (c) Direction estimate error for a hull radius of* $0.05m$*. (d) Direction estimate error for a hull radius of* $0.1m$*.*

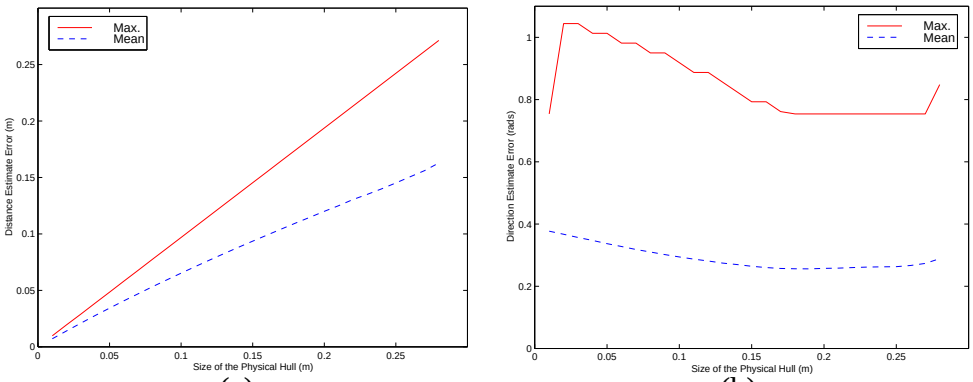


Figure 3.35: *Estimate errors calculated from a simple two A-Cell model over a range of hull sizes. (a) Mean and maximum error in the distance estimate. (b) Mean and maximum error in the direction estimate.*

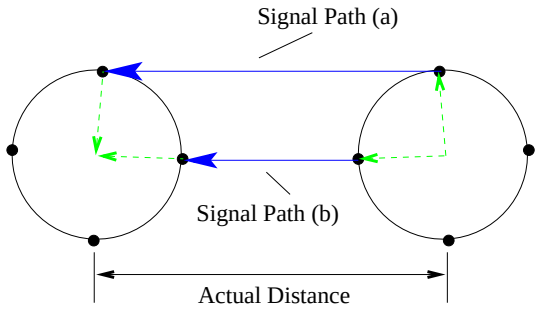
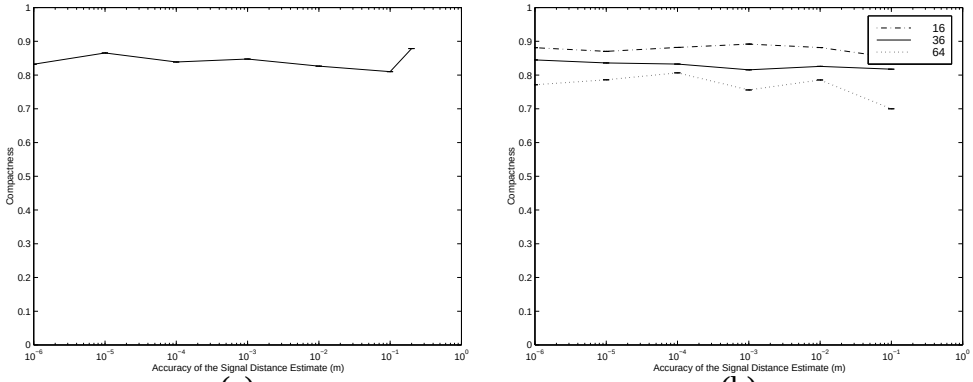


Figure 3.36: *Orientations causing the maximum error in the distance estimate.*



(a) (b)

Figure 3.37: CC *values of A-Cell aggregates with a range of signal distance estimate accuracies. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

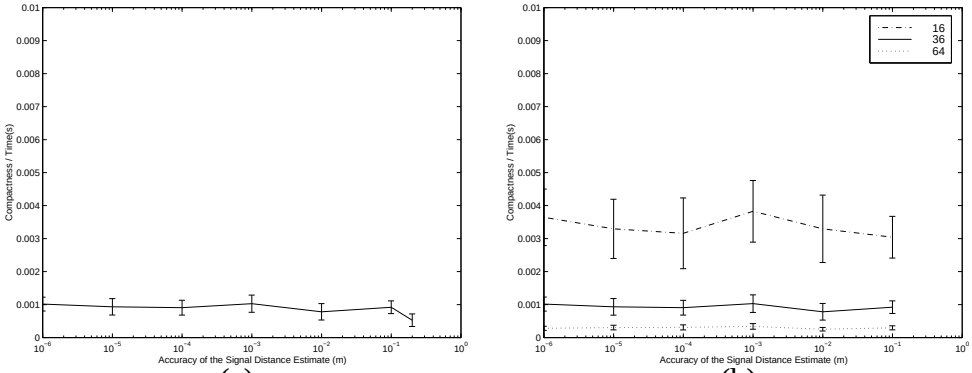
The most interesting discovery from these results is that some error in the distance estimate can actually be beneficial to the model as it naturally increases the fluctuations in the A-Cell's positions, causing them to explore more of their neighbouring space.

3.5.2 Signal Accuracy

In our implementation of the model, the simple but crude method used by the A-Cells to estimate the distance of their neighbours relies, to some extent, on how accurately the distance travelled by any signal can be calculated.

In this set of experiments, the signal distance estimate was rounded to the nearest multiple of a specified value. Thus, as the specified value is increased, the accuracy of the estimate is reduced. To cover a much broader range, the values considered were increased in powers of 10 from $10^{-6}m$ to $1m$. The 36 A-Cell aggregate was also evaluated for $0.2m$ multiples between $0.2m$ and $1m$ to provide additional detail.

The results shown in figures 3.37 and 3.38 indicate that the rounding behaviour of the model is tolerant of even very coarse distance estimates upto $0.2m$, after which the aggregates fail to remain intact. Also, the compactness variance graph of the same



(a) (b)

Figure 3.38: CR *values of A-Cell aggregates over a range of signal distance estimate accuracies. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 36 A-Cell aggregate (b) Less detailed range for 16, 36 and 64 A-Cell aggregates.*

results (see fig. 3.39) shows that when the signal estimate is $0.01m$ and below, the stability of the aggregates changes very little. These results match those given in section 3.5.1. As the mean distance estimate error of A-Cells with a hull radius of $0.05m$ was already calculated as just under $0.05m$, the accuracy of the signal distance estimate should have little additional effect as long as it is greater than this value.

The two interesting results are obtained for the lower accuracies of $0.1m$ and $0.2m$. Unexpectedly, the CC results show that rounding to the nearest $0.2m$ produced a slightly more compact aggregate. However, the compactness variance is also higher for this value and, in particular, it is in the same range that has produced the best results in previous experiments (see section 3.4.2). Therefore, it is likely that the extra fluctuations introduced by the error in the distance estimate have resulted in the better performance.

Rounding to the nearest $0.1m$ actually resulted in a higher mean compactness variance without affecting either the CC or the CR . However, the confidence error of the value is very large indicating that this value might not be representative of the actual population.

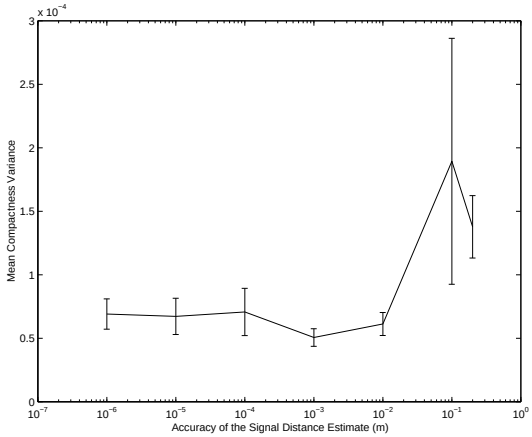


Figure 3.39: *Compactness variance for a 36 A-Cell aggregate over a range of signal distance estimate accuracies.*

3.5.3 Sensor Fields of View

In this chapter the A-Cells use a very basic sensor layout with only four transmitter-receiver pairs equally distributed around each A-Cell's hull. As discussed previously (see section 3.5.1), the orientation of the A-Cells and the distance between them can affect the accuracy of both their distance and direction estimates. In this section we investigate the additional property of sensor field of view (FOV).

In this set of experiments, a range of transmitter and detector FOVs are considered simultaneously to reveal any combined effects. Also, in addition to the main simulations, an additional two A-Cell model similar to that described in section 3.5.1 was evaluated to investigate the respective distance and direction estimate errors in more detail.

The histograms of the main simulations, showing the number of runs where the main A-Cell aggregate did not remain intact, are shown in figure 3.40. These graphs clearly indicate that there is a very distinct threshold, with aggregates of A-Cells with sensor FOVs under $\frac{\pi}{4}$ failing to remain intact in almost 100% of the simulations. Crucially, this angle is the point below which the sensor coverage will diverge with range, thus increasing the number of situations where any two A-Cells will be out of contact (see fig. 3.42). To illustrate this, the simple two A-Cell model was also used to calculate the percentage of orientation pairs where the A-Cells could not communicate (see

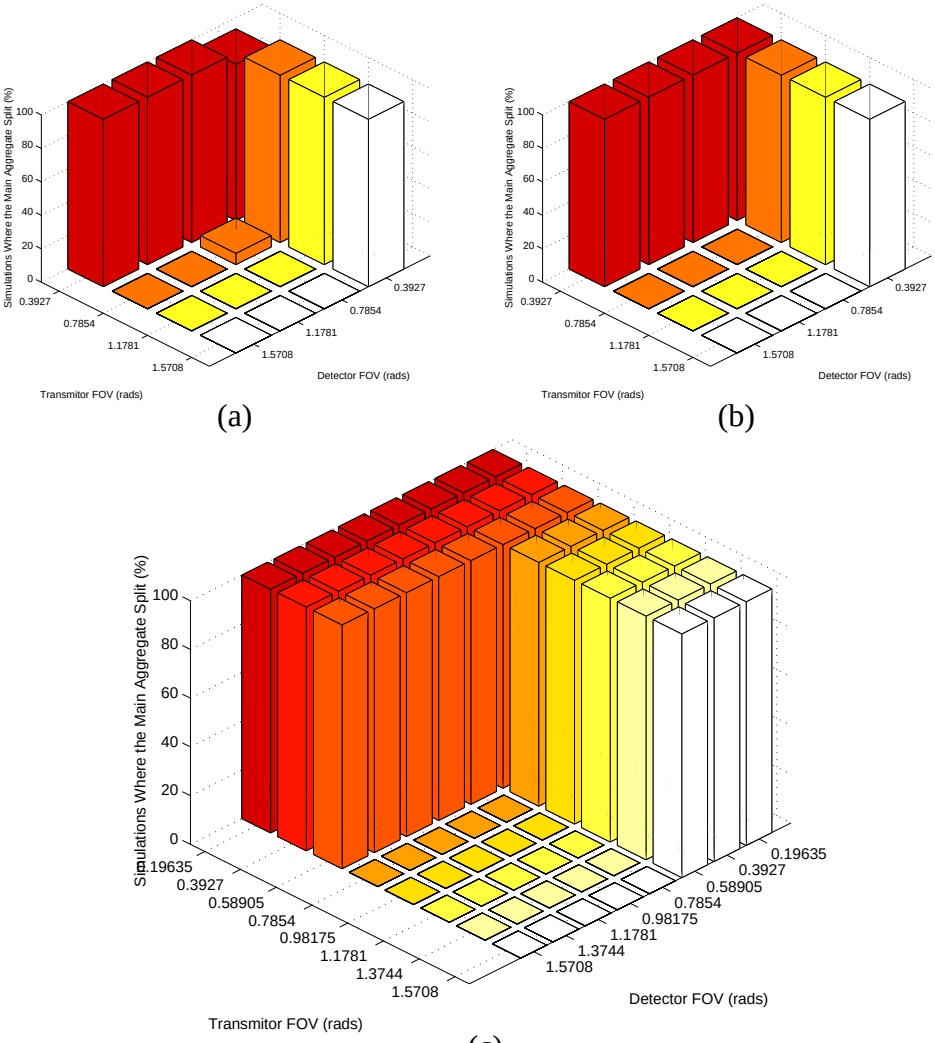


Figure 3.40: *Histograms showing the number of aggregate that did not remain intact. (a) 16 A-Cell aggregate (b) 64 A-Cell aggregate (c) 36 A-Cell aggregate*

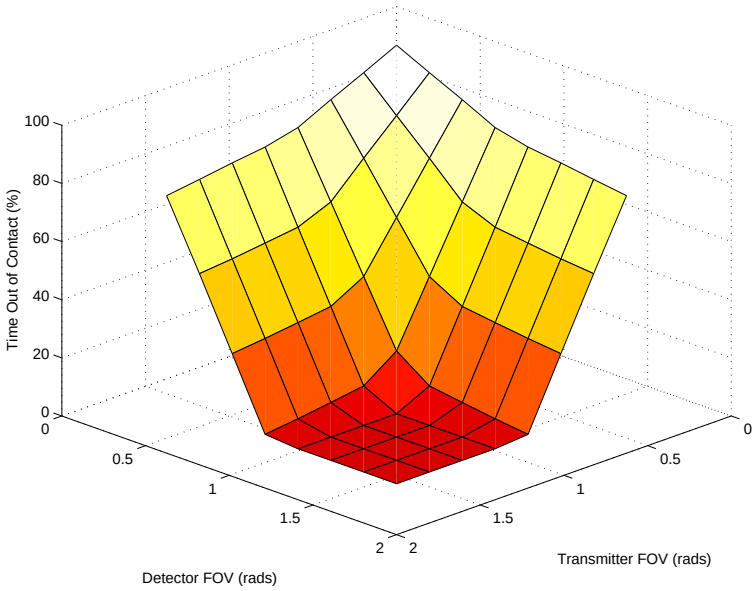


Figure 3.41: *The percentage of orientation pairs when two A-Cells cannot communicate for a range of detector and transmitter FOVs.*

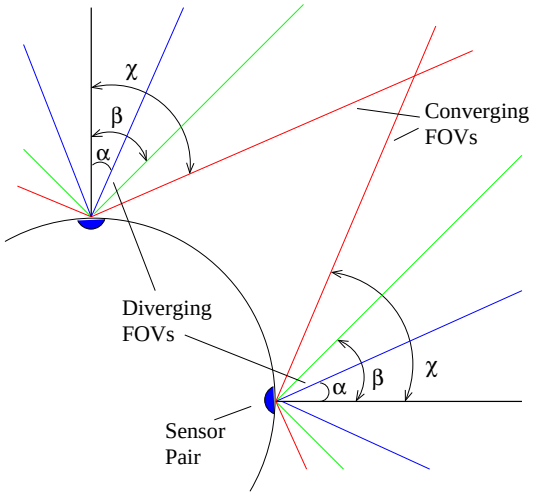


Figure 3.42: *Examples of sensor coverage:* $\alpha = \pi/8$, $\beta = \pi/4$, *and* $\gamma = 3\pi/8$.

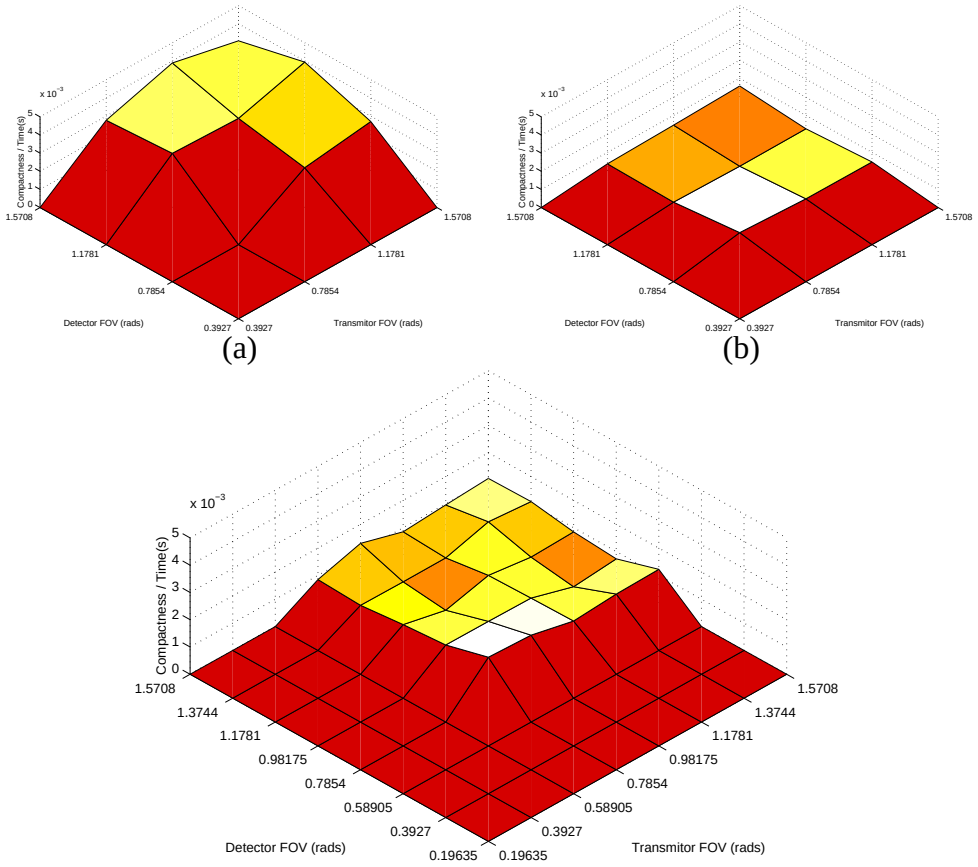
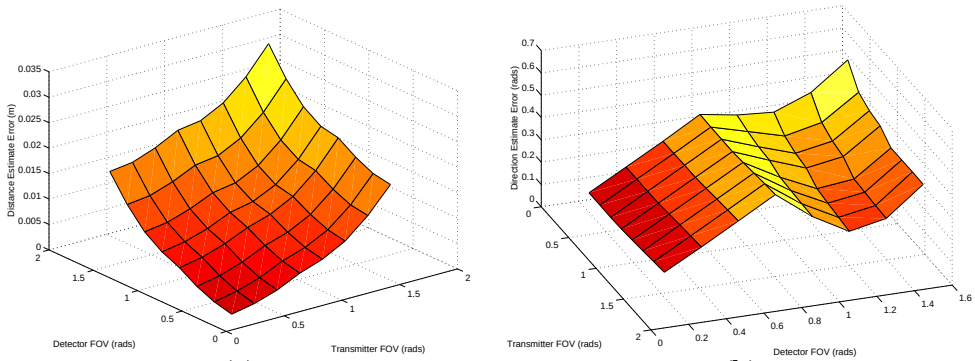


Figure 3.43: CR *values of A-Cell aggregates over a range of transmitter and detector FOVs. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 16 A-Cell aggregate (b) 64 A-Cell aggregate (c) 36 A-Cell aggregate*

fig. 3.41). As expected these results closely match the actual simulation results and demonstrate the same sudden change at $\frac{\pi}{4}$. More specifically these results also show a subtle increase in the time out of contact when both the transmitter and detector FOVs are equal to $\frac{\pi}{4}$, a feature which is also mirrored by an equivalent rise in the number of 16 A-Cell aggregates which failed to remain intact.

The CC and CR graphs for this set of experiments are shown in figure 3.43 and 3.46 respectively. One of the clearest trends in these results is the drop in CC values when the detector FOV is equal to $1.1781 rads$, regardless of the transmitter FOV (see fig. 3.47). A comparison of these results with the mean direction estimate errors (see fig. 3.44)



(a) (b)

Figure 3.44: *Estimate errors calculated from a simple two A-Cell model over a range of transmitter and detector FOVs (a) Mean error in the distance estimate. (b) Mean error in the direction estimate.*

calculated from the two A-Cell model shows that this feature actually corresponds to the lowest direction errors. Calculating the compactness variance results (see fig. 3.45) also reveals that this drop corresponds with a low level of fluctuations in the A-Cells' positions. In combination, these results suggest that increasing the accuracy of the direction estimate increases the stability of the aggregate. However, by eliminating the random fluctuations in the A-Cells' positions it becomes more difficult for the aggregate to break out of locally minimal configurations.

3.6 Summary

The results of the experiments presented in this chapter demonstrate that the cell adhesion model is capable of robust self-aggregation. More specifically, the structures that are formed actively minimise their exposed perimeter while maintaining a stable configuration of units in what is approximate hexagonal close pack (e.g. the optimal packing of spheres on a plane). These behaviours are a direct result of the simple dynamics resulting from the mutual adhesion exhibited by the A-Cells and match those predicted by the DAH.

One of the most significant results that was obtained from the presented simulations relates to the observed effect of random forces. In line with evidence from other

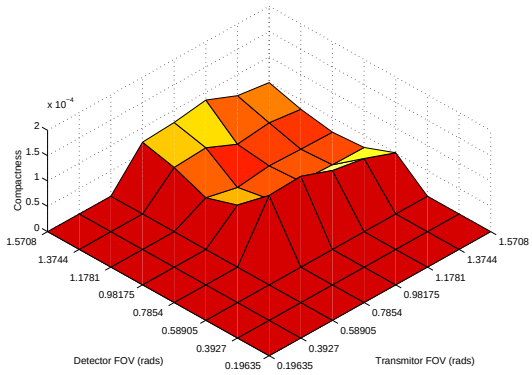


Figure 3.45: *Compactness variance for a range of transmitter and detector FOVs.*

models [40, 45], it has been shown that random fluctuations in the A-Cells' positions directly relate to the aggregate's ability to escape from locally minimal configurations (see section 3.3). This can be seen in the form of a significant increase in the compactness of the final aggregates that are achieved in the presence of higher random movement forces. In addition, the resulting increase in local exploration has the additional effect of increasing the rate at which the aggregates converge to their equilibrium state. However, more specifically to the robotic structure of the A-Cells, the simulations revealed the importance of breaking the symmetry of the alignment between the A-Cells' sensors. This can be achieved by simply exploiting slight turbulence in the environment which randomly rotate the A-Cells around their vertical axis[10].

The other simulations investigated the specificity of various parameter choices, both in the model and the A-Cell's physical structure. In each case the model demonstrated a high tolerance to the exact values chosen, achieving reasonably high performance in most cases where the aggregate was able to remain intact. In addition, the robust nature of this type of model was repeatedly highlighted by the fact that the inherent errors present in the system simply led to slight fluctuations in the A-Cell movements which ultimately helped drive the system to its equilibrium state.

The following general results are of particular significance:

[10]The HYDRON module the A-Cells were modelled on actually has a natural rotation around this axis introduced by the force generated between the module's central impeller and the surrounding fluid.

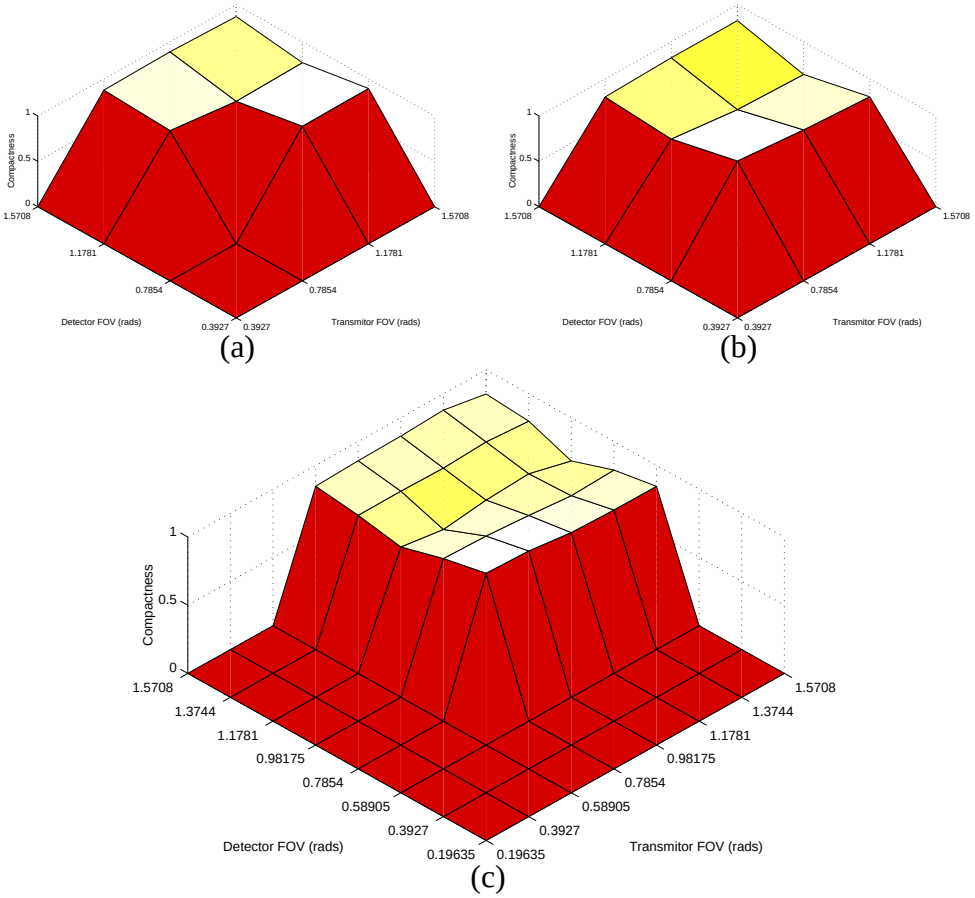
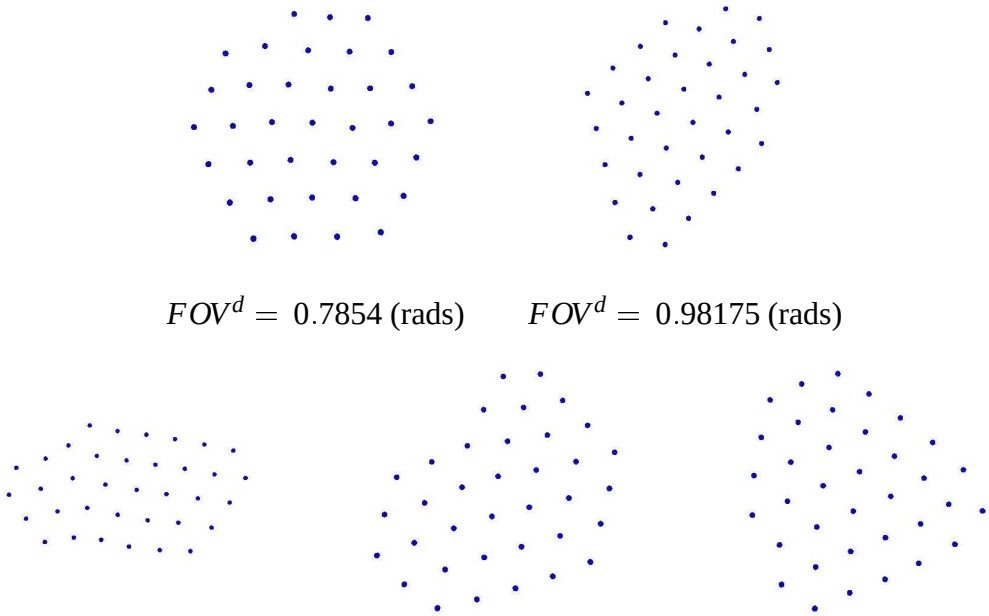


Figure 3.46: CC *values of A-Cell aggregates over a range of transmitter and detector FOVs. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown. (a) 16 A-Cell aggregate (b) 64 A-Cell aggregate (c) 36 A-Cell aggregated*



$FOV^d = 0.7854$ (rads) $FOV^d = 0.98175$ (rads)

$FOV^d = 1.1781$ (rads) $FOV^d = 1.3744$ (rads) $FOV^d = 1.5708$ (rads)

Figure 3.47: *A selected of the final configurations that were achieved for a selection of detector FOVs (FOV^d) and a transmitter FOV of 0.7854 (rads).*

- Equilibrium distance should be less than or equal to the A-Cell virtual radius (see section 3.4.3).
- Sensor FOV should be sufficient to prevent coverage from diverging with distance from the A-Cell hull (see section 3.5.3).

The other simulations did reveal particular optimum values. However, these are most likely to be a consequence of the base parameter set which was chosen. Despite this, the general trends that were revealed should still be informative for fine tuning the performance of any parameter set.

Chapter 4

Exploration of 2D Sorting

In this chapter we investigate the more complex case where two A-Cell types are intermixed. Again the A-Cells should move toward the most thermodynamically stable configuration. However, in addition to the rounding behaviour, this also causes the distinct A-Cell types to form one of four possible hierarchical relationships. The full details of this are discussed in section 1.3.

As in chapter 3, the A-Cell movements in the following experiments are restricted to the horizontal plane.

4.1 Methodology

4.1.1 Measures

The experiments in this chapter are concerned with the hierarchical relationship between two A-Cell types in an aggregate and more specifically, the level of mixing or sorting that occurs between them. The most appropriate method of measuring this is to observe the total area of A-Cell membranes which bond between the different A-Cell types and additionally any area of the membranes which remains unbonded. In total, this gives five values which can be considered in an aggregate of two A-Cell types (*a*and *b*).

1. Bonded area of membrane between type *a*A-Cells (**a** $\rightarrow$ **a**).

2. Bonded area of membrane between type *a*and type *b*A-Cells (**a** $\rightarrow$ **b**).

3. Bonded area of membrane between type *b*A-Cells (**b** $\rightarrow$ **b**).

4. The area of membrane on type *a*A-Cells which is unbonded.

5. The area of membrane on type *b*A-Cells which is unbonded.

As the A-Cells in the aggregate move toward some stable configuration, these values should increase and decrease respectively, clearly indicating the state of the aggregate at any point.

There are two possible methods of measuring this bonded area of membrane. Either the information communicated between the A-Cells during a simulation can be used to show the area of membrane each A-Cell makes available for bonding with its various neighbours or, alternatively, the position of the A-Cells and their membrane size can be used to calculate the actual areas of an A-Cell's membrane which overlap with those of its neighbours.

Each of these approaches has advantages and disadvantages. For example, the A-Cells can only use estimates of their neighbours' positions when determining the areas of membrane they make available for bonding. Therefore, these values may not represent the true configuration of the aggregate. However, it is possible to simply strip these values from the transmission signals passed between the A-Cells without carrying out any further computation. Conversely, calculating the actual overlaps from the A-Cells' positions and membrane sizes does require extra processing and, unlike the former approach, this will not account for areas of membrane which are overlapped by multiple other membranes [1].

Figures 4.1, 4.2, 4.3 and 4.4 show the results that were calculated by applying each of the approaches to a simulation where the A-Cells sort from an evenly mixed grid into the typical 'onion' configuration (see fig. 4.5). Two general observations of these results relate to the methods by which they were calculated. Firstly, it can be seen that the area of bonded membrane results calculated from the values communicated between the A-Cells (see fig. 4.1) all start at zero and immediately jump to a much

[1] This could be accounted for, but would add considerable complexity and consume additional processing time. In addition, the results that are presented show that the simpler approach is adequate.

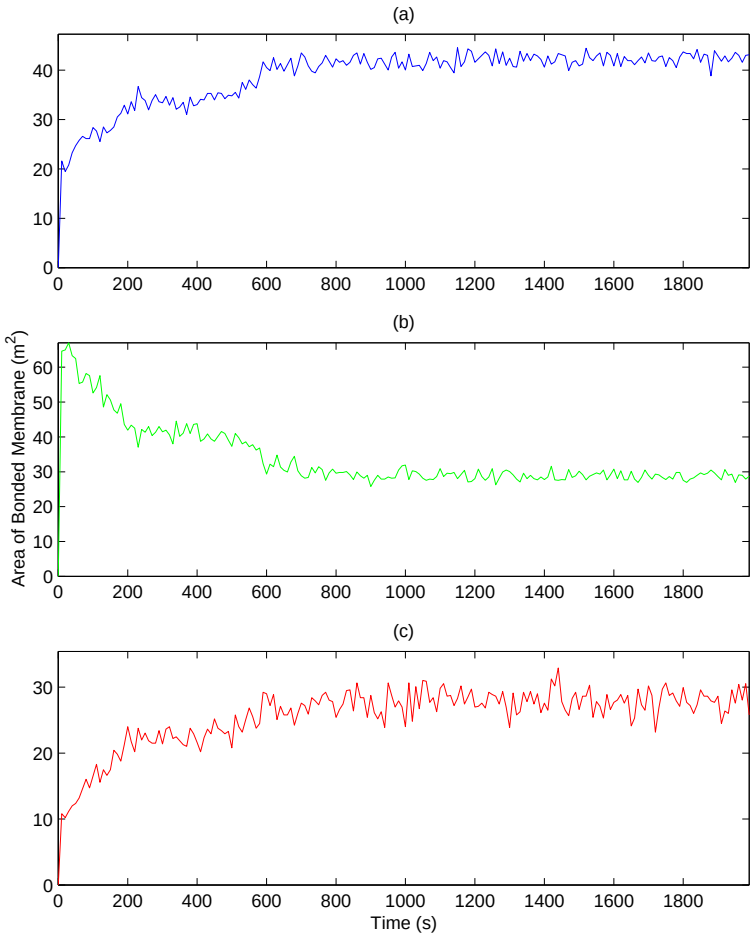


Figure 4.1: *The areas of bonded membrane for the three combinations of the two A-Cell types* a*and* b*obtained using position estimates. (a) Bonded area of membrane for* $\mathbf{a} \rightarrow \mathbf{a}$ *. (b) Bonded area of membrane for* $\mathbf{a} \rightarrow \mathbf{b}$ *. (c) Bonded area of membrane for* $\mathbf{b} \rightarrow \mathbf{b}$ *.*

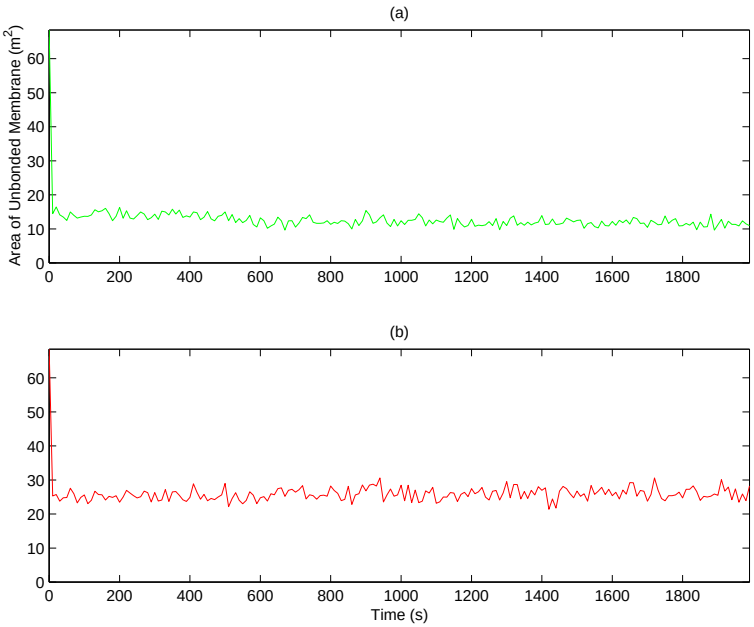


Figure 4.2: *The unbonded areas of membrane for the two A-Cell types* a *and* b *obtained using position estimates. (a) Unbonded area of membrane for type* a *A-Cells. (b) Unbonded area of membrane for type* b *A-Cells.*

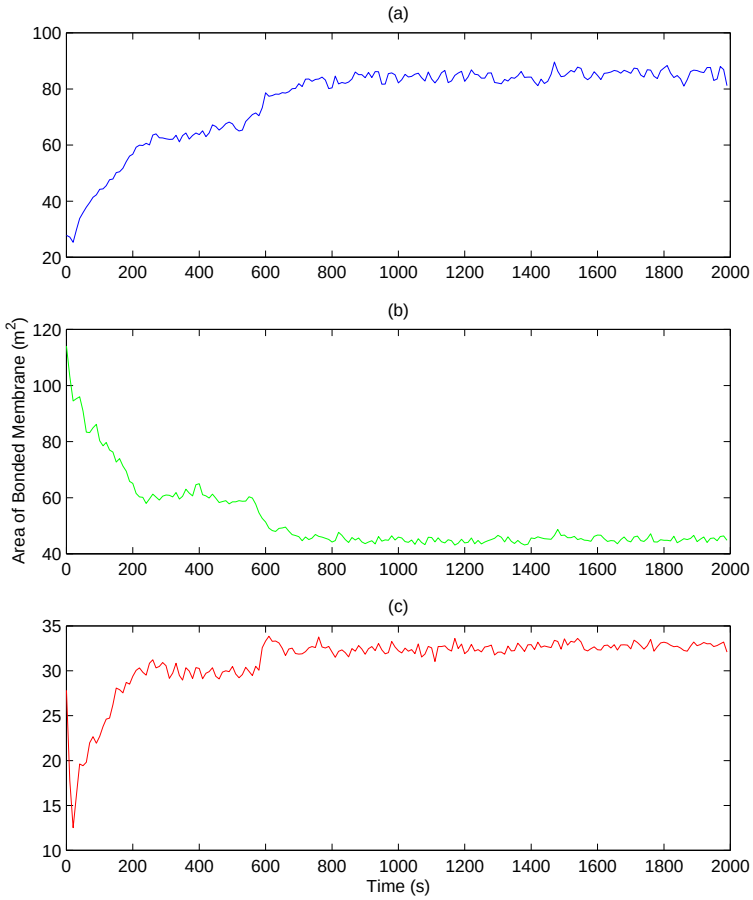


Figure 4.3: *The areas of bonded membrane for the three combinations of the two A-Cell types* a*and* b*obtained using known positions. (a) Bonded area of membrane for* $\mathbf{a} \rightarrow \mathbf{a}$ *. (b) Bonded area of membrane for* $\mathbf{a} \rightarrow \mathbf{b}$ *. (c) Bonded area of membrane for* $\mathbf{b} \rightarrow \mathbf{b}$ *.*

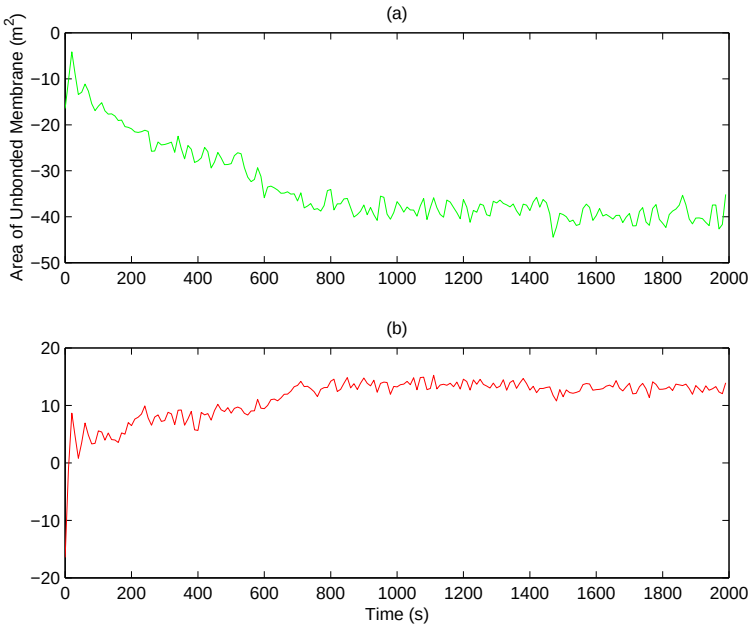


Figure 4.4: *The unbonded areas of membrane for the two A-Cell types* a*and* b*obtained using known positions. (a) Unbonded area of membrane for type* a*A-Cells. (b) Unbonded area of membrane for type* b*A-Cells.*

higher value (the area of unbonded membrane results show a corresponding drop). However, this is simply due to the fact that the A-Cells in the simulation start with no knowledge of their neighbours. Therefore, initially they have no bonds even though their membranes are in contact. Secondly, the area of unbonded membrane calculated using the actual A-Cell positions show some negative values (see fig. 4.4). In this case the cause is the failure to account for multiple membrane overlaps. Therefore, the area of bonded membrane may appear larger than the actual area available and thus the unbonded area appears negative.

The two approaches can be compared in more detail by considering how well they reflect the observed changes in the aggregate's configuration as the simulation progresses. The screen shots from the simulation (see fig. 4.5) show that during the first 200 seconds there is a very rapid change with the more strongly adhering type *a*A-Cells rapidly sorting into two clusters. These then round up slightly before merging at approximately 600 seconds into the simulation. Finally, after almost 900 seconds the aggregate is completely sorted with both the central cluster of type *a*A-Cells and the whole aggregate rounding up as expected. In each set of results, the two largest changes in the aggregate configuration which occur when the initial clusters of type *a*A-Cells form and finally merge are marked by significant changes in each type of bonding. However, it can be seen that these are most clearly defined in the results calculated from the A-Cell positions (see fig. 4.3 & 4.4). In addition, only this set of results show the more subtle changes, as initially both the type *a*A-Cell clusters and finally the single type *a*A-Cell cluster rounds up (see fig. 4.3 - (a) 200s–600s & 600s–900s). There is also a clear difference between the two approaches in the results showing the unbonded area of membrane. The results calculated from the communicated values show no significant changes. However, the results calculated from the A-Cell positions show both the decrease in the area of unbonded membrane for the type *a*A-Cells as they round up and form a central engulfed cluster and also the corresponding increase for the type *b*A-Cells as they are forced out to the external boundary of the aggregate.

This comparison shows quite clearly that the latter approach of calculating the actual areas of overlapped membrane from the A-Cell positions produces the most

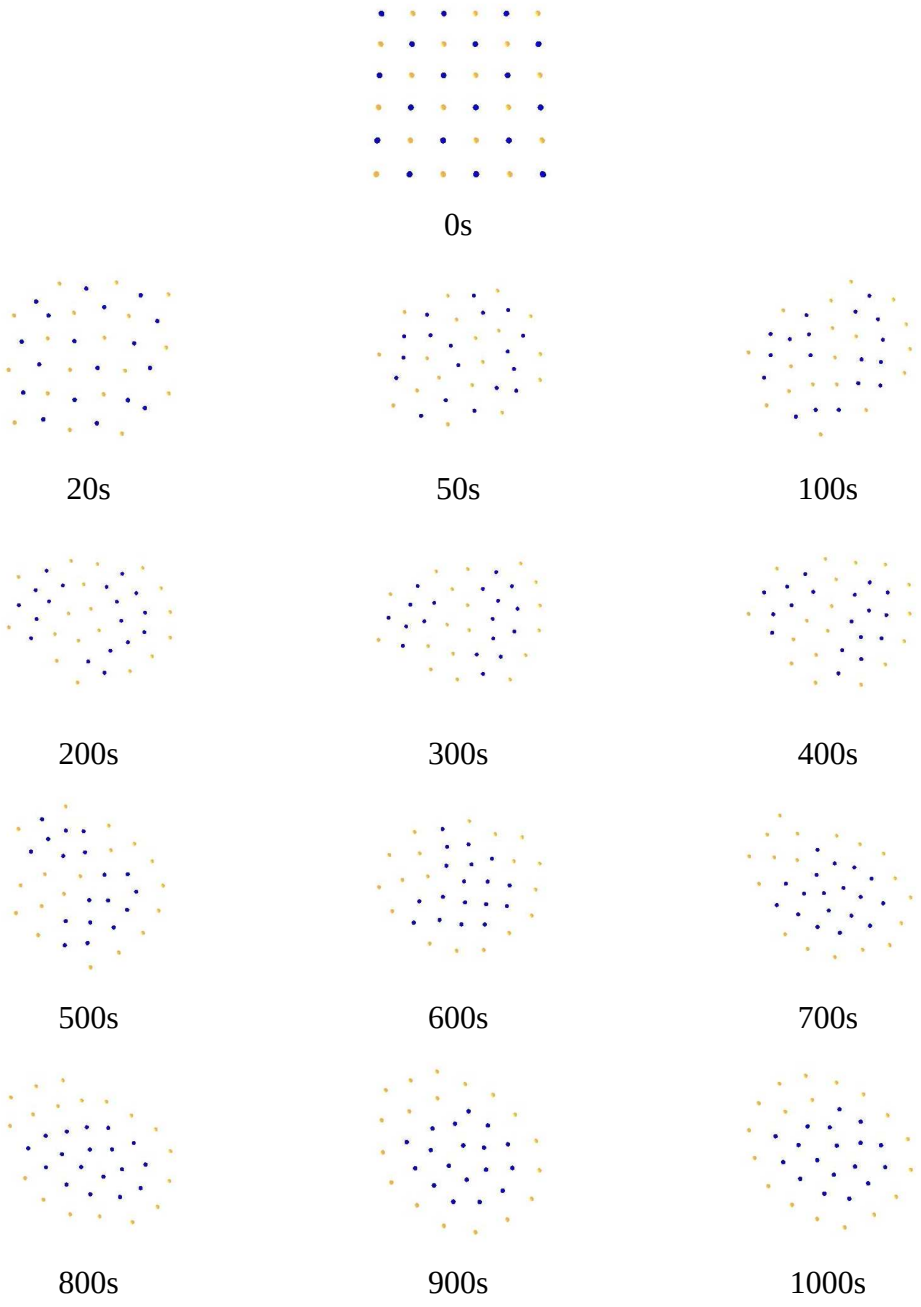


Figure 4.5: *Screen shots from the simulation of an aggregate consisting of two A-Cell types. The simulation shows the sorting case where the A-Cells form the onion configuration (see section 1.3).* $W_{aa} > W_{ab} \geq W_{bb}$ *and* $W_{ab} \leq (W_{aa} + W_{bb})/2$*, type* a= *blue (dark) A-Cells, type* b= *orange (light) A-Cells, W = the work of adhesion*

detailed results which very closely match the observed changes in the configuration of the A-Cells. Therefore, it is this method which is used when evaluating the experiments presented in the chapter. However, as the equilibrium distance between the A-Cells is dependent on the relative adhesions between them (see section 4.1.5.2), if the A-Cells have a greater level of adhesion more of their membrane area will bond. Therefore, to fairly compare the bonded membrane area values between A-Cells with differing levels of adhesion, the values must first be normalised.

The approach taken in this work is to divide the values by the estimated maximum area of membrane which could bond for A-Cells with a specific level of adhesion.

$$B = \frac{O}{E} \tag{4.1}$$

where O is the observed area of bonded membrane (m^2), E the estimated maximum area of membrane that could bond (m^2), and B the bonding index.

This maximum value will occur when the specific A-Cells round up to form a disc. Therefore, it is possible to calculate these estimates by simply carrying out additional simulations consisting of only the required number of A-Cells with the specific level of adhesion. Thus in the previous example, three additional simulations would be required with the following properties:

1. 18 A-Cells with the same adhesion as $\mathbf{a} \rightarrow \mathbf{a}$.
2. 36 A-Cells with the same adhesion as $\mathbf{a} \rightarrow \mathbf{b}$.
3. 18 A-Cells with the same adhesion as $\mathbf{b} \rightarrow \mathbf{b}$.

The convergence value for the area of bonded membrane can be calculated for each simulation, giving the estimated maximum area of membrane which can bond. In practice, to improve the accuracy of the estimate, the simulations are carried out in samples of 15 and the maximum value from each sample is used.

Unfortunately, it is not possible to normalise the unbonded area of membrane values in a similar way, as the areas of membrane which are unbonded ultimately depend on all the bonds made by an A-Cell. For example, if one type of A-Cells form a single

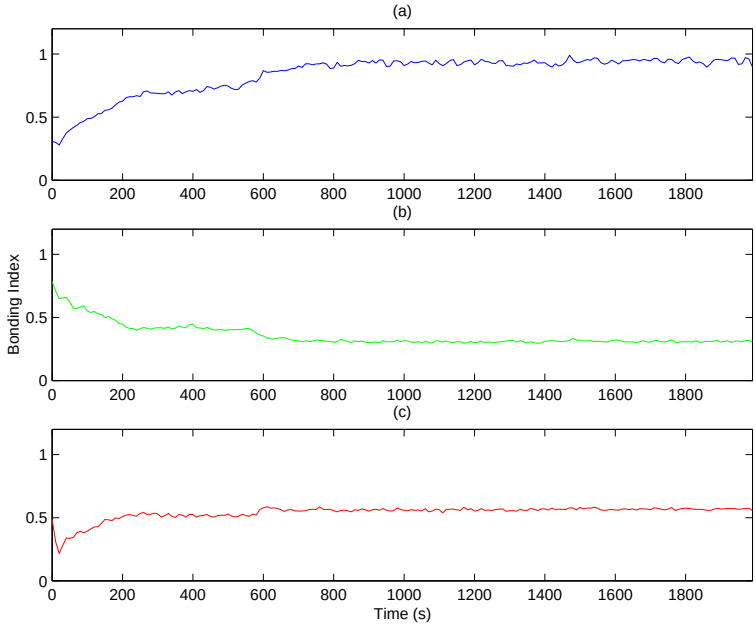


Figure 4.6: *The normalised areas of bonded membrane for the three combinations of the two A-Cell types* a*and* b*. (a) Bonded area of membrane for* $\mathbf{a} \rightarrow \mathbf{a}$ *. (b) Bonded area of membrane for* $\mathbf{a} \rightarrow \mathbf{b}$ *. (c) Bonded area of membrane for* $\mathbf{b} \rightarrow \mathbf{b}$ *.*

rounded cluster, their level of homotypic bonded membrane will remain the same. This applies whether they are engulfed by another type of A-Cell which adheres strongly or weakly with them. The area of unbonded membrane on the other hand, will be different in each case. However, it can be seen from the previous simulation that it is possible to tell the configuration of the A-Cells using only the bonding information. Therefore, the areas of unbonded membrane can be safely ignored.

Figure 4.6 shows the normalised results from the previous simulation. These adjusted results make the state of the aggregate much clearer. For example, it can be seen from the $\mathbf{a} \rightarrow \mathbf{a}$ bonding index (see fig. 4.6 - (a)) that the type *a*A-Cells have formed a single rounded cluster. On the contrary, the lower $\mathbf{b} \rightarrow \mathbf{b}$ index indicates that this is not the case for the type *b*A-Cells. This in conjunction with the presence of some heterotypic bonding indicates that the type *b*A-Cells have at least partially engulfed the type *a*A-Cells, the level of engulfment relating directly to the level of the

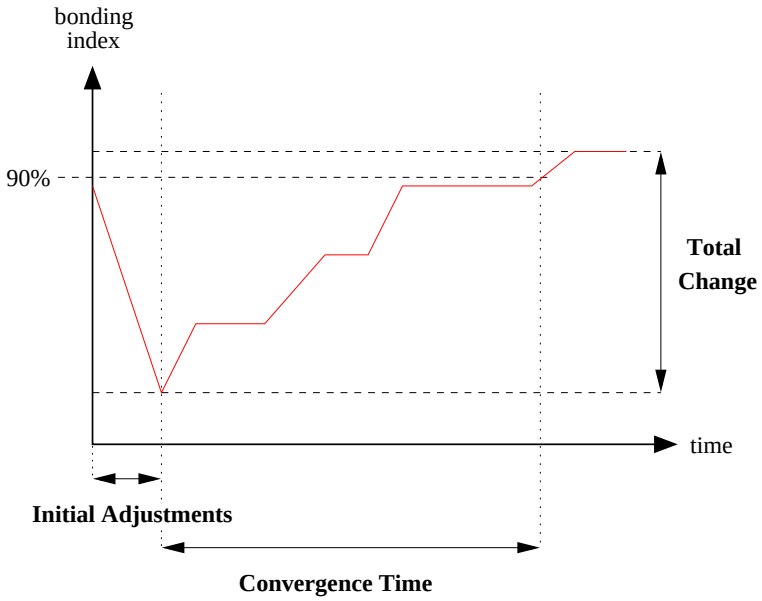


Figure 4.7: *The convergence time of an A-Cell bonding index.*

a $\rightarrow$ **b** bonding index.

4.1.2 Convergence Time and Converged Bonding Indices

In this chapter, to help highlight any trends which result from parameter changes in the model, the individual sets of bonding indices are decomposed into two key values:

- The Convergence Time (CT)
- The Converged Bonding Indices (CBI)

As the bonding indices converge when an aggregate reaches some equilibrium configuration, the CBI can be calculated by simply averaging the values for the last *n* seconds of a simulation (where *n* is large enough to account for any minor fluctuations). As with the measures described in section 3.1.1.2 this assumes that the simulation time is sufficient for convergence to occur. However, in the following experiments the accompanying CT measure should indicate any cases where this is not the case.

The CT measure, like the CR (see section 3.1.1.2), uses a variation of ‘rise time’. However, unlike the compactness measure the CR is calculated from, the bonding in-

dices can either increase or decrease over the course of a simulation. Therefore, in this case, the CT is defined as the length of time for a value to change by ninety percent of the total difference between the initial value and the converged value (see fig. 4.7). Additionally, during the first seconds of a simulation, the A-Cell bonding can change rapidly as the A-Cells quickly move from the initial configuration to reach an equilibrium distance with their neighbours. It is only after this point that the values actually begin to converge. Therefore, it is necessary to allow some period of time for this to occur. Thus, when the CT is calculated, the first minute of the simulation is ignored. This extra time is then added on to the calculated value to determine the final result.

Bonding	$\mathbf{a} \rightarrow \mathbf{a}$	$\mathbf{a} \rightarrow \mathbf{b}$	$\mathbf{b} \rightarrow \mathbf{b}$
CT (s)	680	610	590

Table 4.1: *The* CT *values calculated for an example simulation (see fig. 4.6).*

Table 4.1 shows the results of using this approach to calculate the CT values for the example simulation presented in section 4.1.1. These results could be summarised using either the maximum, mean, or median (see table 4.2). However, as the measure should show the total convergence time for the whole aggregate, the maximum value is most suitable. Therefore, it is this value that is used in the results presented in this chapter.

	Max	Median	Mean
CT (s)	680	610	626.6667

Table 4.2: *Statistics calculated for the results shown table 4.1.*

4.1.3 Sample Size

As in chapter 3 the sample size used is 15. Again, this lower value has been chosen to allow adequate exploration of the model, making best use of available resources and time. However, as before, the accuracy of the confidence error estimates has been improved through the use of bootstrapping[2].

[2]The original sample was resampled 5000 times to determine the error estimates. Full details of this procedure are given in chapter 3.

4.1.4 Initial Configurations

When evaluating aggregates of two A-Cell types there are two extreme configurations which can be considered. Either the A-Cells can be evenly mixed or alternatively they can be sorted into two homotypic clusters. In the following sections, the model's performance is evaluated for both these situations by using two contrasting grid formations as the initial configuration:

1. An evenly mixed grid
2. A sorted grid

The main behaviours which are investigated are A-Cell sorting, where the A-Cells should form the typical 'onion' configuration, and A-Cell mixing, where the A-Cells should form an evenly mixed aggregate. In the mixing case, these configurations test the model for conditions which are near and far from the goal state. However, in the sorting case, these two configurations allow two distinct types of sorting to be evaluated. In the first case with an evenly mixed grid, the A-Cell types will reach the equilibrium state through dramatic rearrangements which cause one A-Cell type to migrate centrally, while in the latter case, the more weakly adhering A-Cells will slowly envelop the existing other cluster until it is fully engulfed.

As in the rounding experiments, each set of simulations are carried out on aggregates of 36 and 64 A-Cells. However, the smaller 16 A-Cell aggregate size was not considered for these experiments, as sorting and intermixing in such a small aggregate should be reasonably trivial.

4.1.5 Base Parameter Set

The base parameter set presented in section 3.1.3 is again used throughout this chapter with a few additions to account for two A-Cells types.

4.1.5.1 A-CAM Relationships

In the cell adhesion model there are two possibilities for creating different adhesions between different A-Cell types:

- The A-Cell types can simply have different numbers of the same A-CAM. Therefore, the different densities of A-CAMs will result in the different homotypic and heterotypic adhesions (see section 2.2.4).
- The A-Cell types can express different types of A-CAM. Therefore, the bonding relationships between the A-CAM types will influence the different adhesions.

In this chapter the latter of these approaches is used such that the A-Cells express the following types of A-CAM:

- Type a A-Cells – A-CAM cam1 and A-CAM cam2
- Type b A-Cells – A-CAM cam3 and A-CAM cam4

with the following bonding relationships:

- cam1 bonds with itself
- cam2 bonds with cam3
- cam4 bonds with itself

However, for simplicity, the numbers of cam2 and cam3 bonds are kept at equivalent levels. This allows the actual adhesion for each of the three bonds ($\mathbf{a} \rightarrow \mathbf{a}$, $\mathbf{a} \rightarrow \mathbf{b}$, and $\mathbf{b} \rightarrow \mathbf{b}$) to be very precisely specified.

In addition, the purpose of the following experiments is to demonstrate that only differences in the adhesion between the A-Cells are necessary to create the required sorting dynamics. Therefore, all other membrane properties are kept constant. The consequence of this is that the more strongly adhering A-Cells will have a slightly shorter equilibrium distance (e.g. the adhesion has been increased while the repulsion factor has been kept constant). However, where this obscures the results, additional experiments have been carried out to control for any effects (see section 4.2.2).

The alternative approach would be to adjust the repulsion factor of the membranes for each A-Cell type independently, such that a particular equilibrium distance is achieved. However, this would only specify the equilibrium distance for homotypic

bonds and the equilibrium distance for the heterotypic bonds would still be determined by the interaction between the different membranes. In addition, a more complex membrane collision model would also be required as it could no longer be assumed that each membrane would compress equally (see section 2.2.3).

4.1.5.2 Homotypic and Heterotypic Work of Adhesion

In the traditional sense, work of adhesion refers to the the work done in the formation of an adhesion over unit area (see section 1.3). However, this assumes that all the units have an equal capability of forming a unit area of bonded surface. This is not the case with the A-Cells because of the force field nature of the membrane (i.e. stronger adhering A-Cells can overcome more of the repulsive force generated by the membranes and form a larger area of bonded surface) (see section 2.2.3). Therefore, the actual measure used is the attraction force which exists between the A-Cells when they are at an equilibrium distance, as this relates directly to the overall strength of bond which the A-Cells may actually form.

The default homotypic adhesion values which are considered throughout this chapter are $\mathbf{b} \rightarrow \mathbf{b}$ $0.15125N$ and $\mathbf{a} \rightarrow \mathbf{a}$ $0.3025N$ $(2 * 0.15125)$. However, a comparison of various differences in homotypic adhesion is given in section 4.2.2. These specific values were chosen for several reasons. The lower of the two values was selected primarily as this was the adhesion used throughout the rounding experiments. Therefore, appropriate levels of random movement forces and torques for this level of adhesion were already established (see section 4.1.5.3). Additionally, in conjunction with the other membrane properties, this level of adhesion produces an equilibrium distance of $0.55m$ and the experiments carried out in section 3.4.3 show that the performance of the model drops off rapidly when the equilibrium distance increases above this. The higher value of $0.3025N$ was selected as it creates a reasonable level of difference between the two adhesion values and the equilibrium distance it produces ($0.322183m$) is still relatively large.

The heterotypic adhesion has two default values. These have been selected such that the equilibrium state is either the onion configuration or intermixed. The actual values used are:

- Sorted onion: 0.18675 or $W_{bb} + ((W_{aa} - W_{bb})/4)$
- Intermixed: 0.2646875 or $W_{bb} + 3((W_{aa} - W_{bb})/4)$

These values were chosen as they lie exactly half way between the homotypic adhesion values and their average. In the sorting case, as this is half way between two different states it should result in the best performance (see fig. 4.8).

4.1.5.3 Random Forces

The final difference between this configuration and that outlined in the previous rounding experiments is that random movement forces with a maximum of $0.02N$ are also applied by default. These are included as it has been established that random movements are beneficial to the dynamics of the model and it has already been established that aggregates of A-Cells with the lowest homotypic adhesion outlined above are easily able to remain intact for this level (see section 3.3).

4.2 Differential Adhesion

The experiments presented in this section evaluate how well the model's behaviour corresponds with that predicted by the DAH. It is clear from the adhesion relationships presented in section 1.3 that if the homotypic adhesions are fixed such that $\mathbf{a} \rightarrow \mathbf{a} > \mathbf{b} \rightarrow \mathbf{b}$ then the final configuration is dictated by only the heterotypic adhesion. In particular, there are three key values of heterotypic adhesion which border changes in behaviour (see fig. 4.8).

1. $W_{ab} = 0$
2. $W_{ab} = W_{bb}$
3. $W_{ab} = (W_{aa} + W_{bb})/2$

where W_{aa}, W_{ab} and W_{bb} represent the work of adhesion between the cell types, or the attraction at equilibrium values for the model.

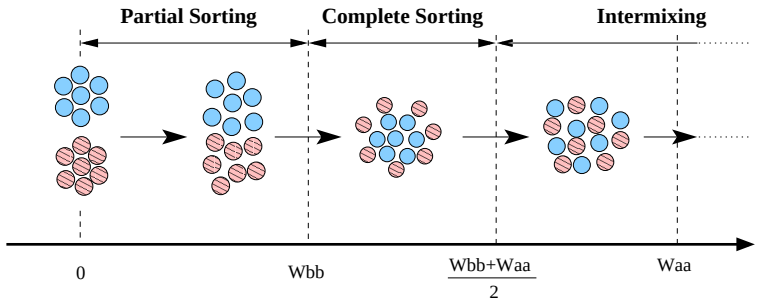


Figure 4.8: *The effect of heterotypic adhesion on the final configuration of an aggregate consisting of two cell types (*a*and* b*, Waa > Wbb).*

The experiments presented in this section consist of two sets. In the first set, a detailed range of heterotypic adhesions covering all the key values are considered. In the second set, a smaller range of heterotypic adhesions are used to evaluate the effect of increasing or decreasing the difference between the homotypic adhesions.

4.2.1 Heterotypic Adhesion

In the following experiments, the two A-Cell types used have the same properties as those defined in the base parameter set (see section 4.1.5). In particular the homotypic adhesions are fixed at:

- $\mathbf{b} \rightarrow \mathbf{b}$ – attraction at equilibrium of $0.15125N$
- $\mathbf{a} \rightarrow \mathbf{a}$ – attraction at equilibrium of $0.3025N$

The range of heterotypic attraction at equilibrium values, $0N$ to $0.45375N(3{*}\mathbf{a} \rightarrow \mathbf{a})$, has been selected to cover all the key values discussed at the beginning of this section. The range also extends above the $\mathbf{a} \rightarrow \mathbf{a}$ value to investigate whether there is any significant change in the level of intermixing when the strength of $\mathbf{a} \rightarrow \mathbf{b}$ adhesions exceeds both homotypic values.

Figures 4.10, and 4.15 show the CBI results that were obtained using 36 and 64 A-Cell aggregates starting from intermixed grids. In general it can be seen that the CBI for both aggregate sizes differ very little, especially when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion reaches

at least $0.15125N$. In addition, in both cases the error bars indicate that there is very little variance, which suggests that the aggregates are consistently achieving the same equilibrium state. There is one clearly erroneous value for the $\mathbf{a} \rightarrow \mathbf{b}$ CBI of the the 64 A-Cell aggregate results when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion is $0N$. However, this is simply a result of normalising the values. As there is no $\mathbf{a} \rightarrow \mathbf{b}$ adhesion, the maximum area of bonded membrane (the normalisation value) should be 0. However, as these values were determined by simulation, and areas of overlapping membrane are considered to be bonded (see section 4.1.1), the calculated results although very small were non-zero due to chance collisions. Therefore, although the actual area of bonded membrane which was calculated was very small, the normalised value appears exceptionally large.

More specifically, the results closely match the expected behaviour, with the $\mathbf{a} \rightarrow \mathbf{b}$ CBI steadily increasing to a maximum value when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion becomes greater than the average of the homotypic adhesions. This increase initially corresponds with a decrease in the $\mathbf{b} \rightarrow \mathbf{b}$ CBI as the type *b*A-Cells increasingly engulf the type *a*A-Cells. However, both homotypic CBIs then drop as the aggregates begin to form intermixed equilibrium states. These changing equilibrium states can be clearly seen in figure 4.14 which shows some of the final configurations observed in the simulations.

The only results which had a slightly different outcome from that predicted by the DAH were those when $\mathbf{a} \rightarrow \mathbf{b}$ adhesions were below $0.1134375N$. In these few cases the DAH suggests that the A-Cells should round to form two clusters with one partially engulfing the other. However, the simulations show that these low $\mathbf{a} \rightarrow \mathbf{b}$ adhesions are not sufficient to hold the clusters together and the random forces eventually cause them to break away from each other. Also, the simulations for adhesions of $0.075625N$ and $0N$ show that these lower values cause the aggregate to split into many unconnected clusters. The cause of this is the rapid formation of small homotypic clusters as the initially intermixed A-Cells begin to sort out. These small clusters then break away from each other before they have time to merge. However, the results show that there is an initial increase in the homotypic CBIs as the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion increases, indicating that the increasing $\mathbf{a} \rightarrow \mathbf{b}$ adhesion allows the clusters to remain bonded for longer, and thus increases the chance that they will merge.

(a) (b)

Figure 4.9: *The pattern formed by intermixed A-Cells. (a) Low heterotypic adhesion. (b) Higher heterotypic adhesion pulling the A-Cells closer together.*

Another interesting trait which can be observed in the results is the linear increase in the homotypic CBIs as the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion increases above $0.2646875N$. The $\mathbf{a} \rightarrow \mathbf{b}$ CBI and the screenshots in figure 4.14 show that the aggregates are remaining fully intermixed so the rise is not caused by a change in equilibrium state. Instead, figure 4.9 shows that this is simply an artifact of the increasing $\mathbf{a} \rightarrow \mathbf{b}$ adhesion pulling all the A-Cells closer together. This causes an increase in the overlap of the homotypic A-Cells' membranes, and thus a corresponding increase in their CBIs . No increase can be observed in the $\mathbf{a} \rightarrow \mathbf{b}$ CBI as the normalisation values also increase and scale the results appropriately.

The CT results for these simulations are shown in figures 4.11 and 4.16. Again, there is a reasonably close correspondence between the results obtained for each size of aggregate. In addition, these results also show several interesting features. Most significantly, the largest peak in CT occurs at the boundary value which separates the cases where the A-Cells form intermixed equilibrium states and those where they form sorted equilibrium states. At this boundary value the A-Cells become partially sorted with some type *b*A-Cells remaining trapped in the central cluster. This explains the intermediate values occur in the CBI . The cause of this is that there is no great preference for one equilibrium state over the other. In the simulations this results in type *b*A-Cells occasionally passing in either direction between the central cluster and the outer shell.

There is a second peak when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion is just less than or equal to the $\mathbf{b} \rightarrow \mathbf{b}$ adhesion. In this case, particularly for the lower value of $\mathbf{a} \rightarrow \mathbf{b}$ adhesion, the type *a*A-Cells very quickly form into a few clusters. However, as the type *b*A-Cells

have no preference for heterotypic bonds they do not fully engulf these smaller clusters and instead act as an obstacle which prevents them merging. The smaller clusters then only fully merge when the random forces in the simulation eventually cause them to make contact. This explains both the large CT and the high variance as indicated by the error bars.

The results also show that the CT is much lower for the well defined cases where the equilibrium state is either the onion or intermixed configuration. In the latter case, this is expected as the initial configuration is already mixed. However, the graph does show that there is still a slight decrease in the CT, reaching a minimum only when the heterotypic adhesions become dominant.

The only anomalous result occurs when there is no $\mathbf{a} \rightarrow \mathbf{b}$ adhesion. As discussed previously, when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion is below $0.1134375N$ the aggregate tends to split quickly into smaller homotypic clusters. As these then change very little, the CT should be low. However, an $\mathbf{a} \rightarrow \mathbf{b}$ adhesion of zero produces a much higher CT than the other values shown, even though the aggregate should actually split up more quickly. The cause of this is the higher number of small clusters which are formed when there is no $\mathbf{a} \rightarrow \mathbf{b}$ adhesion. Viewing the actual simulations revealed that although these small clusters are formed rapidly, the high number of them increases the chance that any two will randomly collide later in the simulation, causing a distinct change in the CBI . The random nature of these collisions also explains the high variance shown for the CT .

Figures 4.12 and 4.13 show the CBI and CT results of simulations of the 36 A-Cell aggregates. These were carried out for a more detailed range of $\mathbf{a} \rightarrow \mathbf{b}$ adhesions which covers the change in equilibrium state from onion to the intermixed configuration. The CBI results show two definite states either side of the boundary value where the CBI remains reasonably constant while the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion increases. In addition, these results also suggest that only a slight change is required to flip the equilibrium state from one configuration to the other.

Figures 4.17, 4.21, 4.18 and 4.22 show the CBI and CT results of similar simulations carried out with the alternative initial configuration of a sorted grid. It is clear that the CBI results share many similarities with the same simulations carried out from the intermixed initial configuration. However, there are two significant differences. Firstly,

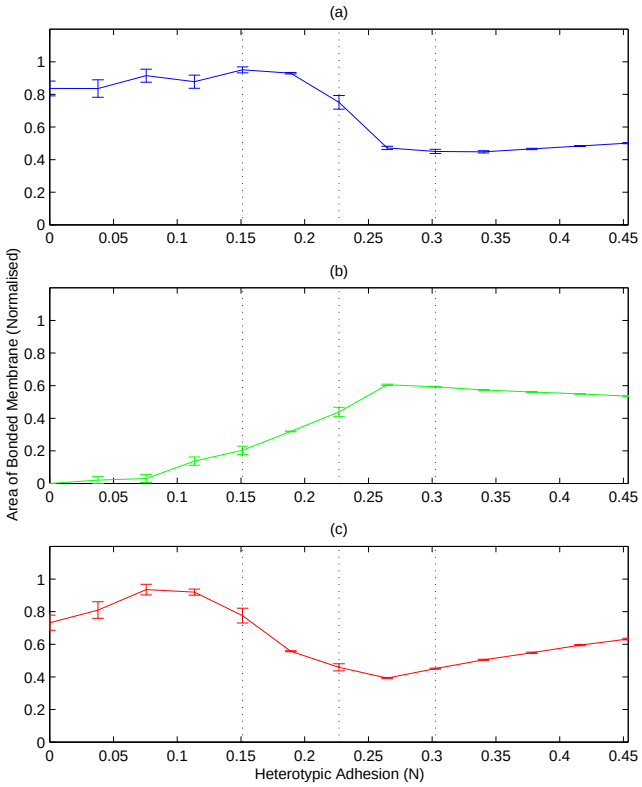


Figure 4.10: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

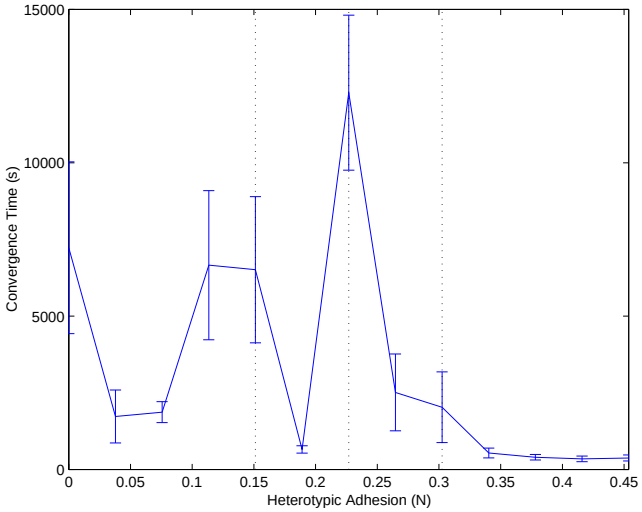


Figure 4.11: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid*

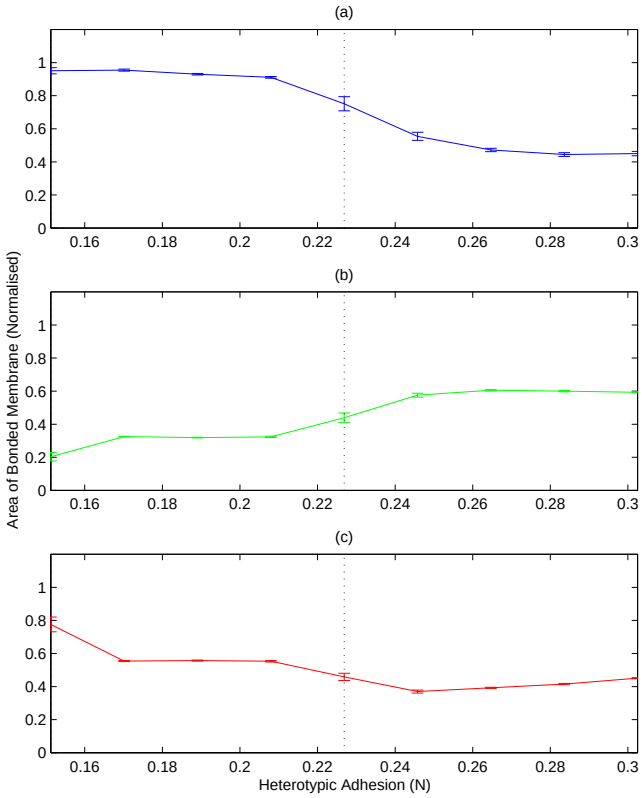


Figure 4.12: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

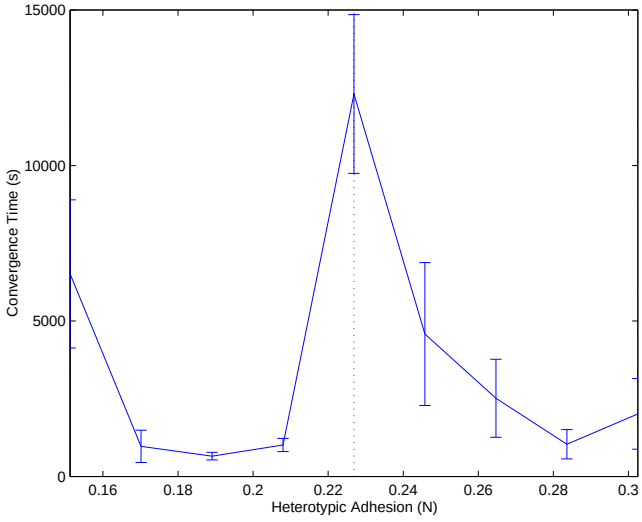


Figure 4.13: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid*

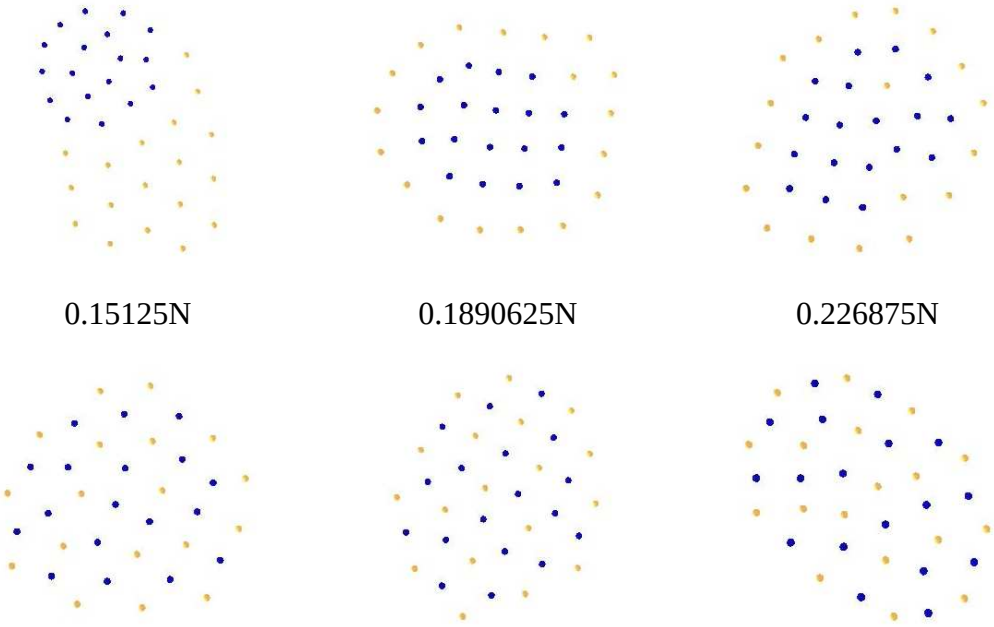


Figure 4.14: *Screen shots showing the final configuration of aggregates consisting of two A-Cell types with various heterotypic attraction at equilibrium values. The aggregates consisted of 36 A-Cells and started with the initial configuration of an intermixed grid.*

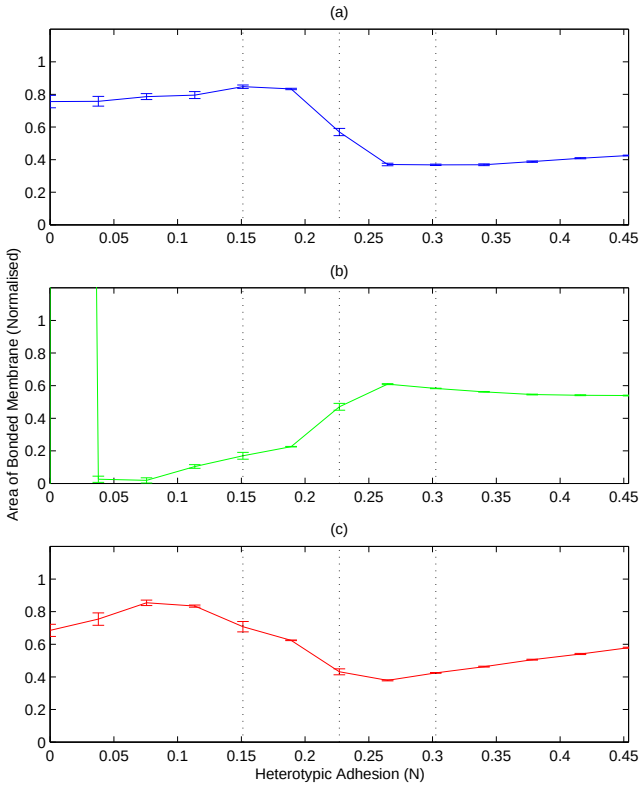


Figure 4.15: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed grid. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

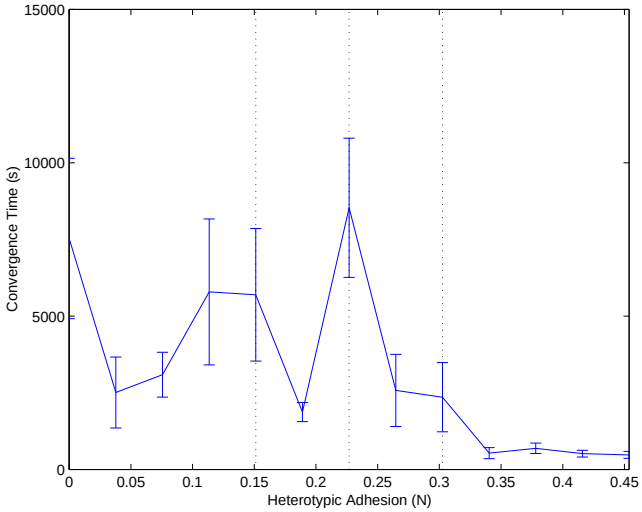


Figure 4.16: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed grid*

the CBI results for the aggregates with $\mathbf{a} \rightarrow \mathbf{b}$ adhesions less than or equal to $0.15125N$, show behaviour which corresponds better with the behaviour predicted by the DAH. The homotypic CBI for these values remain reasonably constant with values just below one, which indicates that as expected they are rounding into homotypic clusters. In addition the $\mathbf{a} \rightarrow \mathbf{b}$ CBI shows a slight rise indicating the type *b*A-Cells are increasingly engulfed. However, it is clear that the low $\mathbf{a} \rightarrow \mathbf{b}$ adhesions below $0.1134375N$ are still too low to hold the clusters together in the presence of the random forces. This results in the two consecutive peaks which can be seen in the CT results. Initially, the increasing $\mathbf{a} \rightarrow \mathbf{b}$ adhesion increases the length of time until the two homotypic clusters split apart. However, once the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion is strong enough to hold the clusters together, the CT drops sharply as the aggregate quickly reaches an equilibrium state. This is followed by a second rise as the increasing $\mathbf{a} \rightarrow \mathbf{b}$ adhesion causes an increase in the final level of engulfment.

The second noticeable difference occurs at the boundary value of $0.226875N$. In the previous results the CBI at this boundary value are intermediate between the CBI for the onion equilibrium state and the intermixed equilibrium state. However, although this is still the case for the 64 A-Cell aggregate, the results for the 36 A-Cell aggregate remain at approximately the same levels as when full sorting occurs. Again, this appears to be a result of the initial configuration. In aggregates starting from an intermixed grid, it is easy for the type *b*A-Cells to become trapped in the cluster of type *a*A-Cells as it forms. Therefore, the final configuration is more intermixed. However, when the aggregates start from a sorted grid, the type *a*A-Cells are quickly engulfed so intermixing can only occur at the border between the A-Cell types. This also explains the difference in behaviour for the two aggregate sizes. In the 36 A-Cell aggregate this boundary is smaller and thus the opportunity for intermixing is lower, resulting in CBI which are closer to those of the fully sorted onion configuration.

The same simulations were, as before, carried out for the 36 A-Cell aggregate with a more detailed range of $\mathbf{a} \rightarrow \mathbf{b}$ adhesions covering the change in equilibrium state from the onion to the intermixed configuration (see fig. 4.19 & 4.20). In this instance the CT results show that there is again a decease in CT for the intermixing case as the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion increases. However, in contrast to the previous results there is

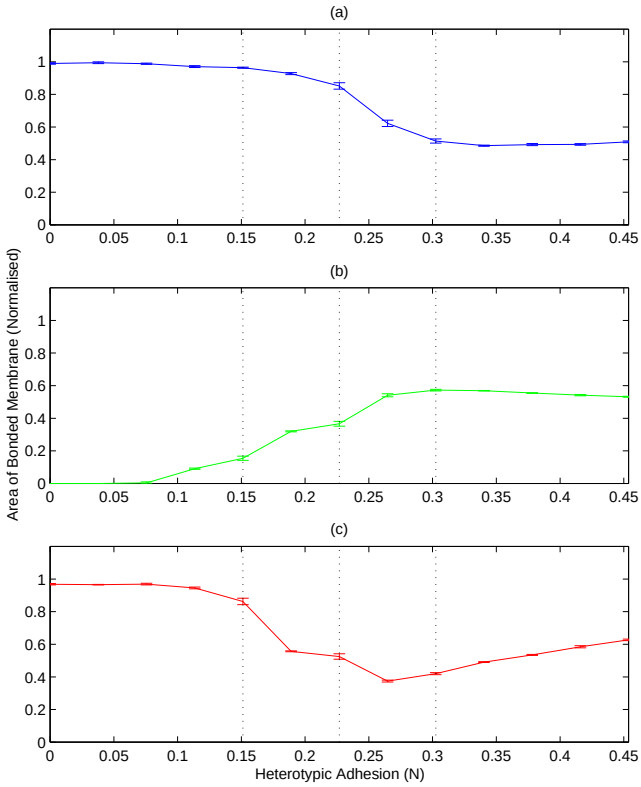


Figure 4.17: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

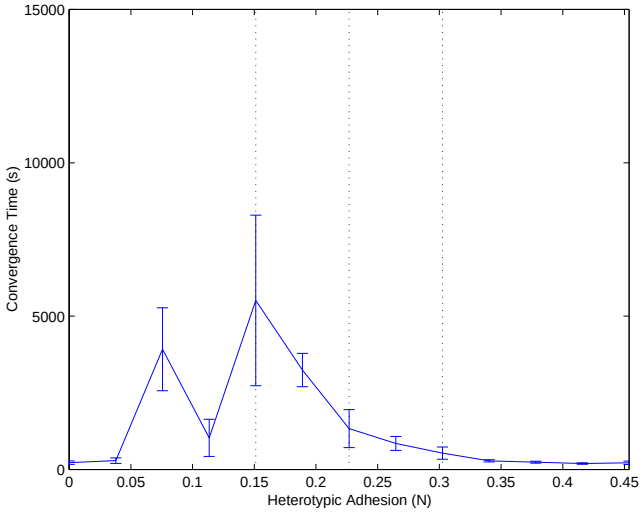


Figure 4.18: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid*

also a decrease in CT for the sorting case as the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion increases. This can be explained by considering figure 4.23. This figure shows a mass of type *b*A-Cells which should engulf the stronger adhering type *a*A-Cells. If random movements cause a type *b*A-Cell (A), which previously had no heterotypic adhesion, to come into contact with a type *a*A-Cell (D) then the stronger heterotypic adhesion should help pull it closer. The net effect of this will be that the A-Cells on either side (B and C) will be pushed left and right respectively, helping move the type *b*A-Cells further round. If the strength of the heterotypic adhesion increases, A-Cell (A) would experience a greater pull toward the type *a*A-Cells helping to speed up the engulfment.

When the sorted onion configuration is achieved from an initially intermixed population, the stronger homotypic adhesion will help drive the sorting through the creation of new bonds between the stronger adhering A-Cells. Therefore, the slight rise in heterotypic adhesion shouldn't have the same positive impact. This matches the trends observed in the actual results.

4.2.2 Difference Between Homotypic Adhesions

In this section, simulations were carried out for aggregates in which a range of homotypic adhesion pairs were investigated. In each case, the weakest adhesion ($\mathbf{b} \rightarrow \mathbf{b}$) was kept constant at $0.15125N$ and the stronger $\mathbf{a} \rightarrow \mathbf{a}$ adhesion was varied. The following four $\mathbf{a} \rightarrow \mathbf{a}$ adhesions were chosen to provide a good coverage of the possible magnitudes of difference.

1. $\mathbf{a} \rightarrow \mathbf{a} = 0.1890625N$
2. $\mathbf{a} \rightarrow \mathbf{a} = 0.226875N$
3. $\mathbf{a} \rightarrow \mathbf{a} = 0.3025N$ (from section 4.2)
4. $\mathbf{a} \rightarrow \mathbf{a} = 0.45375N$

For each pair of homotypic adhesions, simulations were carried out for heterotypic adhesions which increased in range from the weakest homotypic adhesion to the strongest. This range covers the transition in equilibrium state from the onion to the

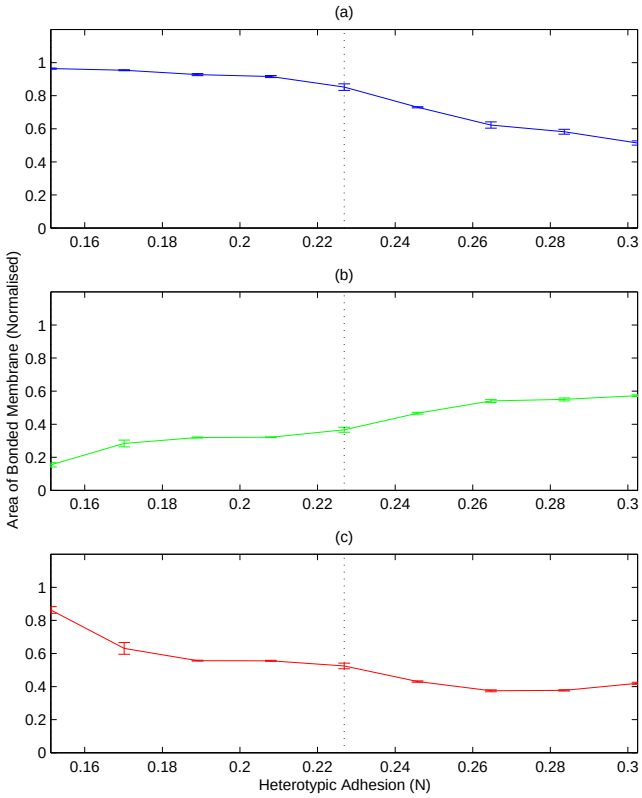


Figure 4.19: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

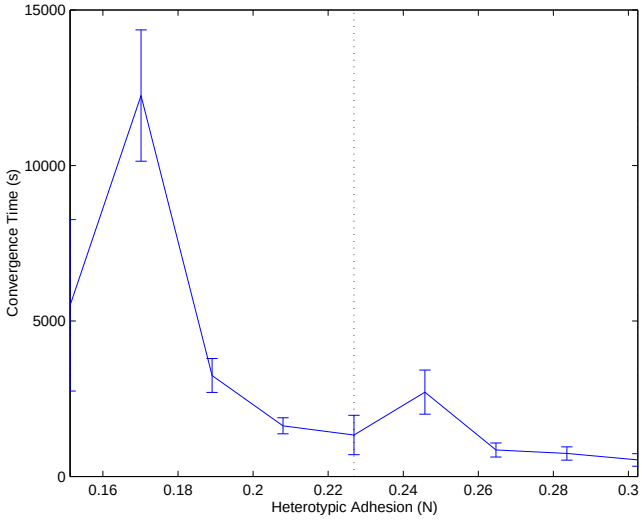


Figure 4.20: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid*

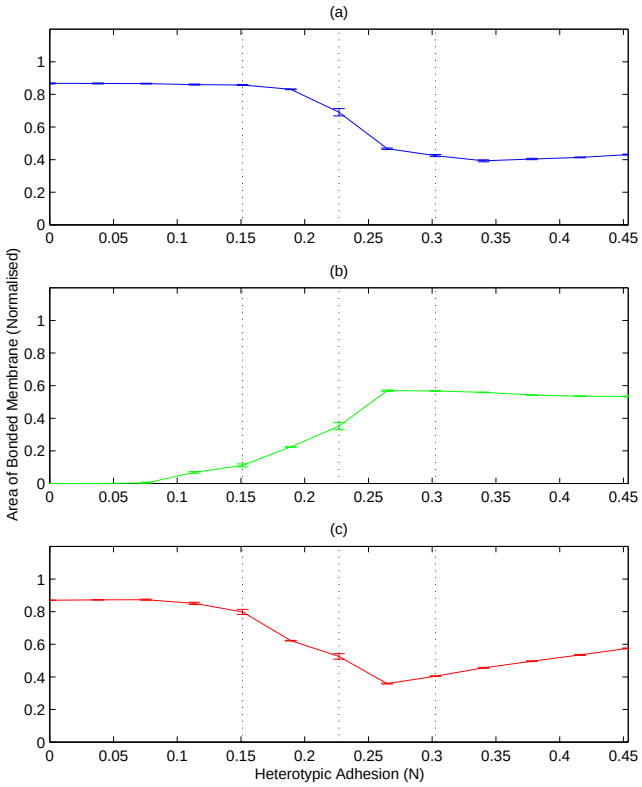


Figure 4.21: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted grid. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

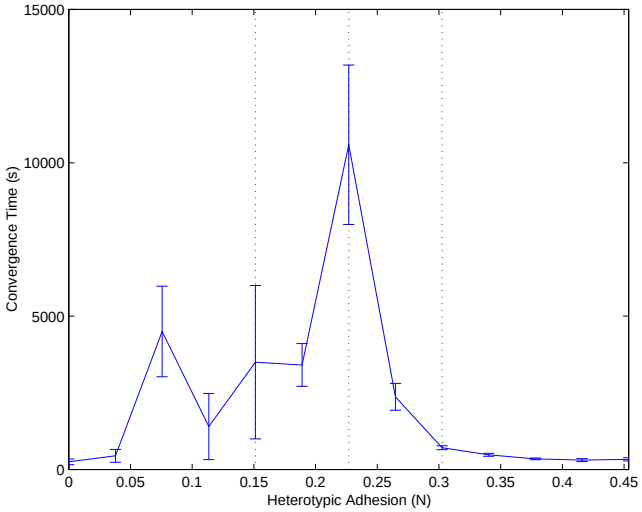


Figure 4.22: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted grid*

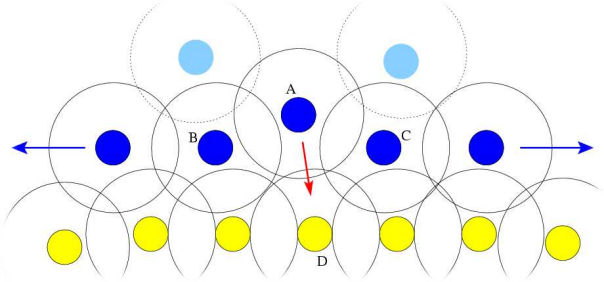


Figure 4.23: *An example of A-Cell engulfment. Type* b*A-Cells are shown in blue (black), type* a*A-Cells are shown in yellow (grey).*

intermixed configuration. In addition, the simulations were carried out for both 36 and 64 A-Cell aggregates starting from initial configurations of an intermixed and sorted grid. Only summaries of these results are presented in this chapter. However, the full range of results for these simulations are included in Appendix B.

The CBI results of the simulations demonstrate that for the majority of the considered homotypic adhesion pairs, the aggregates behaved as expected. The exception to this is the pair with the smallest magnitude of difference (see fig. 4.24 & 4.25). In this case, the CBI results show that the change in the configuration of the aggregates is quite subdued. Furthermore, when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion is low, the aggregates seem to depart very little from the configurations they begin with. This difference in performance is highlighted by the summary figures (see figures 4.26, 4.27, 4.28 and 4.29). These figures show the CBI results for each homotypic adhesion pair with the following $\mathbf{a} \rightarrow \mathbf{b}$ adhesion values:

- $W_{bb} + ((W_{aa} - W_{bb})/4)$ where the aggregates should form the onion configuration.
- $W_{bb} + 3((W_{aa} - W_{bb})/4)$ where the aggregates should become intermixed.

It can be seen that the difference in performance is particularly clear in the cases where the aggregates should form the sorted onion configuration.

The CT results for the simulations are shown in figure 4.34. These results also appear to indicate some link with the model's performance and the difference between

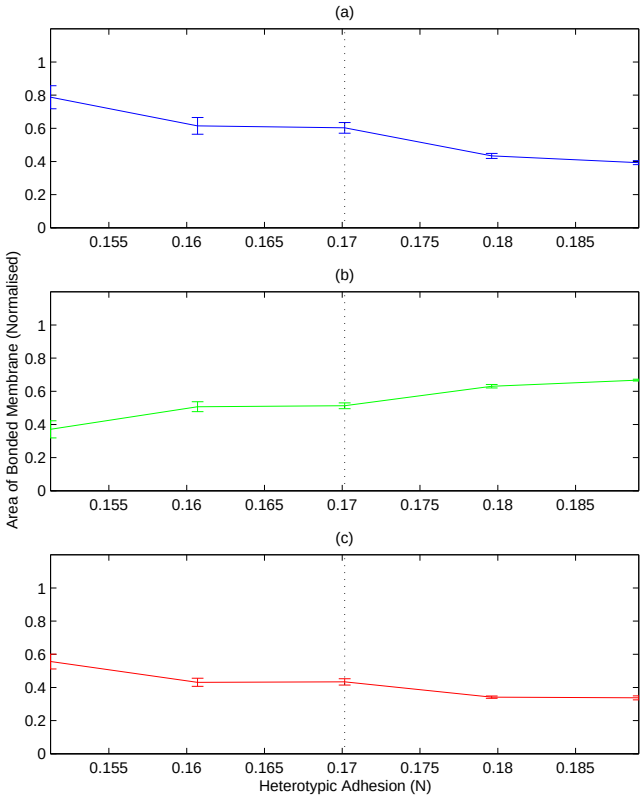


Figure 4.24: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid and* $\mathbf{a} \rightarrow \mathbf{a}$ *and* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesions of* $0.1890625N$ *and* $0.15125N$ *respectively. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

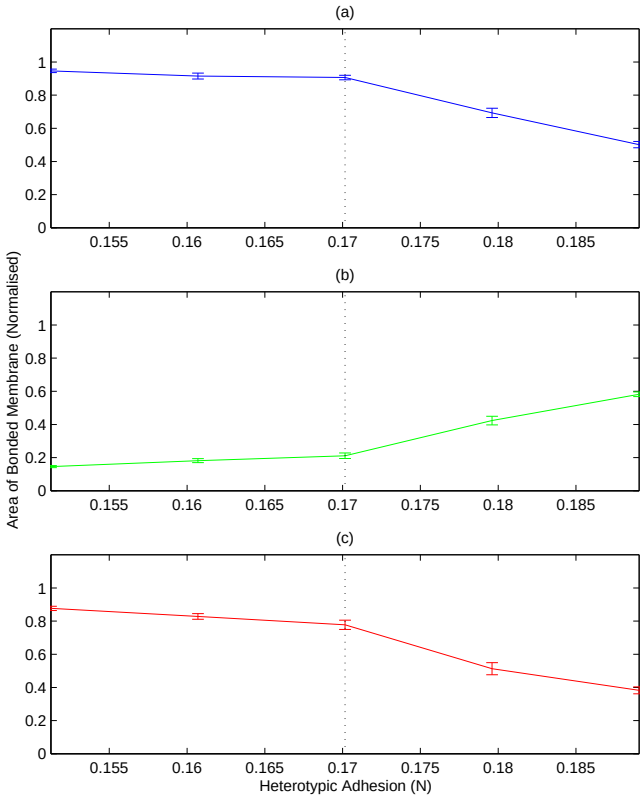


Figure 4.25: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid and* $\mathbf{a} \rightarrow \mathbf{a}$ *and* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesions of* $0.1890625N$ *and* $0.15125N$ *respectively. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

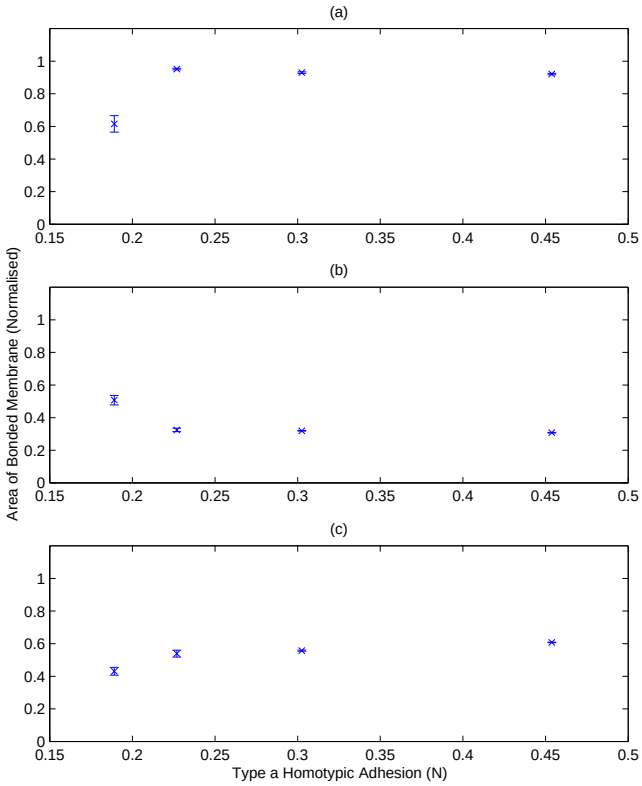


Figure 4.26: CBI *results for 36 A-Cell aggregates which should form the onion configuration from an initially mixed grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

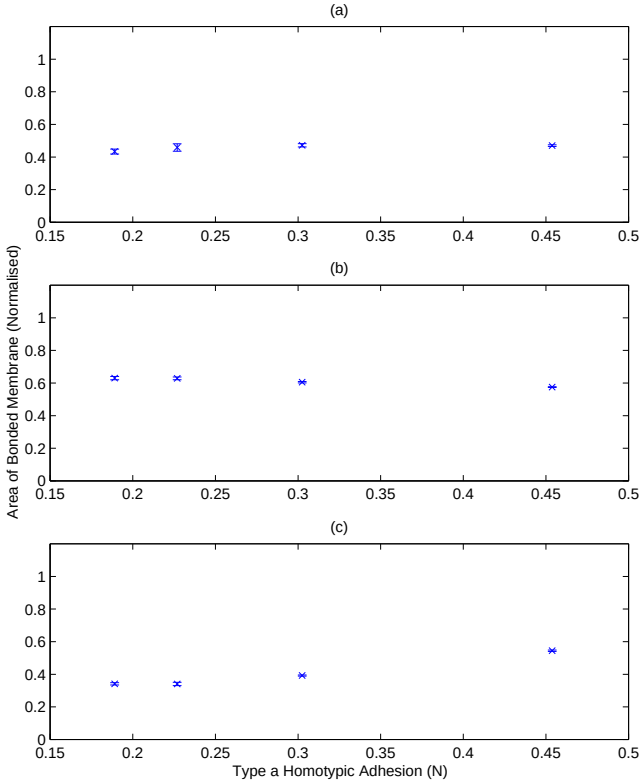


Figure 4.27: CBI *results for 36 A-Cell aggregates which should form an inter-mixed configuration from an initially mixed grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

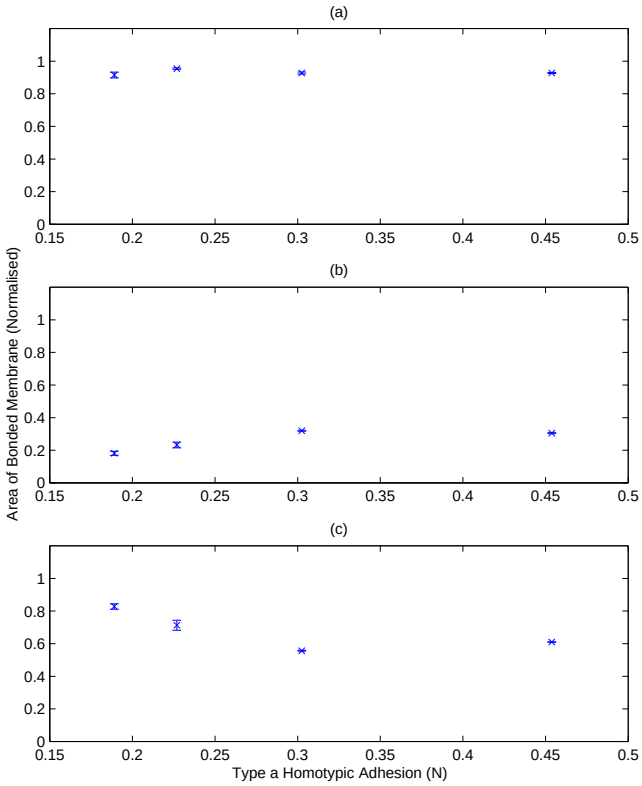


Figure 4.28: CBI *results for 36 A-Cell aggregates which should form the onion configuration from an initially sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

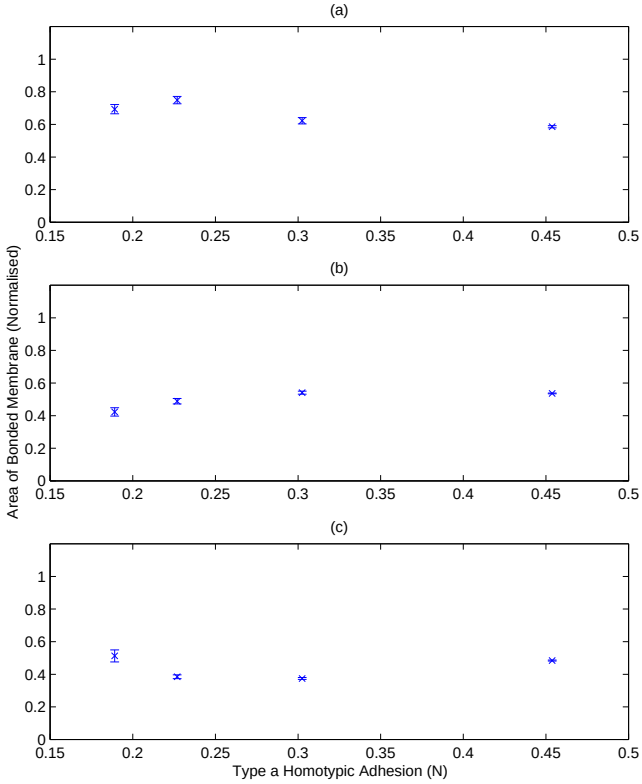


Figure 4.29: CBI *results for 36 A-Cell aggregates which should form an inter-mixed configuration from an initially sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

the homotypic adhesion, and in each case there is a noticeable drop as the magnitude of the difference increases.

As proposed in section 4.1.5.1, it is possible that instead of relating to the difference in the magnitude of the homotypic adhesions, these observations might simply be a by-product of the decreasing equilibrium distance between $\mathbf{a} \rightarrow \mathbf{a}$ and $\mathbf{a} \rightarrow \mathbf{b}$ combinations of A-Cells. To control for this, the experiment was repeated. However, this time the $\mathbf{a} \rightarrow \mathbf{a}$ adhesion was kept constant at $0.45375N$ while the $\mathbf{b} \rightarrow \mathbf{b}$ adhesion was varied as follows:

1. $\mathbf{b} \rightarrow \mathbf{b} = 0.15125N$
2. $\mathbf{b} \rightarrow \mathbf{b} = 0.3025N$
3. $\mathbf{b} \rightarrow \mathbf{b} = 0.378125N$
4. $\mathbf{b} \rightarrow \mathbf{b} = 0.4159375N$

In this way, the same four magnitudes of difference in the homotypic adhesions were considered. However, in contrast to the previous run, the pair with the smallest magnitude of difference have the smallest $\mathbf{a} \rightarrow \mathbf{a}$ and $\mathbf{a} \rightarrow \mathbf{b}$ equilibrium distances. The CBI results for these experiments (see figures 4.30, 4.31, 4.32 and 4.33) show that despite this the homotypic adhesion pair with the smallest magnitude of difference still results in the poor performance. Again, this is most notable in the two cases where the aggregates should form the onion configuration. However, in the CT results (see fig. 4.35), excluding the values for the homotypic adhesion pair where the aggregates failed to reach the desired final configurations, the increase is significantly reduced. This suggests that although the increase in the magnitude of difference between the homotypic adhesions causes some increase in the CT , the effects are reduced by a general improvement in performance caused by greater aggregate fluidity resulting from the general reduction in the equilibrium distances[3]. This increase in the fluidity of the aggregates is reflected in the CBI where aggregates starting from the initial configuration of a sorted grid should achieve the onion configuration. In the original experiments, the more rigid aggregates remained very segregated. However, in the second

[3]This matches the earlier rounding results which show that aggregates with smaller equilibrium distances are more fluid (see section 3.4.3).

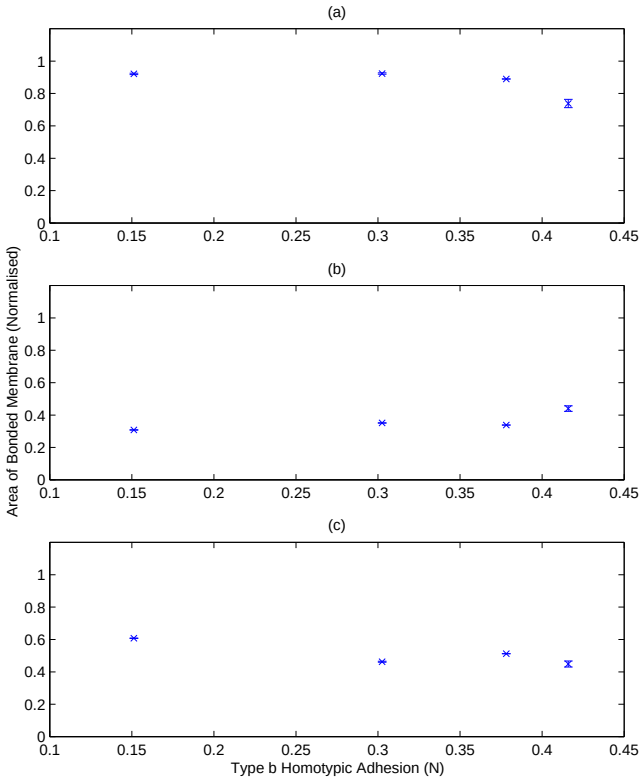


Figure 4.30: CBI *results for 36 A-Cell aggregates which should form the onion configuration from an initially mixed grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion of each pair is fixed at* $0.45375N$ *while the* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions decreases from left to right.*

set of experiments, although the adhesive differences still prevented the aggregate from achieving the expected configuration there is clearly much more interaction between the two A-Cell types as indicated by the higher $\mathbf{a} \rightarrow \mathbf{b}$ and lower $\mathbf{a} \rightarrow \mathbf{a}$ CBIs .

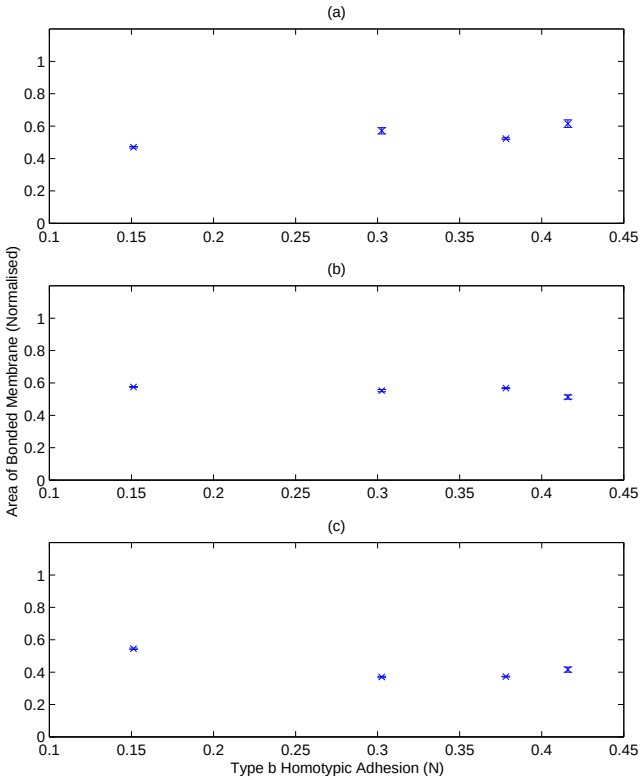


Figure 4.31: CBI *results for 36 A-Cell aggregates which should form an intermixed configuration from an initially mixed grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion of each pair is fixed at* $0.45375N$ *while the* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions decreases from left to right.*

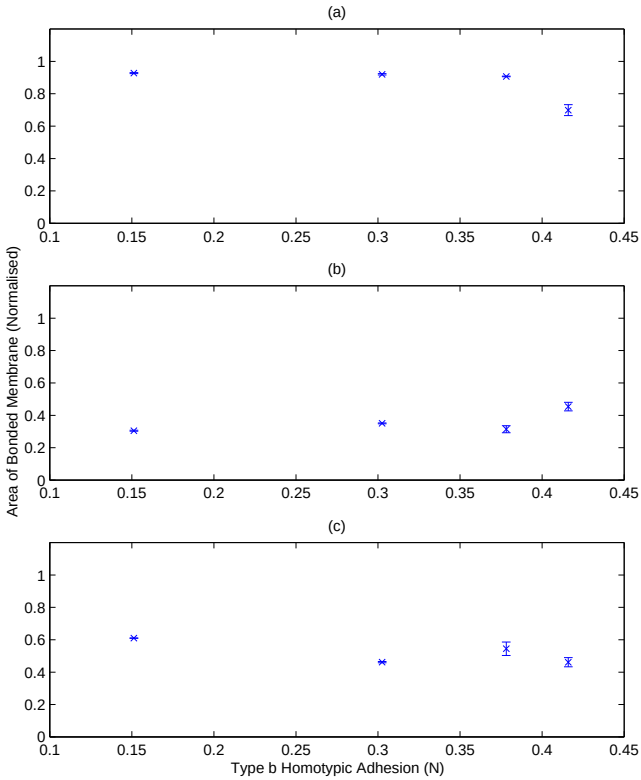


Figure 4.32: CBI *results for 36 A-Cell aggregates which should form the onion configuration from an initially sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion of each pair is fixed at* $0.45375N$ *while the* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions decreases from left to right.*

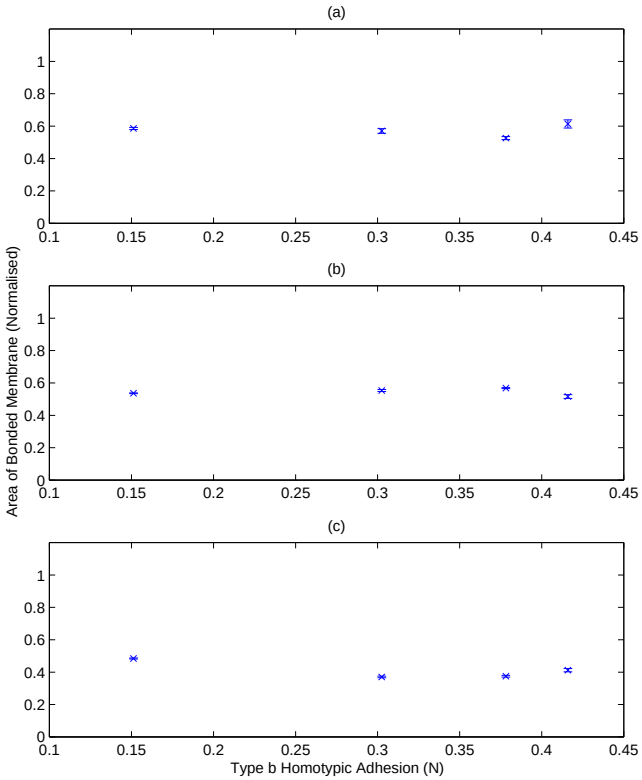


Figure 4.33: CBI *results for 36 A-Cell aggregates which should form an intermixed configuration from an initially sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion of each pair is fixed at* $0.45375N$ *while the* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions decreases from left to right.*

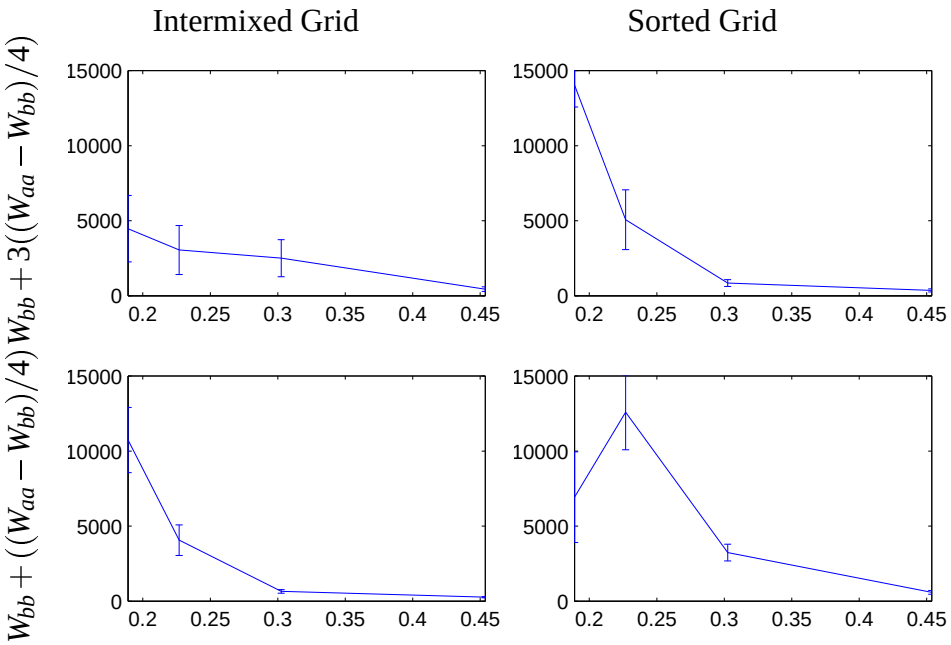


Figure 4.34: CT *results for 36 A-Cell aggregates starting from both an initially mixed grid and a sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

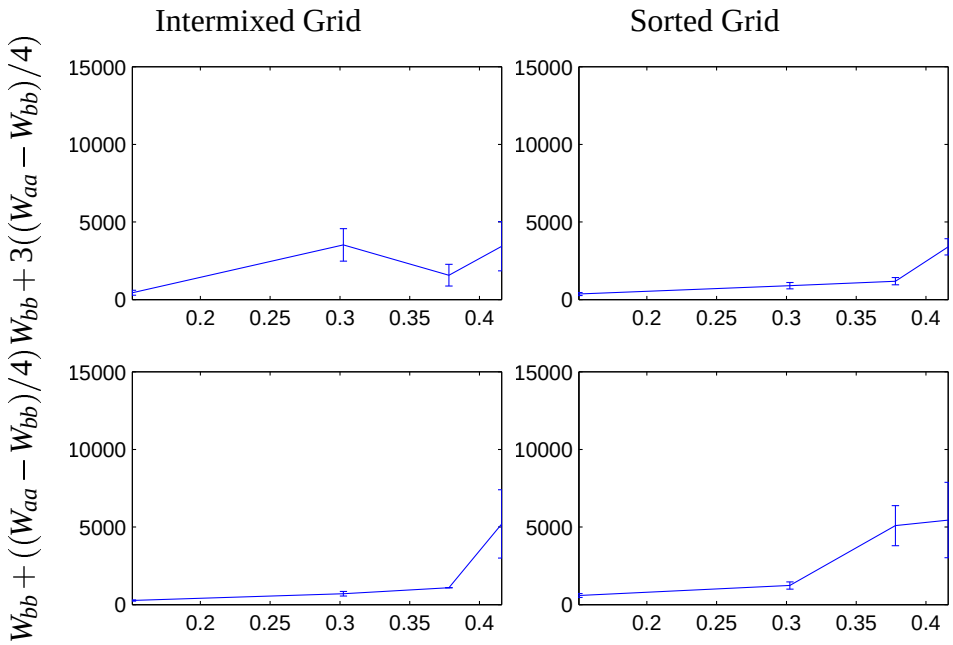


Figure 4.35: CT *results for 36 A-Cell aggregates starting from both an initially mixed grid and a sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion of each pair is fixed at* $0.45375N$ *while the* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions decreases from left to right.*

4.3 Random Forces

The following experiments investigate how small random forces, either actuated or environmental, affect the sorting and mixing ability of the model. The experiments which were carried out follow the same structure as those described in section 3.3. The first two sets of experiments investigate the effects of random movements and random torques independently of each other, while the last set considers their combined effect.

In all the described experiments the homotypic adhesion were kept constant (**a** $\rightarrow$ **a** $0.3025N$, **b** $\rightarrow$ **b** 0.15125N see 4.1.5). However, each set of experiments was carried out starting from both intermixed and sorted grid configurations and in each of these instances the two default heterotypic adhesions were used to investigate both sorting and intermixing behaviours (see section 4.1.5). The simulations were also carried out for the two aggregate sizes consisting of 36 and 64 A-Cells. However, as these results show a very high correspondence, only the results for the 36 A-Cell aggregates are presented here. The remaining results can be found in appendix B.

Figures 4.36, 4.37, 4.38 and 4.39 show the results obtained for the aggregates with only random movement forces acting on the A-Cells. The CBI results show that, as expected, an increase in the size of the random forces does improve the performance of the model, with the aggregates getting closer to the expected equilibrium states. This is particularly clear for the transitions which require greater levels of rearrangement; intermixed grid $\rightarrow$ 'onion' (see fig. 4.36) and sorted grid $\rightarrow$ intermixed (see fig. 4.39). However, in most cases, the CBIs only get close to the expected levels when the random force limit is at the tested maximum of $0.05N$.

Conversely, the CT results for these experiments show the opposite of the expected trend and in the cases highlighted above, the CT actually rises as the random force limit increases. Comparing these results with the CBIs reveals that the rises correspond closely with improvements in the model's performance. The reason for this correspondence is that the regular alignment of the A-Cells' sensors (as there are no torque forces) makes the aggregates more rigid. Therefore, when the random movement force limits are low, the aggregates very quickly become trapped in poor configurations. Increasing the size of the limits means that through increased exploration the aggregates can eventually break out of some of these configurations and move closer to the equi-

librium state before becoming trapped again. However, as a result of this, it takes longer for the CBIs to converge. Once the size of the random movement force limits is high enough for the aggregate to actually reach the equilibrium state, it can be seen that further increases do result in the expected decrease in CT (see fig. 4.39).

The results from the random torque experiments are shown in 4.40, 4.41, 4.42 and 4.43. It is clear that although the presence of some random torque is beneficial, the actual level of torque has no observable effect on the performance of the model. In addition, in the majority of the cases it appears that only torque is sufficient to allow the aggregates to reach the final equilibrium state. The exceptions to this are the simulations of aggregates which should reach an intermixed equilibrium state from an initially sorted grid (see fig. 4.43). In these cases the aggregates only partially mix and become trapped in poor configurations.

The final set of results, which investigate the combined effect of random torque in conjunction with random movement forces are shown in figures 4.44, 4.45, 4.46, and 4.47. As in the similar rounding experiments the random torque limit was fixed at $0.01N$ (see section 3.3). Again, the CBI results show no noticeable effect and remain at the same high level. However, in this case it is likely that this is simply because the aggregates have reached the equilibrium states. Therefore, it is not possible for them to improve further. The CT results, on the other hand, tend to show a drop as the random movement force is increased. The only exception is for the 36 A-Cell aggregates starting from an intermixed grid which should achieve an intermixed equilibrium state (see fig. 4.42) where the CT has no clear trend. However, this is most likely because very little rearrangement is required for the aggregate to obtain the final configuration. Thus, any improvements in CT will be very small and may be obscured by noise.

All of these results match those obtained for earlier rounding experiments (see section 3.3) and suggest that the increased exploration caused by small random forces improves the model's performance. In addition, the results again emphasise the importance of random torque and show that aggregates where the A-Cells remain aligned are far more likely to become trapped in poor configurations, preventing them from reaching the desired equilibrium state.

In contrast to the earlier rounding results, none of the simulations that were carried

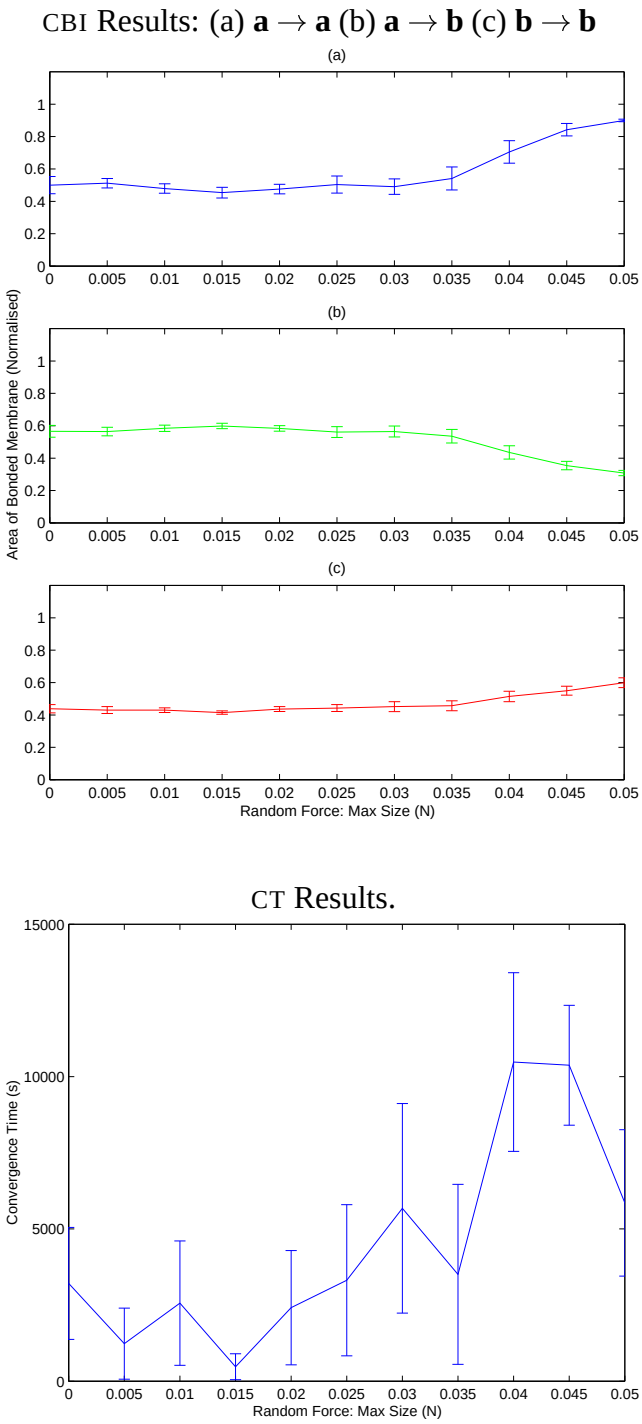


Figure 4.36: *Results for 36 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the onion equilibrium state. The results show that an increase in the magnitude of the random forces causes a significant improvement in the aggregate's ability to sort.*

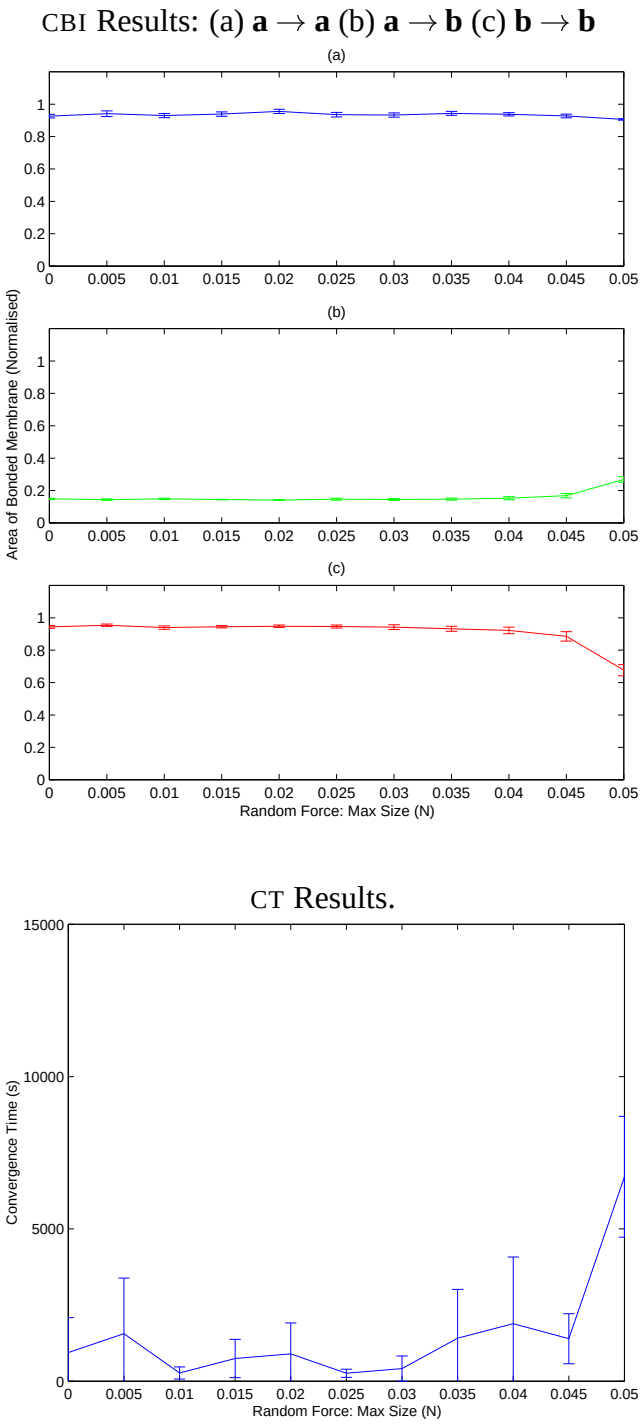


Figure 4.37: *Results for 36 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the onion equilibrium state. The results show that an increase in the magnitude of the random forces has little effect on the engulfment process with the* CBI *values only improving slightly when the random force limit is at the top of the range.*

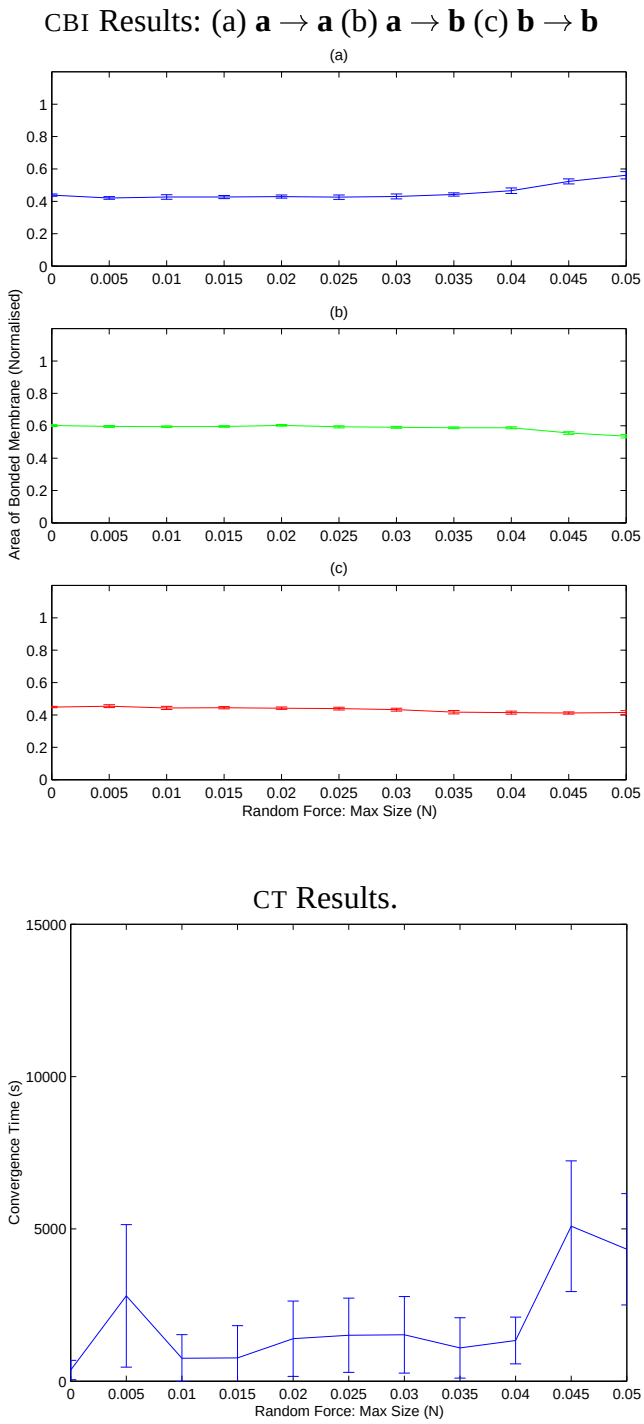


Figure 4.38: *Results for 36 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the intermixed equilibrium state. The results show that the magnitude of the random forces has little effect on the aggregate's ability to remain intermixed.*

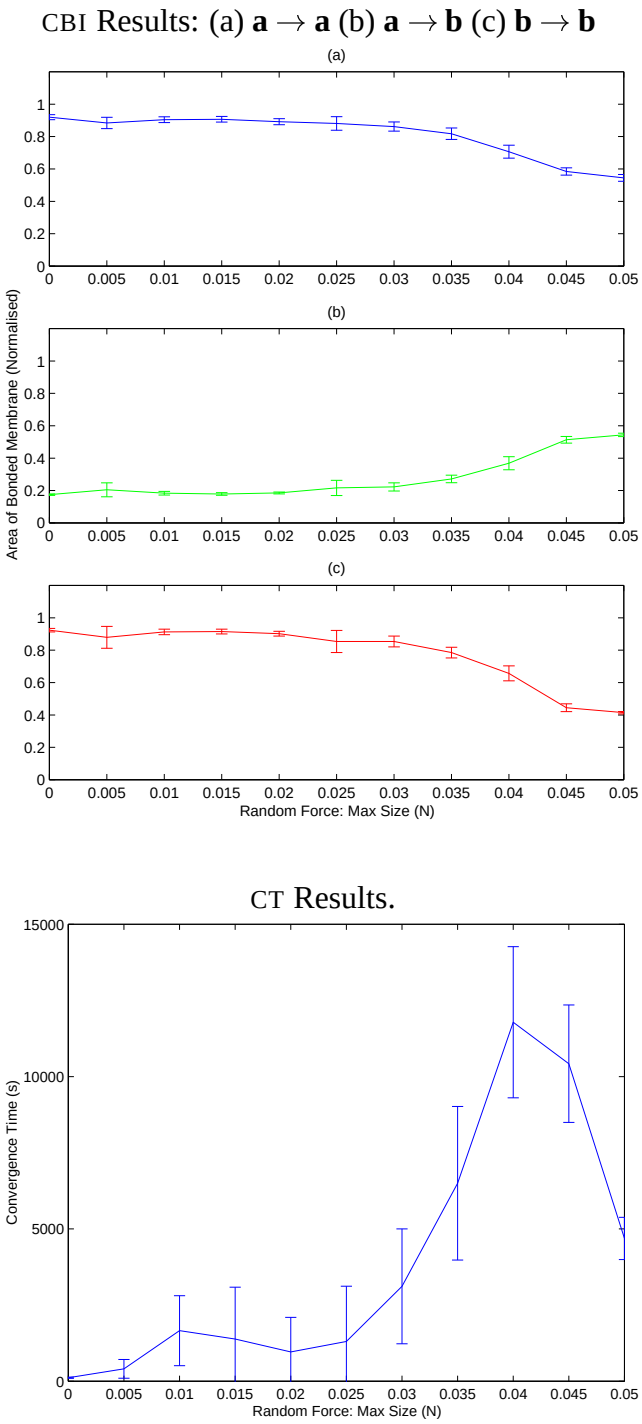


Figure 4.39: *Results for 36 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the intermixed equilibrium state. The results show that an increase in random exploration causes a significant improvement in the aggregate's ability to increase the level of intermixing. For low random force limits the aggregate remains sorted.*

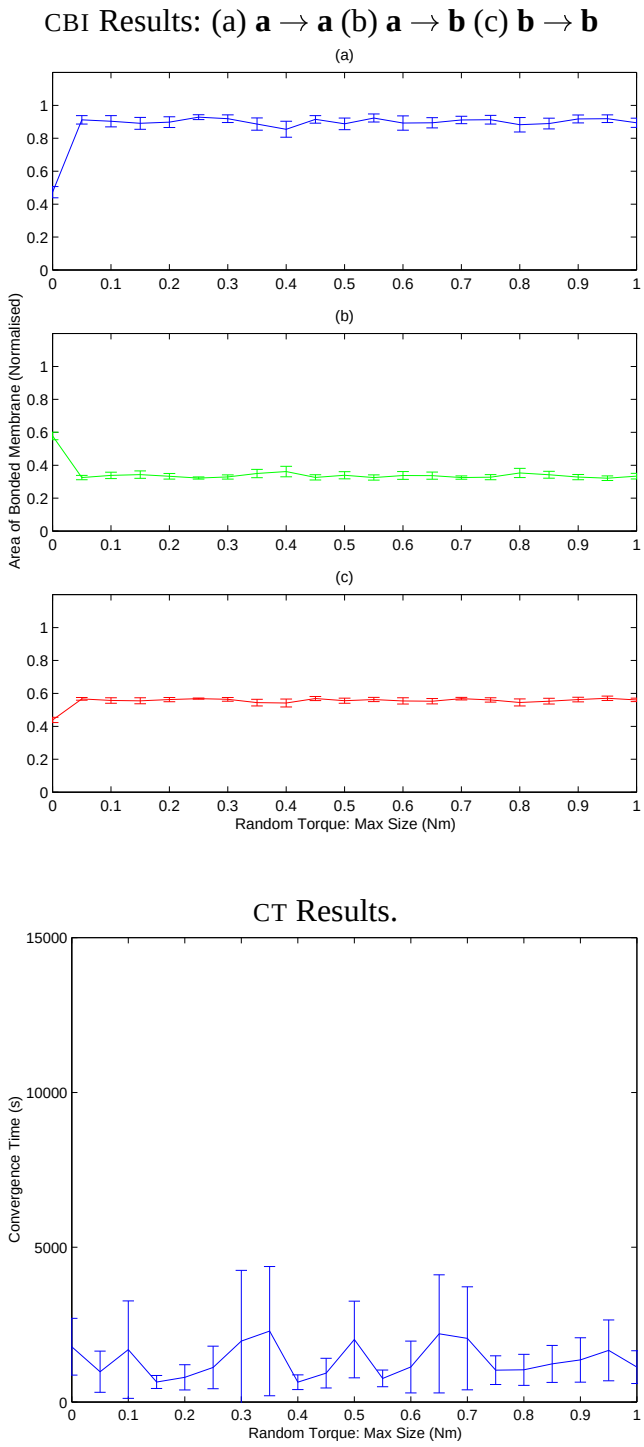


Figure 4.40: *Results for 36 A-Cell aggregates over a range of random torque limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the onion equilibrium state. The results show that any amount of random torque is sufficient to allow the aggregate to fully sort.*

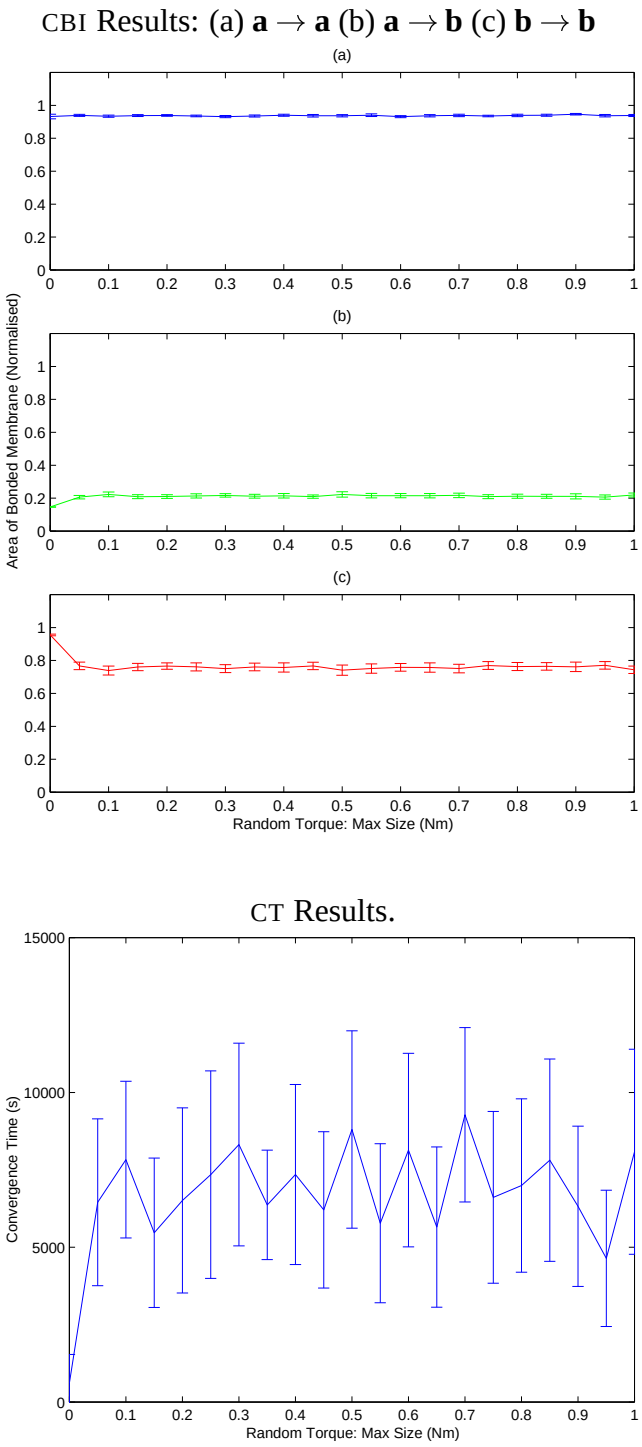


Figure 4.41: *Results for 36 A-Cell aggregates over a range of random torque limits. The aggregates started from the initial configuration of a sorted grid and should move toward the onion equilibrium state. The results show that any amount of random torque is sufficient to allow the aggregate to sort.*

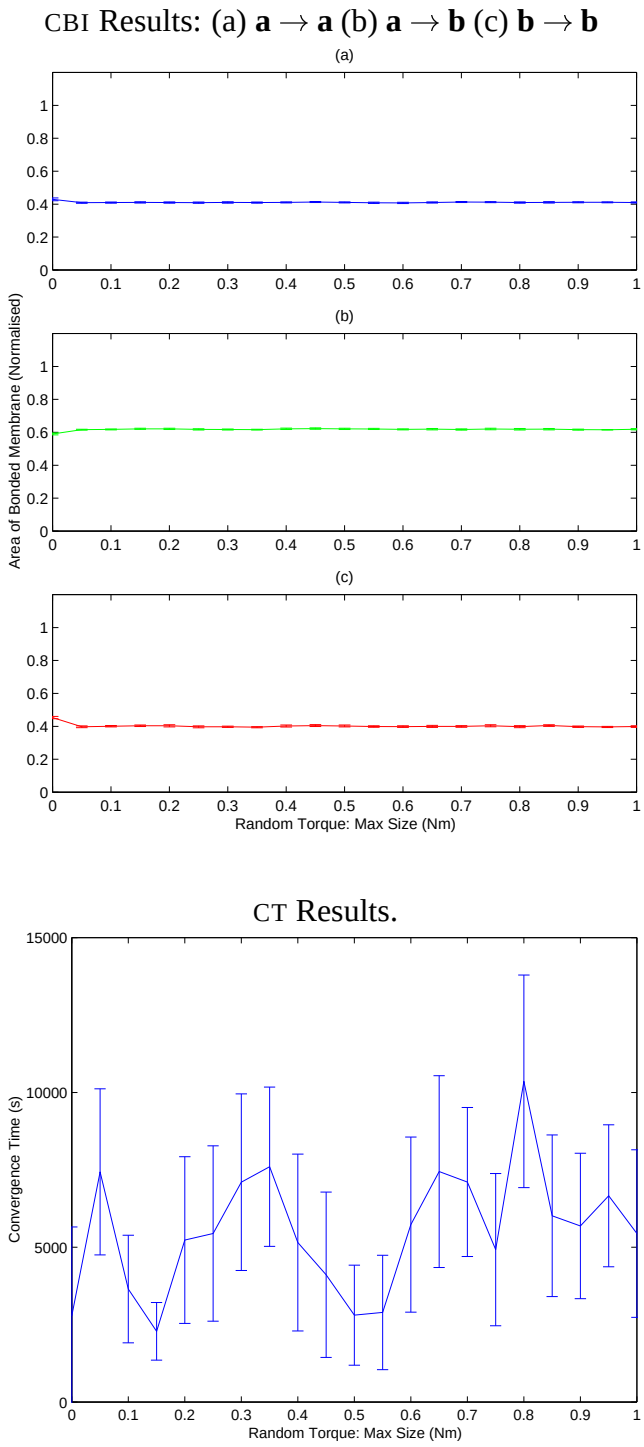


Figure 4.42: *Results for 36 A-Cell aggregates over a range of random torque limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the intermixed equilibrium state. The results show that the presence of some random torque has a very small effect on the results. However, the aggregate is able to remain intermixed over the whole range of values.*

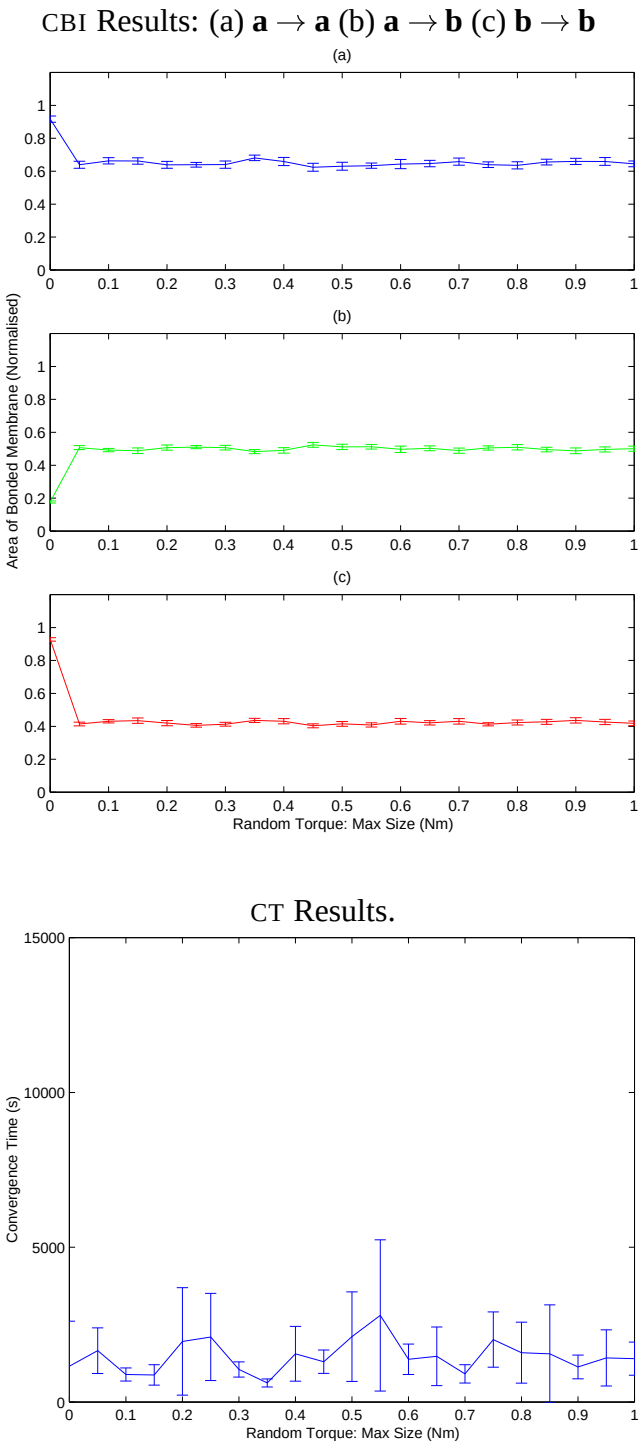


Figure 4.43: *Results for 36 A-Cell aggregates over a range of random torque limits. The aggregates started from the initial configuration of a sorted grid and should move toward the intermixed equilibrium state. The results show that the presence of random torque allows some level of sorting to take place. However, despite the magnitude of the random torques the aggregate fails to reach the fully sorted configuration.*

out resulted in A-Cells breaking away from the aggregate. This is despite the fact that the weakest homotypic adhesion matches the adhesion between the A-Cells in the earlier rounding simulations where random movement force limits greater than $0.04N$ resulted in aggregate splits. However, this is simply because the A-Cells with the low homotypic adhesion have additional heterotypic adhesions which are strong enough to prevent them from breaking away from the main aggregate.

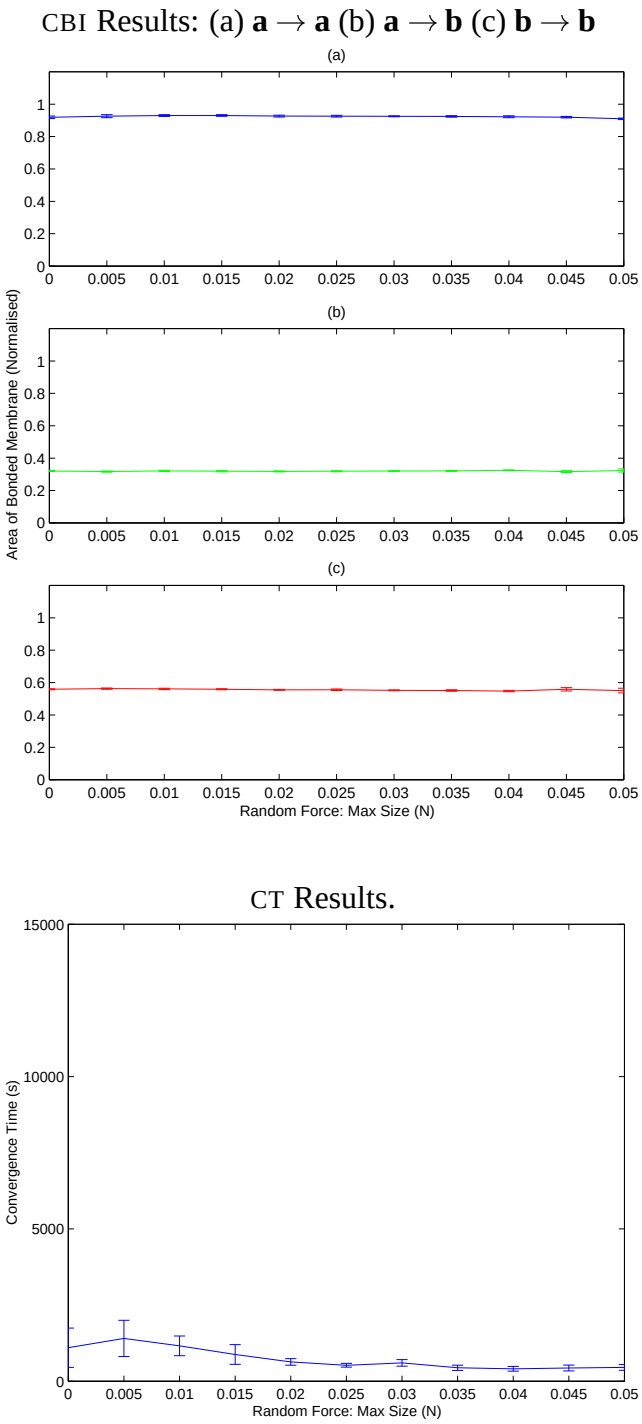


Figure 4.44: *Results for 36 A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the onion equilibrium state. The* CBI *results show that the aggregate reaches the equilibrium configuration in every case. In addition, there is a slight decrease in the* CT *as the magnitude of the random forces increases.*

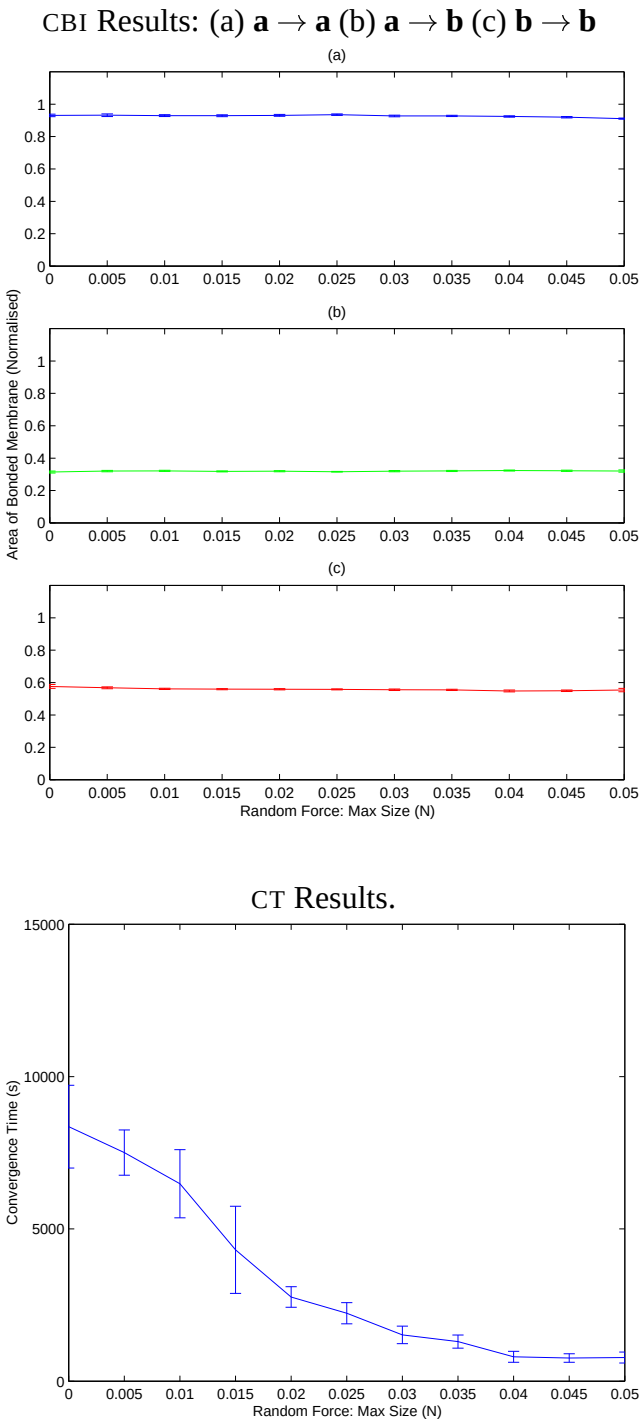


Figure 4.45: *Results for 36 A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the onion equilibrium state. The* CBI *results show that the aggregate reaches the equilibrium configuration in every case. In addition, there is a significant decrease in the* CT *as the magnitude of the random forces increases.*

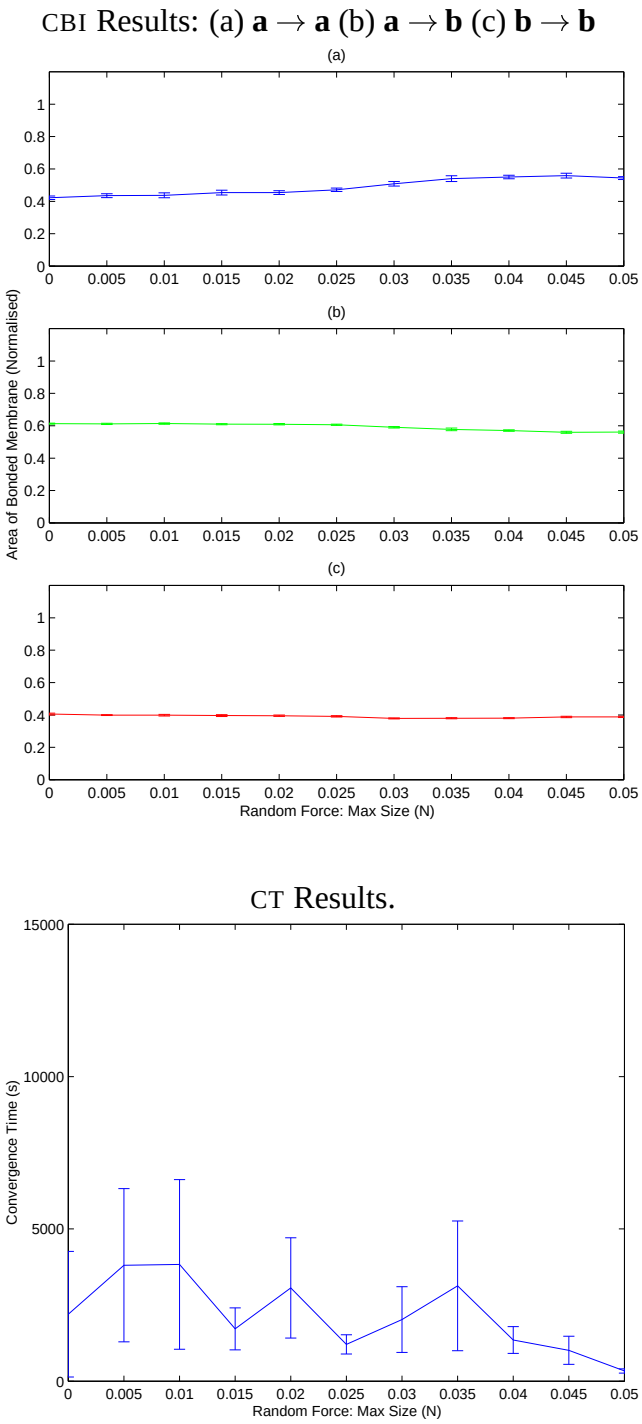


Figure 4.46: *Results for 36 A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the intermixed equilibrium state. The* CBI *results show that the initially intermixed aggregate remains intermixed. In addition, the* CT *shows no clear trend.*

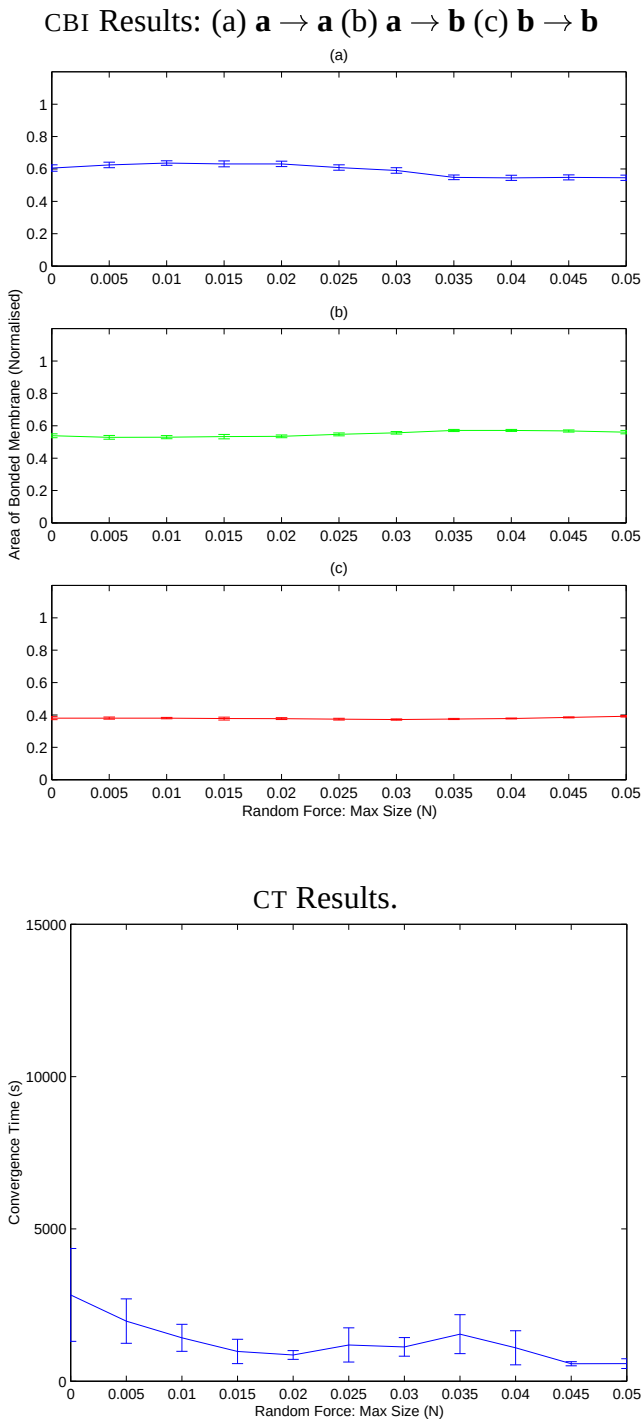


Figure 4.47: *Results for 36 A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the intermixed equilibrium state. The* CBI *results show that the aggregate reaches the equilibrium configuration in every case. In addition, there is a slight decrease in the* CT *as the magnitude of the random forces increases.*

4.4 A-Cell Ratios

Up until now, the experiments discussed have always involved equal numbers of each A-Cell type. However, the ratio of the two A-Cell types should have a great impact on the performance of the model. For example, if the density of one A-Cell type is very low, the chance of any two of those A-Cells coming into contact will be greatly reduced and thus it is less likely that the A-Cells will successfully sort.

The experiments presented in this chapter aim to investigate how the ratio of the A-Cell types affects the overall performance of the model. In total, seven different ratios were considered, ranging from 8:56 to 56:8 such that the total number of A-Cells remained constant at 64. These unbalanced aggregates present an interesting problem of choosing initial configurations which allow the results to be compared fairly. Aggregates with a small number of a particular type of A-Cell have many more configurations in which the initial distance between these A-Cells will be large than aggregates where the same type is in the majority. To minimise this effect, the results should represent the median case where the A-Cells are neither fully sorted or intermixed. This could be achieved by simply simulating a very large number of random configurations for each ratio. If the sample is large enough it should represent the full range of possible configurations, thus giving the mean and expected distribution of results. However, this brute force approach would require considerable resources. To avoid this we have taken an alternative approach where only three small sets of configurations are used to represent the entire range. However, these are carefully selected such that they cover both the extremes and the median case for a particular ratio.

The configurations were selected from the set of all possible configurations of the A-Cell types on an eight by eight grid. 10000 of these configurations were generated at random and, in each case, two indices were calculated from the mean distance between neighbours (the mean distance between A-Cells and their n nearest neighbours of the same type[4]) of each A-Cell type.

Using these measures the configurations were then sorted and five were selected

[4]The actual number of neighbours considered was 6.

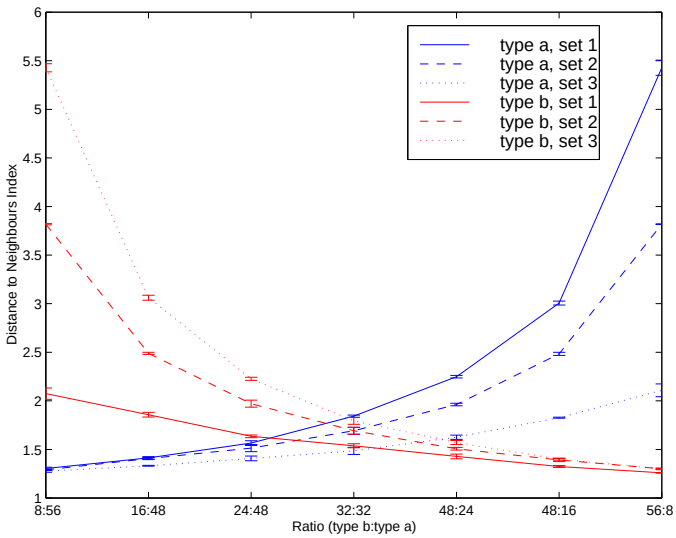


Figure 4.48: *The distance between neighbours measures for each set of initial configurations and each ratio of A-Cells. Set 1 – low type* b*distance between neighbours measure, high type* a*distance between neighbours measure. Set 2 – median between neighbours measures. Set 3 – high type* b*distance between neighbours measure, low type* a*distance between neighbours measure.*

from the extremes and middle of the sorted list [5]. The most obvious way of sorting the configurations is to use a single index. However, this makes the assumption that configurations which represent one extreme for one A-Cell type represent the other extreme for the other A-Cell type and this is not always the case. Therefore, using this approach, twice as many experiments would be required; those in which the configurations are ordered for one measure and those in which the configurations were ordered by the other. To avoid this, a compromise ordering was made by subtracting one measure from the other. This means that although the extreme configurations are not necessarily the most extreme values for either measure, they do generally provide a good approximate for each (see fig. 4.50). The changes in the two measures for the selected configurations are shown in figure 4.48.

Choosing the normalisation values used when calculating the CBIs for these unbalanced aggregates also requires more consideration. The estimated maximum area of bonded membrane for the homotypic bonds can still be determined by simulating the relative numbers of A-Cells with the appropriate adhesion. However, when the A-Cell numbers are unequal, the estimated maximum area of bonded membrane can no longer be determined by simply simulating an aggregate whose size is equivalent to the combined total of A-Cells of each type. Instead, as the area of heterotypic bonded membrane will be limited by the smallest population of A-Cells out of the two types, it makes more sense to use double this value. The drawback with this approach is that it ignores any bonds that may be formed with the excess of A-Cells from the larger population (see fig. 4.51). This underestimate of the true maximum means that the heterotypic CBI will appear higher in such cases. However, it should still provide a reasonable approximation.

The results for the ratio experiments where the equilibrium state is the onion configuration are shown in 4.54, 4.55, and 4.56. In this case the following trends in the CBIs should be observed.

1. The type *a*A-Cells should always form the central aggregate. Therefore, the $\mathbf{a} \rightarrow \mathbf{a}$ CBI should remain constant at a value close to 1 (e.g. the A-Cells always

[5] In practice, the selected configurations were also checked to verify that they contained no mirrored or rotated copies.

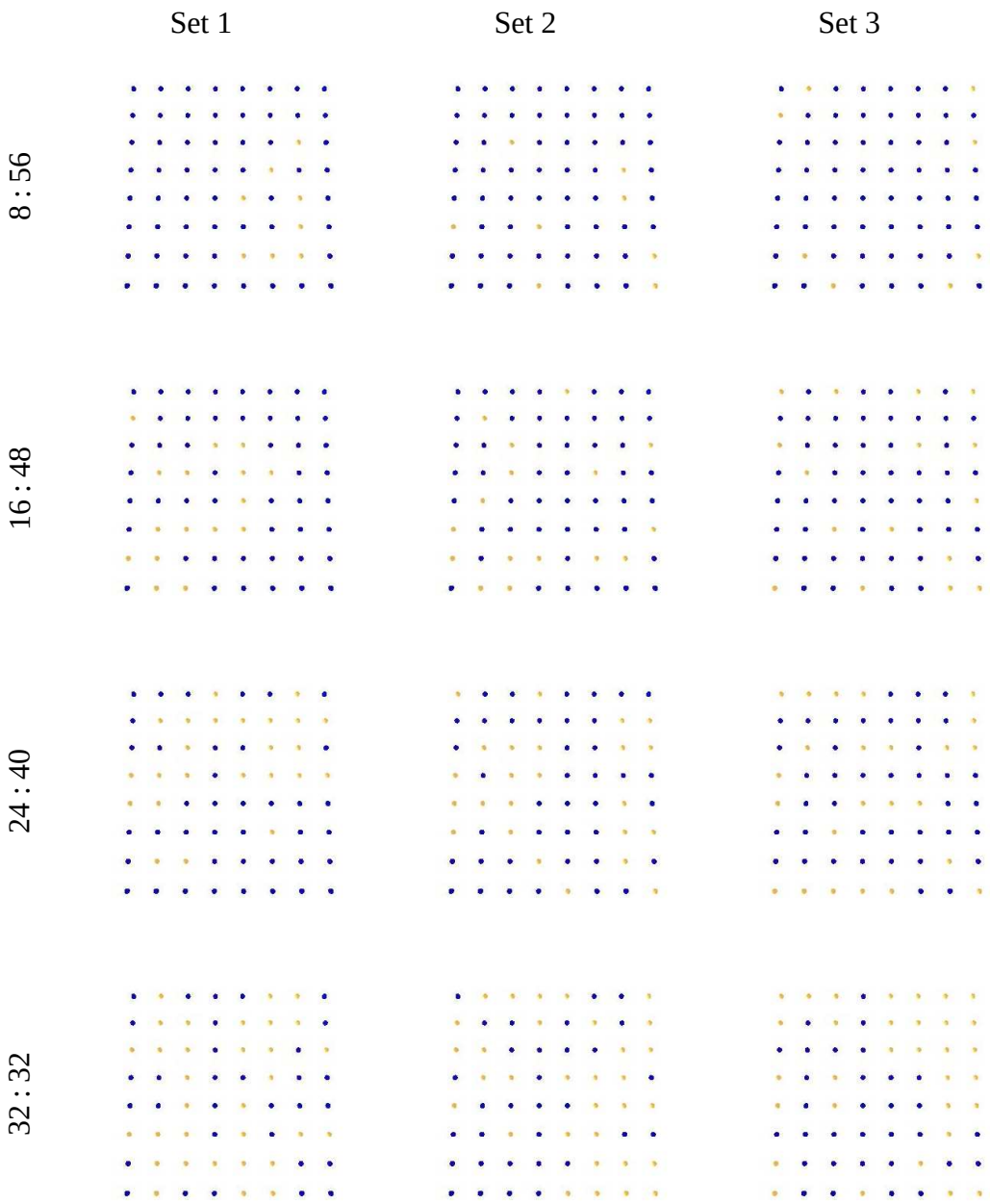


Figure 4.49: *Screen shots showing example initial configurations from the three chosen sets and for different A-Cell ratios (type* b*:type* a*). Set 1 – low type* b*distance to neighbours index, high type* a*distance to neighbours index. Set 2 – median distance to neighbours indices. Set 3 – high type* b*distance to neighbours index, low type* a*distance to neighbours index.*

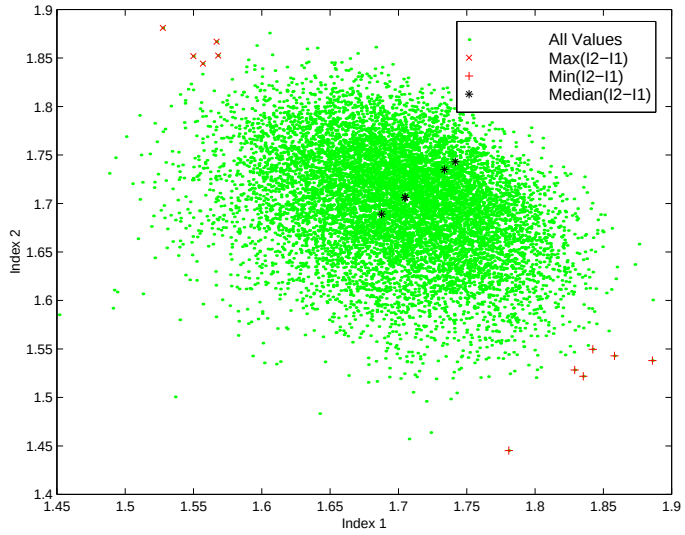


Figure 4.50: *Neighbourhood indices calculated for 10000 random grid configurations consisting of 32 two A-Cells of one A-Cell type and 32 A-Cells of the another. The points for the 15 selected configurations representing the extreme and median values of* $Index2 - Index1$ *are highlighted*

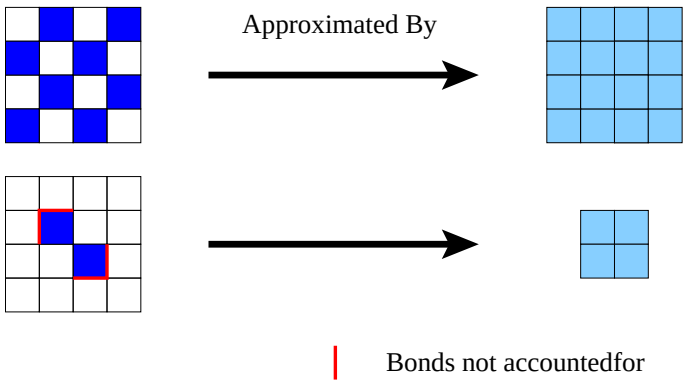


Figure 4.51: *Aggregate sizes used to approximate the area of heterotypic bonded membrane in an aggregate consisting of two A-Cell types.*

maximise their $\mathbf{a} \rightarrow \mathbf{a}$ adhesions).

2. When the ratio of type *b*A-Cells is too low for them to completely surround the central aggregate, they should evenly distribute themselves around its surface. This will result in a reduction in the $\mathbf{b} \rightarrow \mathbf{b}$ CBI and a corresponding increase in the $\mathbf{a} \rightarrow \mathbf{b}$ CBI .

3. As the ratio of type *a*A-Cells drops, the central aggregate will decrease in size. Therefore, the shell the type *b*A-Cells form will increasingly approximate a solid disc where the $\mathbf{b} \rightarrow \mathbf{b}$ adhesions are maximised. This will result in a steady increase in the $\mathbf{b} \rightarrow \mathbf{b}$ CBI .

The actual results largely match these predictions. However, there is a notable exception, with the $\mathbf{a} \rightarrow \mathbf{a}$ CBI actually showing a significant drop in all three sets of results once the ratio of type *a*A-Cells drops below fifty percent. In each case the results show that this drop corresponds with an equivalent rise in the $\mathbf{a} \rightarrow \mathbf{b}$ CBI . Observation of the simulations revealed that in those simulations with a predominance of the weaker type *b*A-Cells, the position of the type *a*A-Cells within the aggregate was reasonably fixed. However, in the aggregates with the converse distribution, the type *b*A-Cells wandered over far greater distances. This matches earlier results which demonstrate that aggregates with smaller equilibrium distances between the A-Cells demonstrate a greater degree of fluidity (see section 3.4.3) and implies that the lack of exploration prevents the type *a*A-Cells from fully clustering together. This also explains the difference in the observed change in the $\mathbf{a} \rightarrow \mathbf{a}$ CBI between the three sets of results. In the set of results where the index for the distance between type *a*A-Cells was maximised, the drop is much greater as the type *a*A-Cells must travel over a greater distance before they will encounter each other. However, at the other extreme, where the index for the distance between type a A-Cells was minimised, the drop is much more subtle.

In the CT results for this set of experiments, the main trend should be an approximate ‘V’ shape with the CT increasing as either A-Cell type becomes increasingly sparse. The reason for this is that the minority A-Cells will have a larger number of the other A-Cell type to migrate through when reaching their final destination (either the

centre of the aggregate or the perimeter). However, the actual results show a definite skew, with a much higher increase when the ratio of the stronger adhering type *a*A-Cells drops. This matches the observations from the CBI results which suggested that in such situations the aggregates are more rigid, and suggests that it results in slower aggregation of the type *a*A-Cells. There is one divergent value where the type *a*A-Cells are most sparsely distributed and their distance between neighbours index was maximised. However, this is a result of the A-Cells frequently failing to cluster at all. Therefore, the CBIs remain quite constant and the CT appears low.

The results for the other set of experiments where the equilibrium state is the intermixed configuration are shown in figures 4.57, 4.58, 4.59. In this case, the CBIs should show a clear relationship between the ratio of an A-Cell type and the amount of homotypic adhesion it forms. The reason for this is that the A-Cells in the minority will be surrounded, to some degree, by the excess of A-Cells from the other type as they attempt to maximise the heterotypic adhesions. This should result in contrasting slopes in the homotypic CBIs and an $\mathbf{a} \rightarrow \mathbf{b}$ CBI which is symmetrical about the middle ratio (32:32).

Once more, the actual results match the prediction very closely. However, the extreme ratios of the type *b*A-Cells (8:56 and 16:48) result in a lower $\mathbf{a} \rightarrow \mathbf{b}$ CBI than expected. The screenshots of the final aggregate configurations for the four most unbalanced aggregates reveal why this is the case (see fig. 4.52). In the aggregates where the type *b*A-Cells are in the minority, the slight preference of the $\mathbf{a} \rightarrow \mathbf{a}$ adhesions over the $\mathbf{a} \rightarrow \mathbf{b}$ adhesions results in a slow migration of the type *b*A-Cells to the perimeter where they become evenly distributed. Clearly this results is less $\mathbf{a} \rightarrow \mathbf{b}$ adhesion than the aggregate could potentially form. In the converse ratios, the preference of the $\mathbf{a} \rightarrow \mathbf{b}$ adhesions over the $\mathbf{b} \rightarrow \mathbf{b}$ adhesions prevents this from happening. Therefore the $\mathbf{a} \rightarrow \mathbf{b}$ adhesions are at their expected levels.

The CT results for these experiments should be affected by the level of intermixing in the initial aggregate (aggregates which are initially more intermixed should have a lower CT). Therefore, they should follow similar trends as the distance-between-neighbours indices for the A-Cell type which is in the minority (see fig. 4.53). This suggests that in each of the extreme cases, the CT should decrease rapidly as the ra-

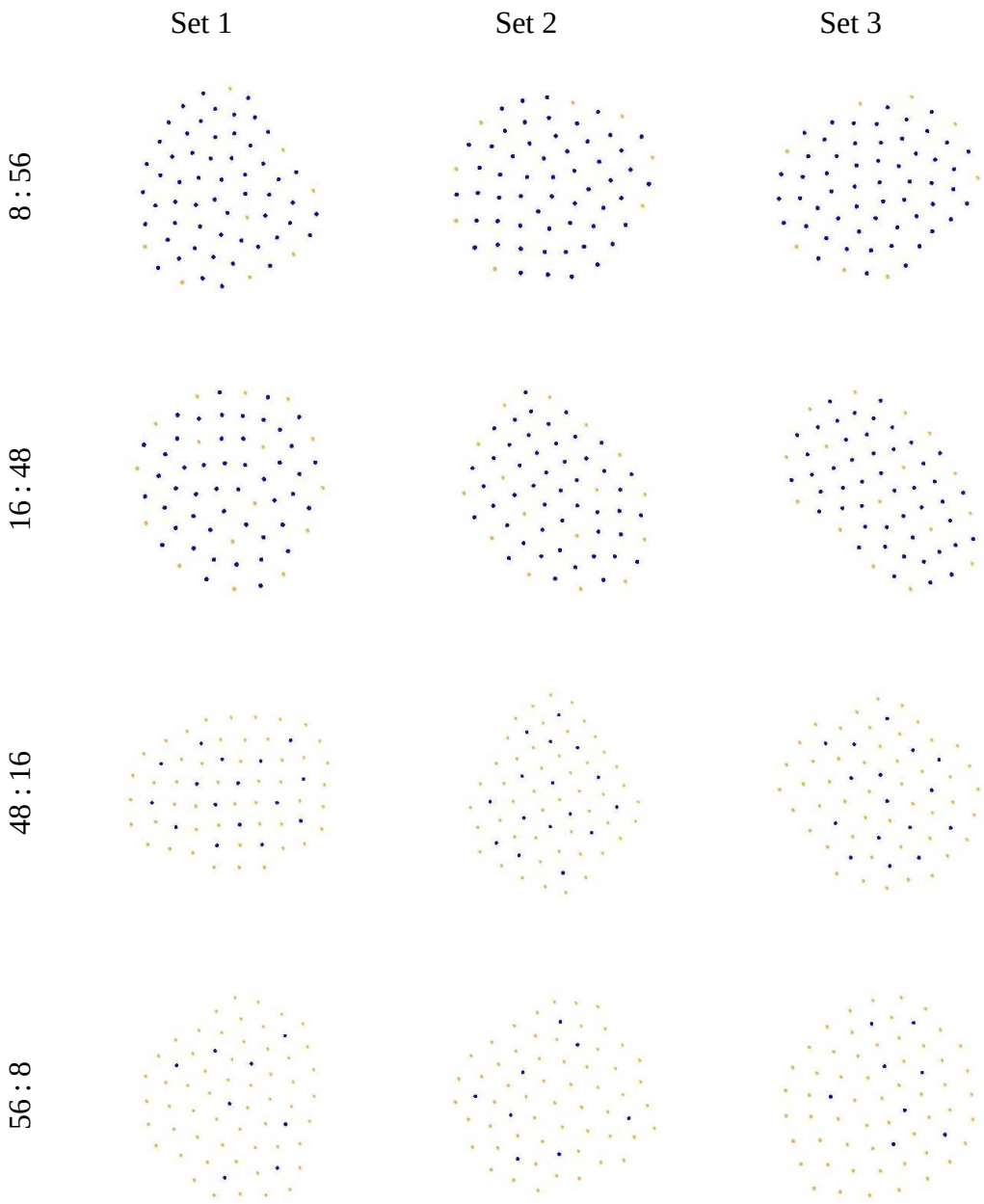


Figure 4.52: *Screen shots showing the final configurations from the three chosen configuration sets and for different A-Cell ratios (type* b*:type* a*). Set 1 – low type* b*distance to neighbours index, high type* a*distance to neighbours index. Set 2 – median distance to neighbours indices. Set 3 – high type* b*distance to neighbours index, low type* a*distance to neighbours index.*

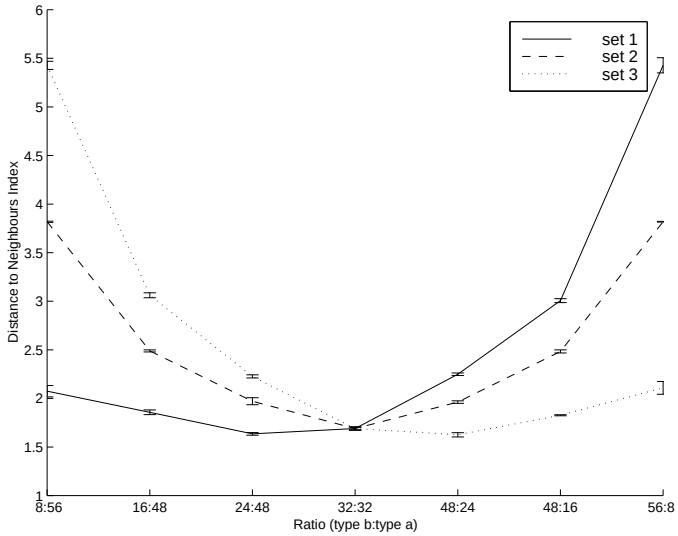


Figure 4.53: *The distance between neighbours measure of the A-Cells in the minority of each ratio. The vales for the 32:32 ratio are given as the average of the two measures. Set 1 – low type* b*distance between neighbours measure, high type* a*distance between neighbours measure. Set 2 – median between neighbours measures. Set 3 – high type* b*distance between neighbours measure, low type* a*distance between neighbours measure.*

tio of A-Cells with the high distance between neighbours measure drops below fifty percent. In addition, when the ratio of the same A-Cells increases above seventy five percent, there should also be a very slight decrease in CT . However, the CTs for the median configurations should be symmetrical about the 32:32 ratio and decrease as the aggregates become more unbalanced.

In the actual results these trends are not very clear. However, it is likely that they have been obscured, to some extent, by the effects which have already been highlighted. For example, when the type *b*A-Cells are in the minority, their migration to the perimeter will increase the CT by an amount relating to their initial distance from it. Therefore, in the simulations where the distance to neighbours index is higher (i.e. the A-Cells are further apart and thus more likely to be closer to the perimeter), this effect will be worse. When this is not the case, these sparse aggregates should be able to reach an equilibrium very quickly.

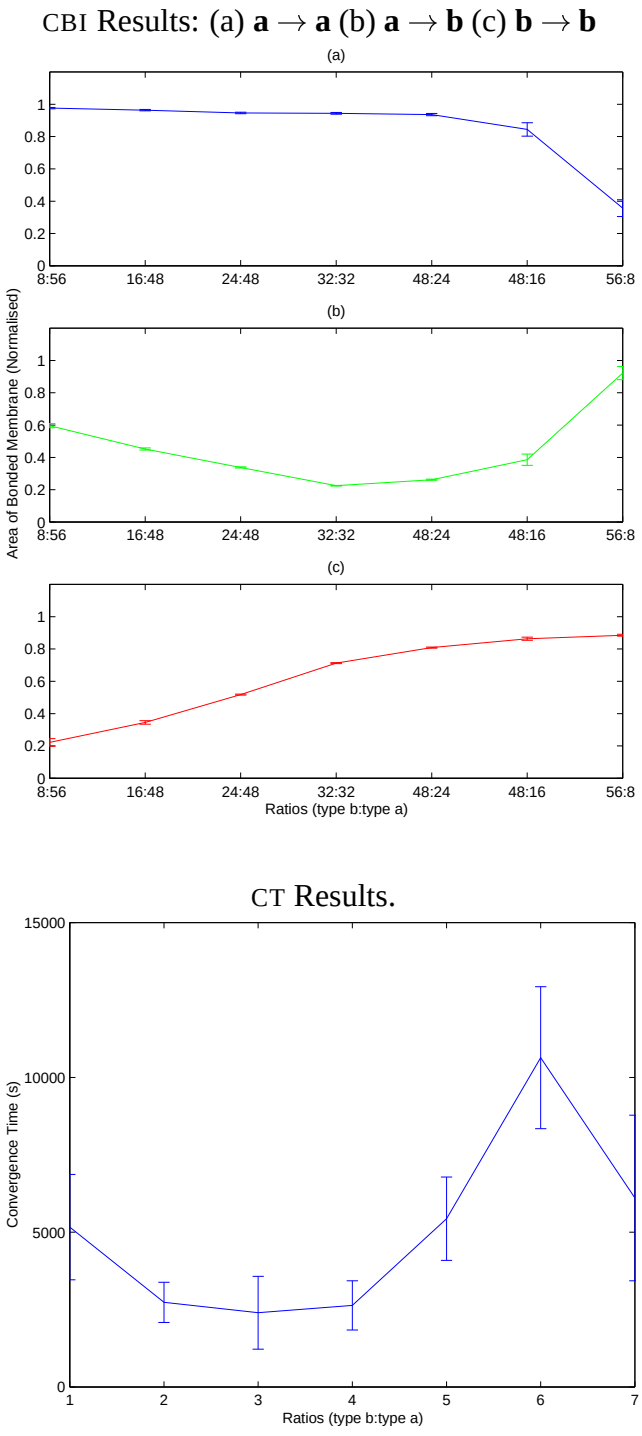


Figure 4.54: *Results for 64 A-Cell aggregates over a range of different ratios of A-Cell types* a*and* b*. The aggregates started from five different initial configurations in which the type* b*distance to neighbours index was low and the type* a*distance to neighbours index was high. The aggregate should move toward the onion equilibrium state.*

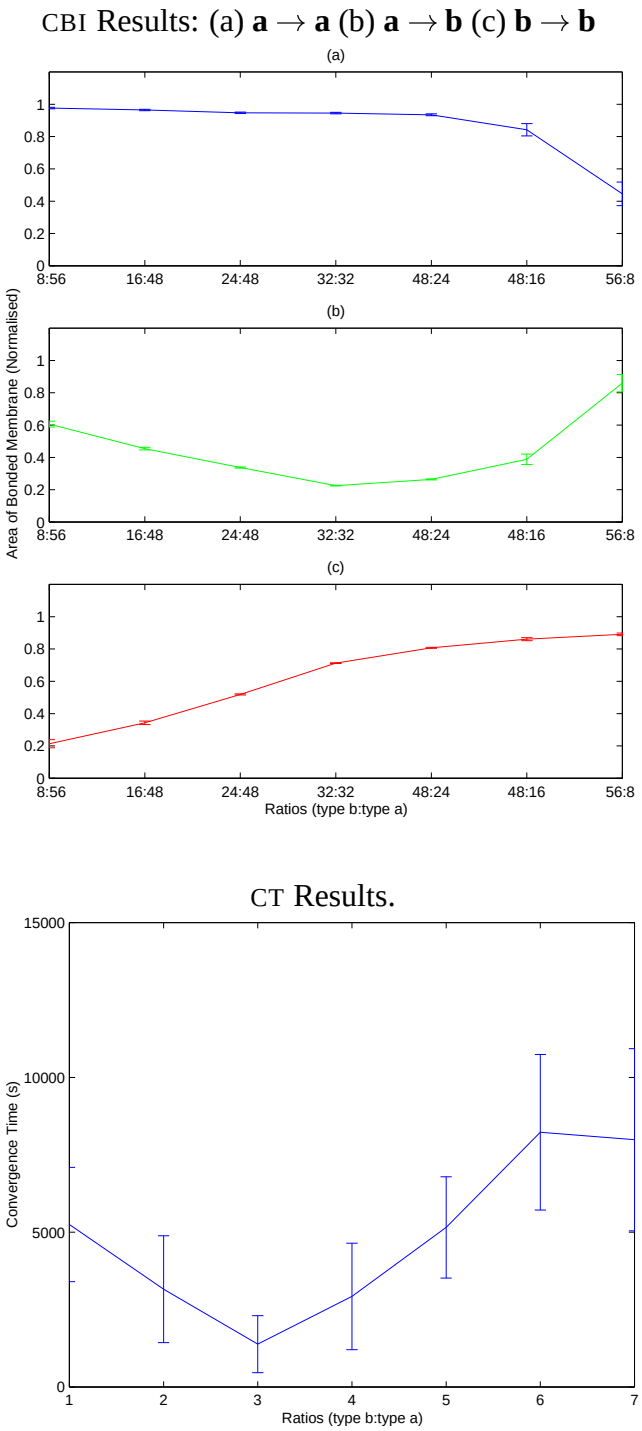


Figure 4.55: *Results for 64 A-Cell aggregates over a range of different ratios of A-Cell types* a*and* b*. The aggregates started from five different initial configurations in with median values for both the type* b*and type* a*distance to neighbours indices. The aggregate should move toward the onion equilibrium state.*

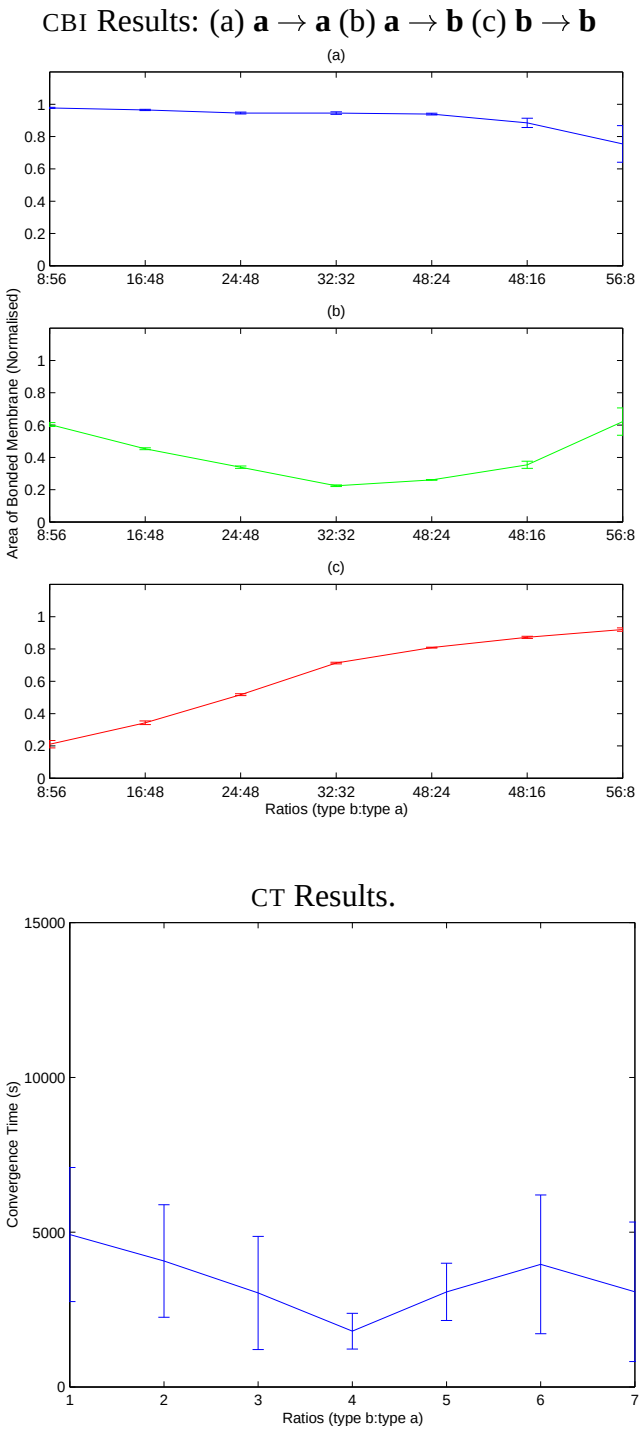


Figure 4.56: *Results for 64 A-Cell aggregates over a range of different ratios of A-Cell types* a*and* b*. The aggregates started from five different initial configurations in which the type* b*distance to neighbours index was high and the type* a*distance to neighbours index was low. The aggregate should move toward the onion equilibrium state.*

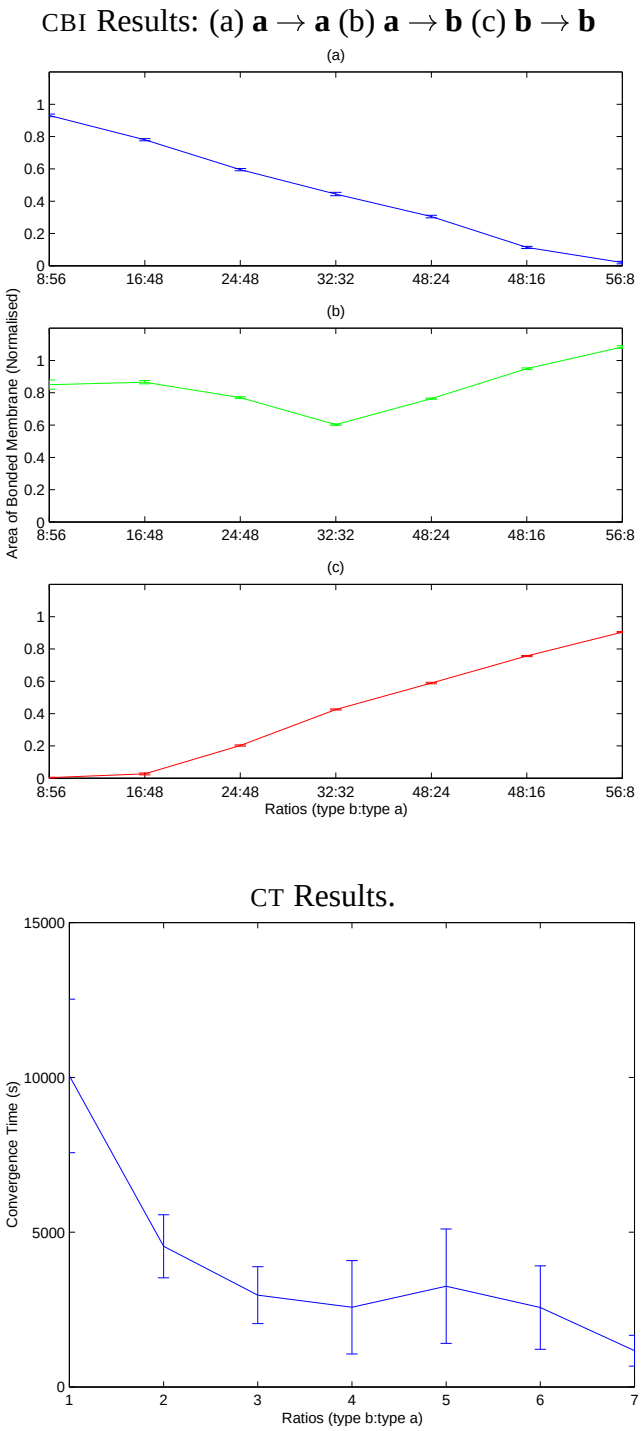


Figure 4.57: *Results for 64 A-Cell aggregates over a range of different ratios of A-Cell types* a*and* b*. The aggregates started from five different initial configurations in which the type* b*distance to neighbours index was low and the type* a*distance to neighbours index was high. The aggregate should move toward the intermixed equilibrium state.*

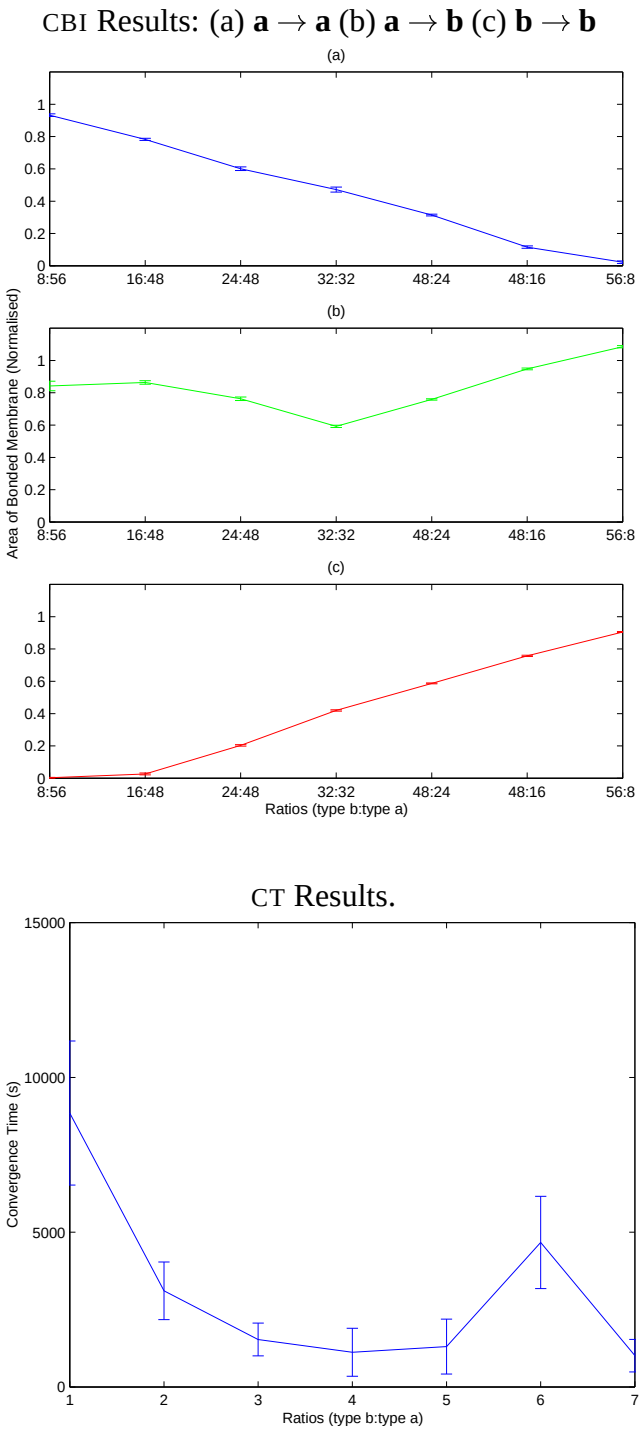


Figure 4.58: *Results for 64 A-Cell aggregates over a range of different ratios of A-Cell types* a*and* b*. The aggregates started from five different initial configurations in with median values for both the type* b*and type* a*distance to neighbours indices. The aggregate should move toward the intermixed equilibrium state.*

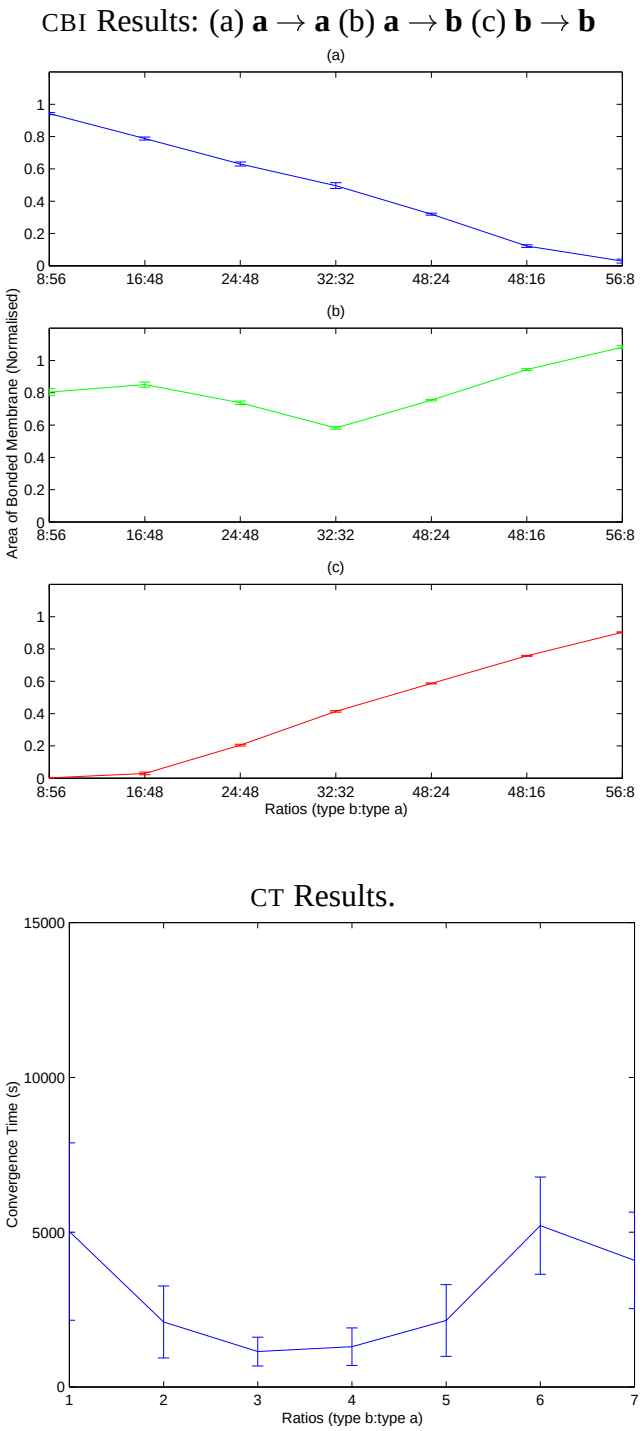


Figure 4.59: *Results for 64 A-Cell aggregates over a range of different ratios of A-Cell types* a*and* b*. The aggregates started from five different initial configurations in which the type* b*distance to neighbours index was high and the type* a*distance to neighbours index was low. The aggregate should move toward the intermixed equilibrium state.*

4.5 Summary

The results presented in this chapter demonstrated that the behaviour exhibited by the cell adhesion model can be clearly predicted by the DAH (see section 4.2.2). This allows simple hierarchical patterns to be encoded solely by specifying the adhesions which exist between populations of units. When the encoded configuration is well defined, the model appears to reach an equilibrium relatively quickly. However, in the specific border cases which are discussed, the convergence time is more variable. It was also demonstrated that there appears to be some minimum threshold in the magnitude of difference in the homotypic adhesions below which only partial sorting can occur. This is in line with simulations carried out by Palsson [45] which indicate a similar result. Once this limit is exceeded, further increases lead to some observable reduction in the required convergence time.

Once more, the experiments also highlighted the important role that local exploration appears to play in the convergence of the model. This was explicitly shown in section 4.3, where increases in the random forces, and thus the level of exploration, were directly linked with improvements in performance. However, it was also shown that, specific to this model, a similar effect can be achieved by reducing the equilibrium distance between the A-Cells thus increasing aggregate fluidity.

Chapter 5

Exploration of 3D Movement

The cell adhesion model presented in this work was designed and implemented to operate in 3D (see section 2.2.1). However, in the preceding chapters, the A-Cell movement was restricted to only two dimensions, allowing a more in depth exploration of the model parameters to be carried out. In this chapter we present a brief analysis of the model's performance without this restriction, rerunning a subset of the experiments from both the rounding and sorting tasks.

5.1 Methodology

5.1.1 Measures

In the earlier chapters (see chapters 3 & 4), the measures that were used to evaluate the adhesion model were selected on the basis that they could be applied to both two and three-dimensional aggregates. The primary measures, from which the others are derived, are the compactness index (used to evaluate A-Cell rounding) and the bonding indices (used to evaluate A-Cell Sorting). However, although the bonding indices can be calculated for two and three-dimensional aggregates using the same technique, in 3D the calculation of the compactness index is slightly different.

In 2D, the compactness equation (see equation 3.2) relates the area of a two-dimensional object with that of a circle which has the same perimeter. To calculate the same measure (C) for three-dimensional objects, it is necessary to calculate the ratio

between the volume of an object (V_{obj}) and the volume of a sphere (V_{sphere}) with the same surface area (S_{sphere}):

$$C = \frac{V_{obj}}{V_{sphere}} = \frac{24\pi^{\frac{1}{2}}V_{obj}}{4S_{sphere}^{\frac{3}{2}}} \tag{5.1}$$

where:

$$V_{sphere} = \frac{4}{3}\pi r^3 \tag{5.2}$$

$$S_{sphere} = 4\pi r^2 \tag{5.3}$$

$$r = \sqrt{\frac{S_{sphere}}{4\pi}} \tag{5.4}$$

In addition, in section 3.1.1.1 it was established that to best approximate the actual behaviour of an aggregate, the compactness index should be calculated from its concave hull (as opposed to the smoother convex hull). In appendix A.1.2, an algorithm for calculating the concave hull of an arbitrary 3D cluster of connected points is presented and discussed. However, in some aggregates, the hull that is determined using this approach may contain single edges which do not form part of a triangular face (e.g. where an A-Cell is connected to only one other A-Cell) (see fig. A.8). As these edges do not contribute to the surface area they will not be accounted for by the compactness index, making the aggregate appear more compact than it actually is. To overcome this, each of these edges is actually treated as a flat rectangular surface with the length of the edge and some small pre-defined width. In this way, aggregates with more of these edges will be suitably punished.

Figures 5.1 and 5.2 show how the new compactness measure relates a typical simulation of a 64 A-Cell aggregate without the restricted movement to the actual aggregate structure. The results show the same characteristic trend observed in early 2D simulations, with an initial rapid rise in compactness as the A-Cells quickly aggregate together followed by a slower convergence to some upper bound.

5.1.2 Initial Configurations

In the 2D rounding experiments (see section 3) the A-Cell aggregates started from a simple linear configuration. This configuration was chosen as it represents the worst

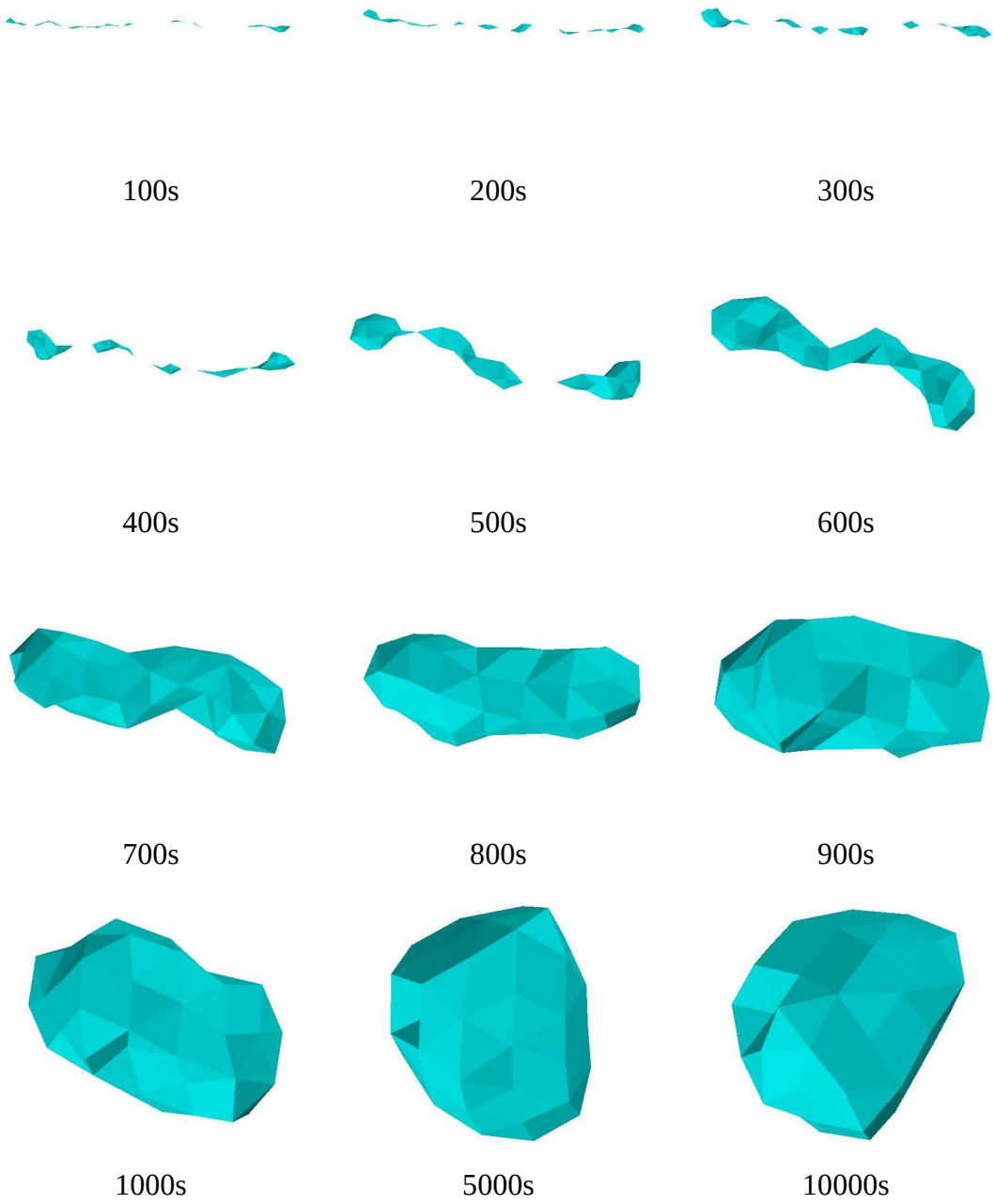


Figure 5.1: *Simulation of the A-Cell model with 3D movement. The concave hulls that are shown were calculated using the algorithm outlined in appendix A.1.2.*

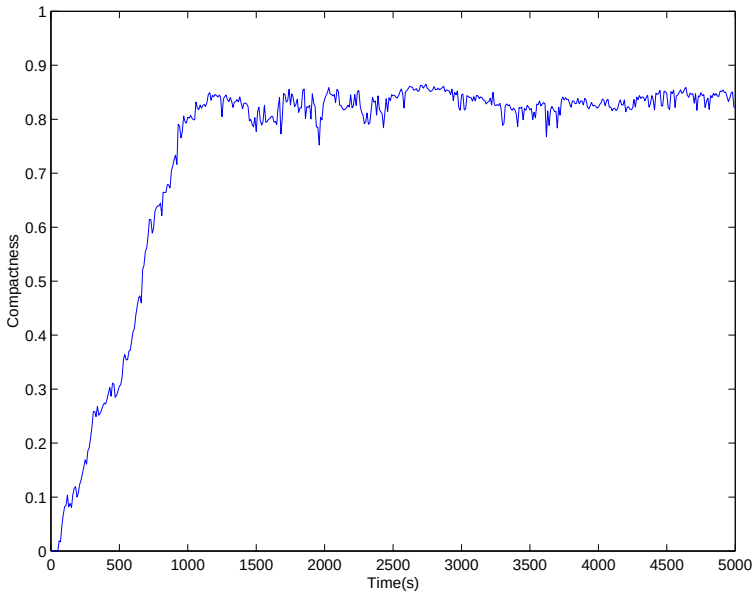


Figure 5.2: *The compactness versus time graph for a simulation of a 64 A-Cell aggregate initially configured in a straight line.*

case starting configuration and thus gives a lower bound on the model's performance. This is still true when considering 3D aggregates. Therefore, the same configuration was used in the similar work presented in this chapter.

In the 2D sorting experiments, two initial configurations were chosen which represent the two extremes of intermixing; an evenly mixed grid and a sorted grid. However, when considering 3D aggregates, the same extremes are more clearly represented by equivalent lattice structures where the A-Cells are arranged in a cube. Therefore, as the purpose of the sorting experiments is to evaluate the changing organisation of the two cell types as opposed to changes in the aggregate shape, the lattice configurations were used in the appropriate experiments presented in this chapter (see fig. 5.3).

Finally, throughout this chapter only aggregates of 64 A-Cells are considered. Smaller numbers were disregarded as the diameter of the 3D aggregates that would be formed would make both rounding and sorting reasonably trivial. In addition, the length of time required for simulations of larger populations would have greatly limited the range of experiments which could be carried out.

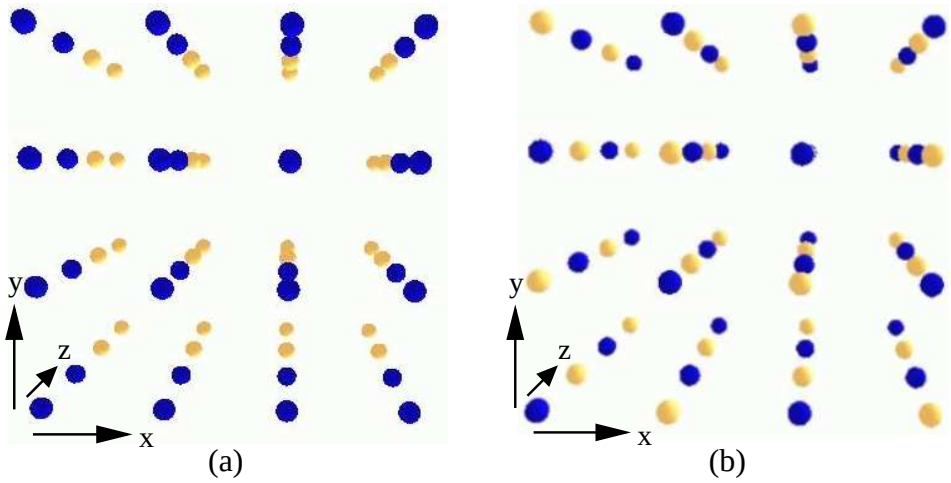


Figure 5.3: *Initial configurations used for the 3D simulations. (a) Sorted lattice (b) Intermixed lattice.*

5.1.3 Base Parameter Set

The base parameter sets used for the three-dimensional simulations are largely the same as those used for their two-dimensional counterparts (see sections 3.1.3 & 4.1.5). The main exceptions to this are in the number and position of sensors and adhesion sites. Obviously, as the A-cells are no longer fixed to the same horizontal plane it is not sufficient to simply place these around the equator of their hull and membrane. Instead, they should be appropriately distributed across the surface of the respective spheres. In the case of the sensors, as the number which are used should be kept to a minimum, the problem is reasonably trivial. In fact, a reasonable coverage can be provided with only eight sensors arranged at the corners of a cube inscribed inside an A-Cell hull (as per the original HYDRON design, see section 2.1.2). Therefore, this is the number used throughout this chapter.

In the case of the adhesion sites, the problem is more complex as it should be possible to generate a suitable arrangement for any number of sites, such that the sites are evenly distributed. This equates to the general problem of distributing n points on a unit sphere. The problem has a long history and has raised much interest. However, excluding certain special values of n, the solutions proposed are limited to approximations which optimise specific criteria [51]. For the purpose of the adhesion sites it is

possible to consider the area of membrane represented by each site as a disc of radius r centred at the location of the site. Therefore, the problem is to maximise r such that no two disks overlap (e.g. minimise the area of membrane which is not covered). This is, in effect, maximising the distance between the two closest points on the sphere. In keeping with the self-organising aspect of the cell adhesion model, in this work this is achieved by using a charged particle model [1]. This involves representing each of the n points as a unique charged particle with its movement restricted to the surface by a unit sphere. The particles are first placed on the sphere at random, then for successive cycles the repulsive charges between the particles are calculated and applied. In this way the points slowly move toward some equilibrium state where the distance between any two particles will be maximised.

To determine the adhesion site positions used in this chapter, the charged particle model was started from 5 random positions and stopped when the largest particle movement dropped below 0.00001. Of these five, the configuration which maximised the distance between the two closest points was selected. The points were then scaled appropriately to ensure that they lie on the surface of each A-Cell's membrane. An example distribution for 300 points is shown in figure 5.4.

As a default value, each A-Cell used 100 adhesion sites. However, a range of possible values were also explored in section 5.2.2.

5.2 Rounding

5.2.1 Random Forces

The following experiments aim to investigate how small random movements and random torques affect the 3D aggregates. The method which is used replicates the previous 2D work presented in sections 3.3 and 4.3. Again, as the A-Cells are assumed to be weighted such that a self righting torque will keep them upright (see section 2.2.1), only random torques around the vertical (y) axis are considered. However, as the A-Cells are no longer restricted to movement in two dimensions neither are the random

[1] The code used was 'diffuse' which was implemented by Jonathan D. Lettvin (Copyright©2003) and offered under the GPL.

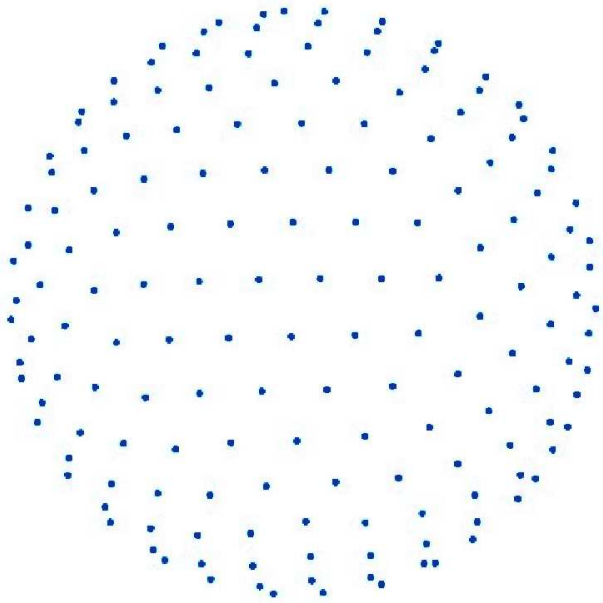


Figure 5.4: *A configuration of 300 points determined by the charged particle model. The points on the rear facing hemisphere have been removed from this view.*

forces (see fig. 5.5).

Figures 5.6, 5.7, 5.8, 5.9 and 5.10 show the results of the simulations where the separate forces were considered independently [2]. The results showing the combined effect of the forces (varying random movement force limit, fixed random torque limit) are given in figures 5.11, 5.12 and 5.13.

These results demonstrate some of the familiar traits exhibited in the earlier 2D simulations.

- In the absence of random torques, increasing the random movement force limit causes a steady rise in the CC , until the point where A-Cells begin to break away from the main aggregate.

- The level of random torque has little effect on either the CR or CC .

[2]No histogram is given for the torque experiments as the aggregates remained intact for all the considered values.

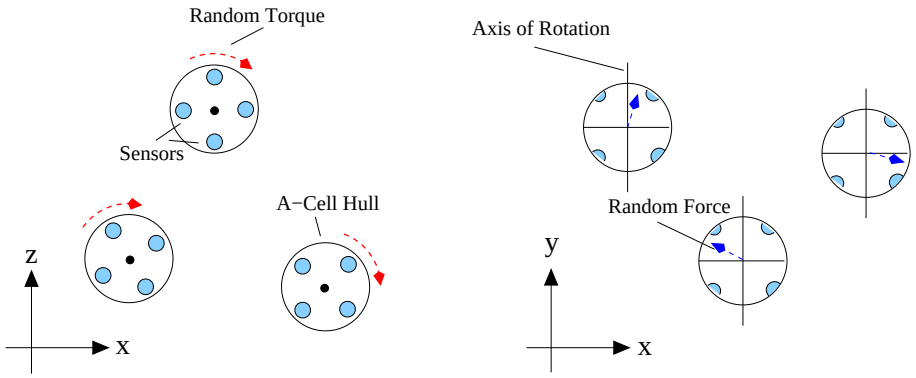


Figure 5.5: *The random forces and random torques acting on the A-Cells.*

- The best results are achieved with a combination of random torques and random movement forces.

However, there are a couple interesting observations which can also be made. Firstly, we would expect the two-dimensional aggregates to exhibit less robustness to the random movement forces than similar 3D aggregates. The reason for this assumption is that A-Cells at the extremities of a 2D aggregate only have neighbours in the same plane to bond with, whereas, similarly placed A-Cells in a 3D aggregate may have additional neighbours positioned above or below which they can also bond with, and thus increase the force required to split them from the main aggregate. Despite this, the results for the simulations exposed to combined random movement and random torque forces (see fig. 5.11 & 3.14) show that, in both two and three dimensions, similarly sized aggregates lose the ability to remain intact when exposed to the same random movement force limits. In addition, when exposed to random movement forces in the absence of random torques (see fig. 5.6 & 3.12), the 2D aggregates actually appear slightly more robust. However, examining the simulation logs it is clear that the aggregates actually break apart very early in the simulations (e.g. within the first few hundred seconds), where both the 2D and 3D aggregates are still very linear in structure. Therefore, the 3D aggregates are unlikely to have any true advantage. In fact, with the additional dimension of movement, the 3D aggregates should actually be at more risk of breaking apart, as observed in the results.

The second observation relates to the CC results obtained when the 3D aggregates

were exposed only to random torque forces. In each case, the aggregates still form reasonably compact structures. However, in the 2D case, the CC values tend to fluctuate above 0.7 while the 3D results appear to be consistently lower. There are two possible explanations for this. Firstly, it could simply be the case that the rounding task is more difficult in 3D. However, this is not backed up by the other results in which neither model outperforms the other in all cases. The second possibility is that the poorer performance is linked to the fact that the random movements generated by the random torques will mainly be horizontal rather than vertical in nature. The reason for this is that the torques will cause the sensors to rotate around the vertical axis, changing the distance and direction estimates (see sections 3.5.1 & 3.5.3). In theory this should result in a slower vertical spread of the A-Cells.

To test the latter hypothesis, the three sets of simulations were reprocessed to determine the mean vertical displacement of the A-Cells in an aggregate's final configuration from the aggregate's centre of mass. As in the calculation of the CC (see section 3.1.1.2) this was achieved by calculating the values for the last *n* seconds of the simulation and taking the average. The results that were obtained (see fig. 5.14, 5.15 & 5.16) show that, as predicted, the mean vertical displacement achieved when only torque is applied is consistently lower than in either case where random movement forces are also provided. To confirm that this was not simply a result of the aggregates producing less random movement in all directions, the same displacement measure was calculated for the z axis (given that the initial configuration is a line parallel to the x axis) (see fig. 5.14, 5.15 & 5.16). These results show that the simulations with only torque have a considerably higher displacement along the z axis. If there was an equal amount of random movement in all directions, the displacement from either of the axes should be equivalent. Therefore, this cannot be the case.

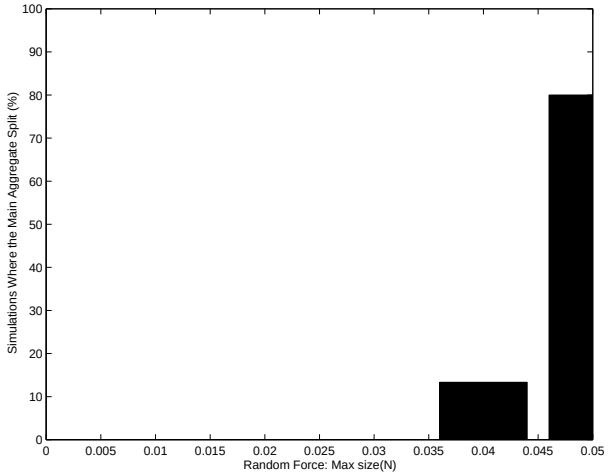


Figure 5.6: *Histograms showing the number of aggregates that did not remain intact over a range of random movement force limits.*

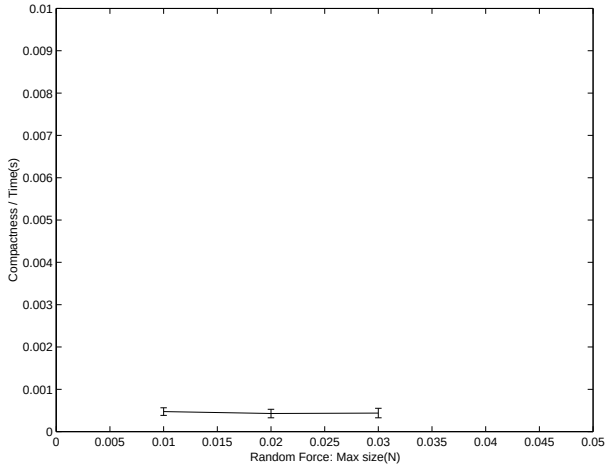


Figure 5.7: CR *values of A-Cell aggregates over a range of random movement force limits. The* x *axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown.*

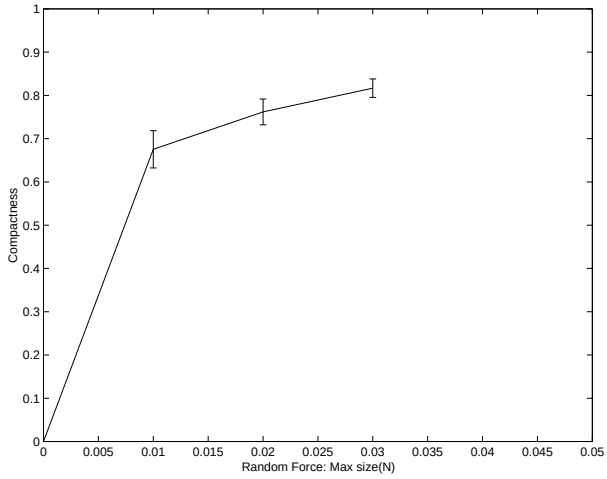


Figure 5.8: CC *values of A-Cell aggregates over a range of random movement force limits. The x axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown.*

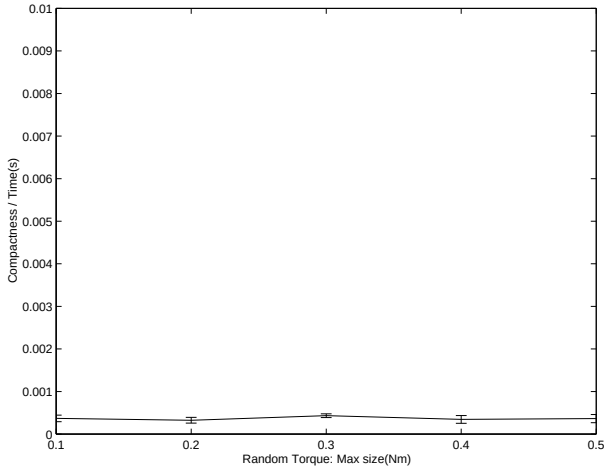


Figure 5.9: CR *values of A-Cell aggregates over a range of random torque limits.*

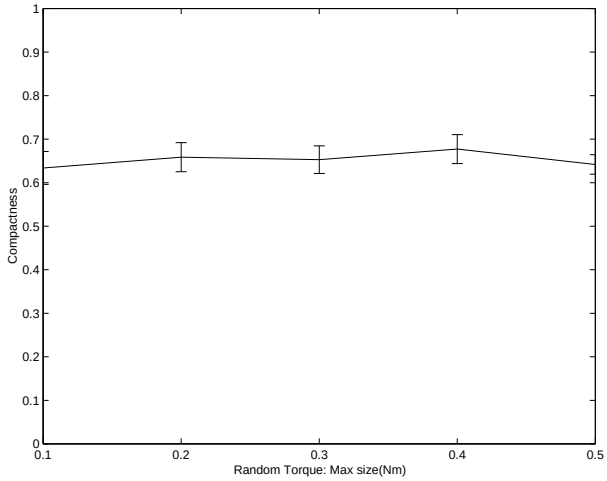


Figure 5.10: CC *values of A-Cell aggregates over a range of random torque limits.*

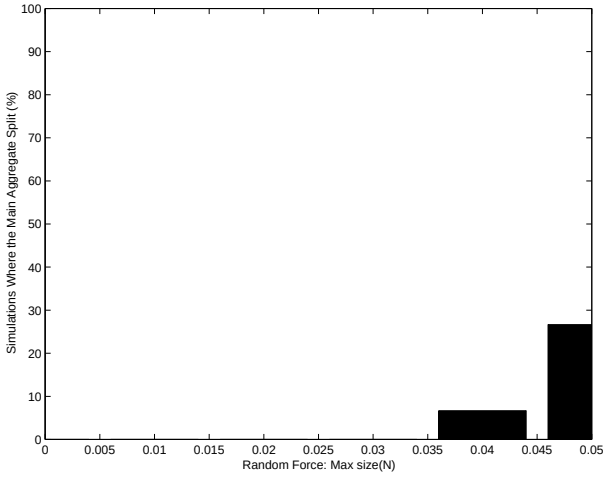


Figure 5.11: *Histograms showing the number of aggregates that did not remain intact with a fixed random torque limit of* $0.01N$ *over a range of random movement force limits.*

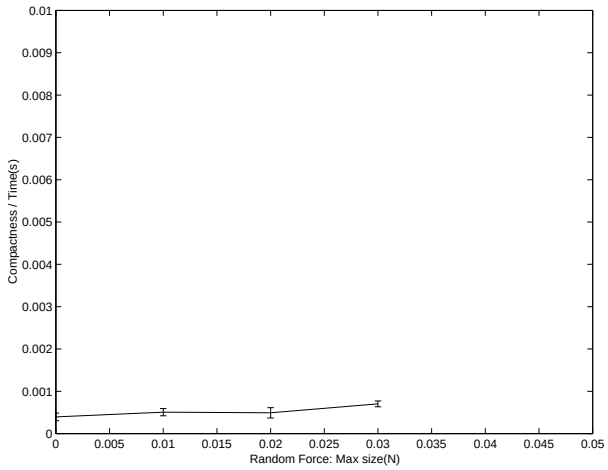


Figure 5.12: CR *values of A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random movement force limits. The* x *axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown.*

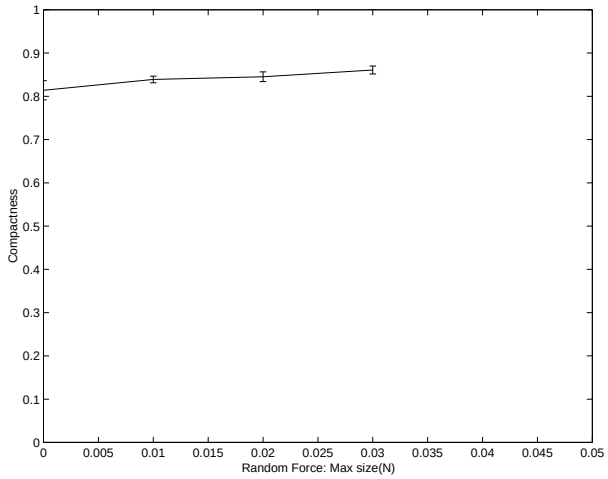


Figure 5.13: CC *values of A-Cell aggregates with a fixed random torque limit of* $0.01N$ *over a range of random movement force limits. The* x *axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown.*

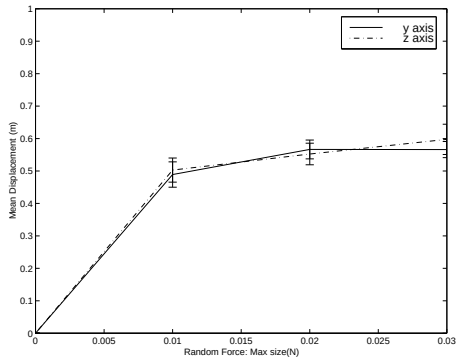


Figure 5.14: *The mean displacement of A-Cells from an aggregate's centre of mass along either the z or y axis for a range of random movement force limits.*

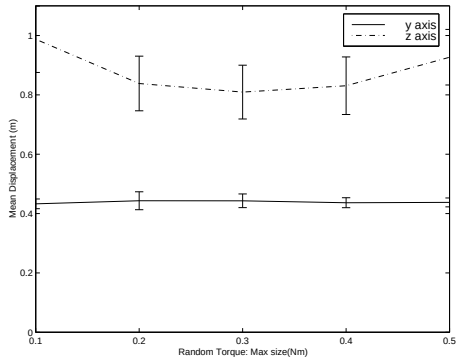


Figure 5.15: *The mean displacement of A-Cells from an aggregate's centre of mass along either the z or y axis given a range of random torque limits.*

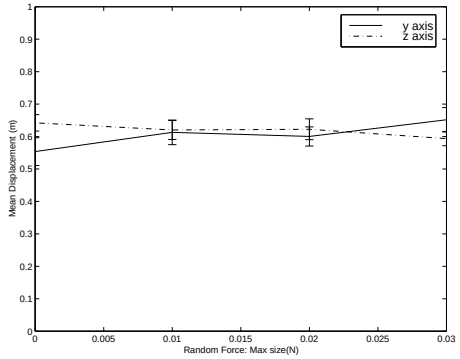


Figure 5.16: *The mean displacement of A-Cells from an aggregate's centre of mass along either the z or y axis given a range of random movement force limits and a fixed random torque limit of* $0.01N$*.*

5.2.2 Adhesion Sites

The number of adhesion sites effectively determines the resolution (granularity) at which the numbers of bonded A-CAMs can be calculated (see section 2.2.5). The following experiments were carried out to determine how this affects the 3D aggregates and in particular, as a larger number of adhesion sites require more of each A-Cell's processing time, what the minimum number of adhesion sites must be to achieve reasonable performance.

The CR and CC results shown in figures 5.20 and 5.21 cover a large range of adhesion site numbers ($50 \rightarrow 300$). However, as in the earlier 2D results (see section 3.4.2) it appears that over this range of higher values, the specific number of adhesion sites has no discernible effect on either measure.

An additional set of experiments covering a range of smaller adhesion site numbers in more detail were also carried out and the results are shown in figures 5.17, 5.18 and 5.19. In this case, as one might expect, when using many of the smaller values, A-Cells simply broke away from the main aggregate (see fig. 5.17). However, a minimum of only 12 adhesion sites allowed the majority of the aggregates to remain intact for the duration of the simulation and once this value was increased to 16 all the aggregates remained intact.

The 2D simulations demonstrated that the smaller values cause the aggregates to break apart because the number of sites on the overlapping areas of membrane between two A-Cells will fluctuate as each A-Cell rotates. As the number of sites drops, these fluctuations increase. Calculating the mean compactness variance of the simulation carried out in this section (see fig. 5.22) indicates that the same behaviour is taking place in the 3D aggregates. In addition, the trends show that in both the 2D and 3D, the compactness variance graphs correspond very closely (see fig. 3.25 & 5.22). The key difference between the results is in the scale, which is significantly larger in the 3D case. However, this is unsurprising as, in the 3D simulations, the adhesion sites are distributed across the entire membrane, thus only a subset of these will ever be involved in the overlap between two neighbouring A-Cells.

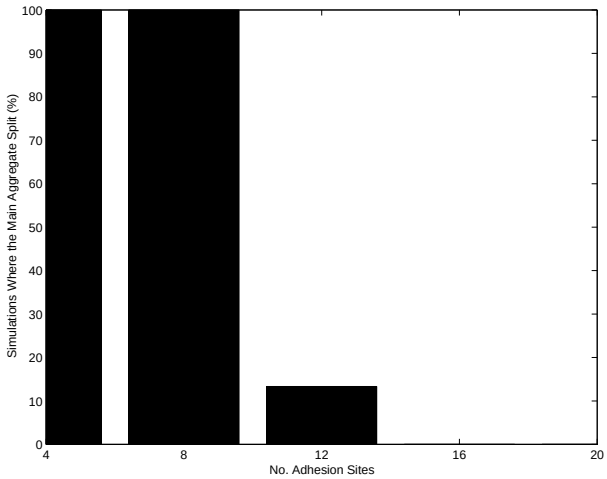


Figure 5.17: *Histograms showing the number of aggregates that did not remain intact over a range of adhesion site numbers.*

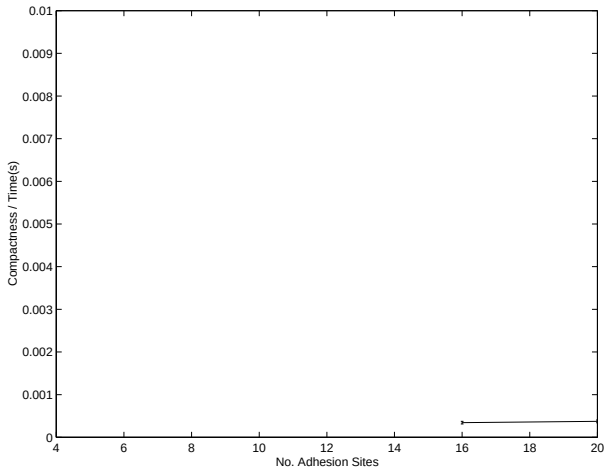


Figure 5.18: CR *values of A-Cell aggregates with a range of adhesion site numbers. The* x *axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown.*

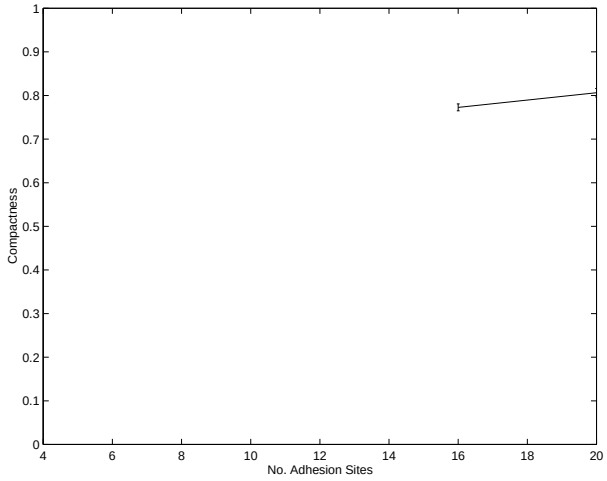


Figure 5.19: CC *values of A-Cell aggregates with a range of adhesion site numbers. The* x *axis shows the complete range. However, only values from samples where all the aggregates remained intact are shown.*

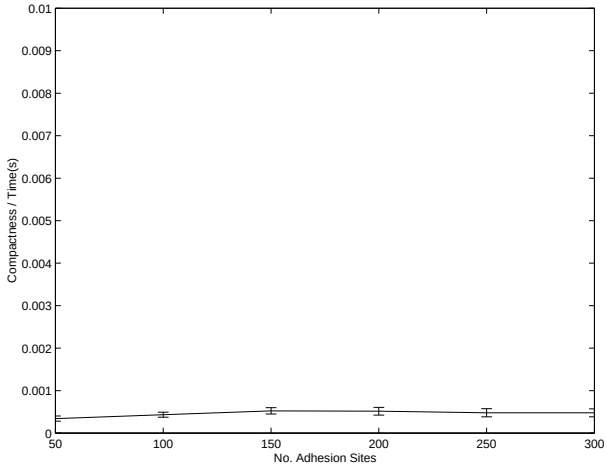


Figure 5.20: CR *values of 64 A-Cell aggregates with a range of adhesion site numbers.*

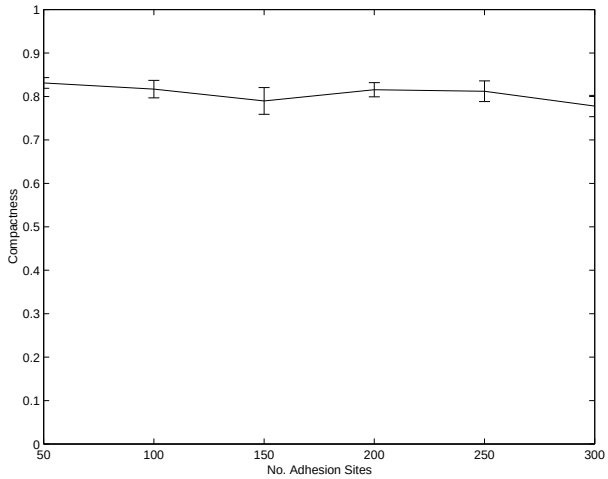


Figure 5.21: CC *values of 64 A-Cell aggregates with a range of adhesion site numbers.*

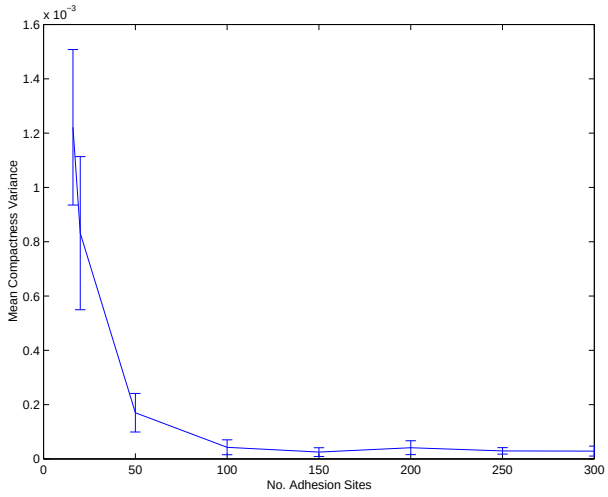


Figure 5.22: *The mean compactness variance of 64A-Cell aggregates for a range of adhesion site numbers*

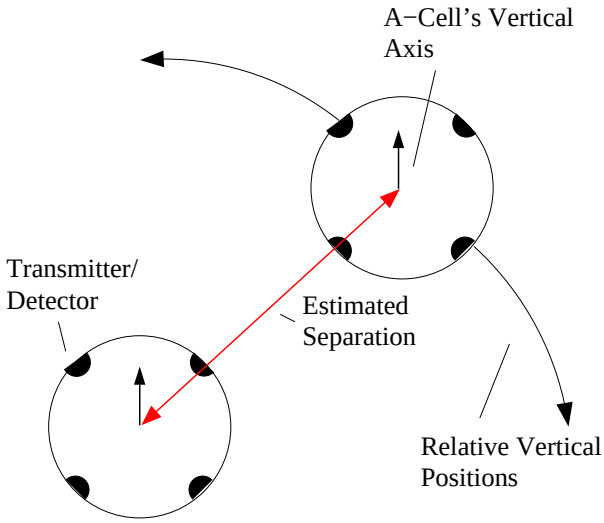


Figure 5.23: *The arc of relative vertical positions that were considered for each orientation pair in the simple 2 A-Cell model.*

5.2.3 Sensor Field of View

The experiments presented in this section investigate how the performance of the model is affected by varying the FOV of the available transmitters and detectors. As discussed in section 5.1.3, to account for the 3D movement, a greater number of sensors are used. However, these must cover the whole A-Cell hull, unlike the earlier 2D simulations where all signals were simply transmitted on the same plane.

Once more, the experimental results are presented alongside estimates obtained from a simple two A-Cell model. The model is essentially the same as the one presented in section 3.5.1. However, to account for the addition of 3D movement, each orientation pair must also be considered in a number of relative vertical positions (see fig. 5.23). In addition, a slightly larger experimentally determined separation distance of $0.7597m$ was used[3].

As in the 2D simulations, the histogram results show the most significant data and demonstrate the same sudden change in performance when either the transmitter or detector FOV drops below $\frac{\pi}{4}$ (see fig. 5.24). Unexpectedly, the results also show

[3]This distance was calculated as the mean distance between neighbouring A-Cells for a sample of 15 simulations with the base configuration.

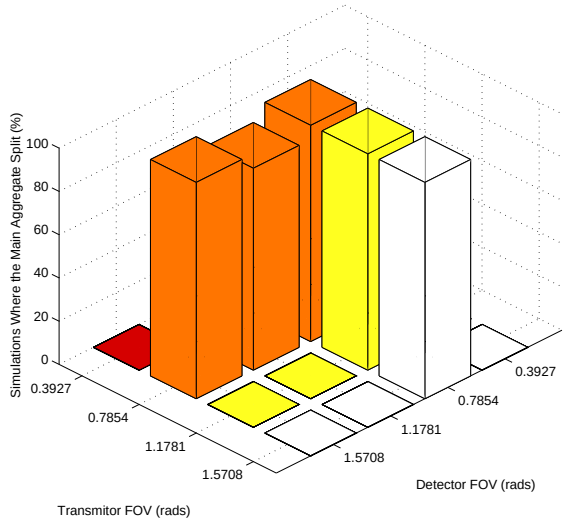


Figure 5.24: *Histograms showing the number of aggregates that did not remain intact over a range of sensor fields of view.*

that the aggregates remain intact for one hundred percent of the simulations when either FOV is only $\frac{\pi}{8}$. However, in this case the explanation is simply that the FOV is too small to allow any communication between neighbouring A-Cells in the linear starting configuration. Therefore, although the A-Cells spin due to the random torques, as they are not exposed to any random movement forces, they remain in their initial configuration and, thus, the aggregates appear to remain intact.

In section 3.5.3 this sudden change is linked to the value below which the sensor coverage diverges with distance from the A-Cell hull. This should also be the case in the 3D aggregates and explains the results obtained from the two A-Cell model (see fig. 5.27 - a). As predicted, when this value is crossed, there is a rapid increase in the number of configurations where any two A-Cells will be out of contact. In addition, the rate of change is significantly higher which reflects the fact that the effect is magnified in the 3D case (see fig. 5.27 - b).

The CC and CR results which were calculated from the simulations are shown in figures 5.25 and 5.26. However, although the results show that the model's performance was good when the aggregates were able to remain intact, there are too few results to reveal any trends. Despite this, the mean distance and direction estimate errors were

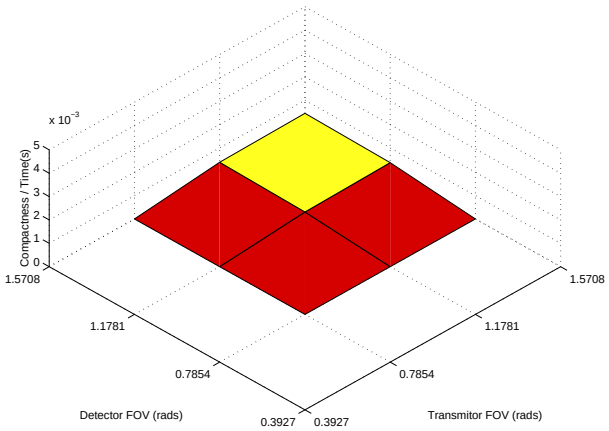


Figure 5.25: CR *values of A-Cell aggregates over a range of sensor fields of view.*

still calculated using the simple two A-Cell model. The distance estimate errors show the same basic trend as the similar 2D results (see fig. 5.28 - a). However, comparing the two sets reveals that in the 3D case, the errors are slightly worse (see fig. 5.28 - b). The direction errors that were calculated are also reasonably similar to the 2D results (see fig. 5.29). However, they don't show the same local minima (see fig. 3.44).

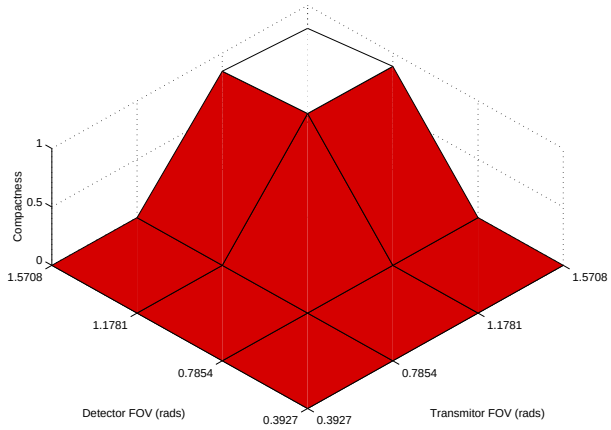


Figure 5.26: CC *values of A-Cell aggregates over a range of transmitter and detector FOVs.*

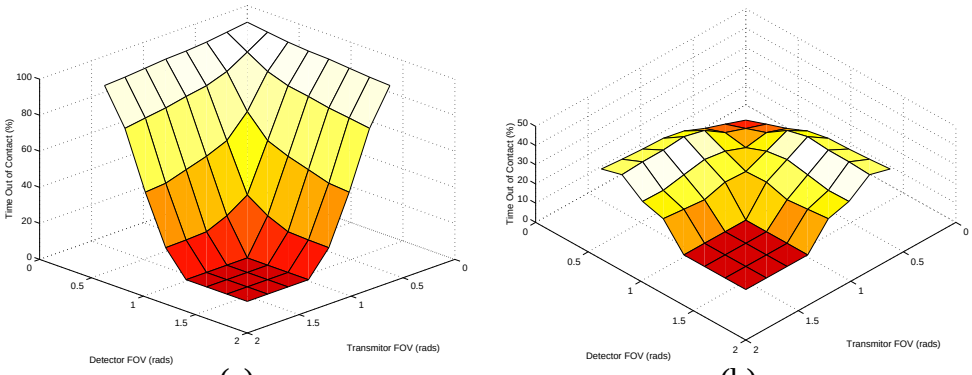


Figure 5.27: *(a) The percentage of orientation pairs when two A-Cells with 3D movement cannot communicate over a range of detector and transmitter FOVs. (b) The difference between the 3D and 2D results (2D subtracted from 3D).*

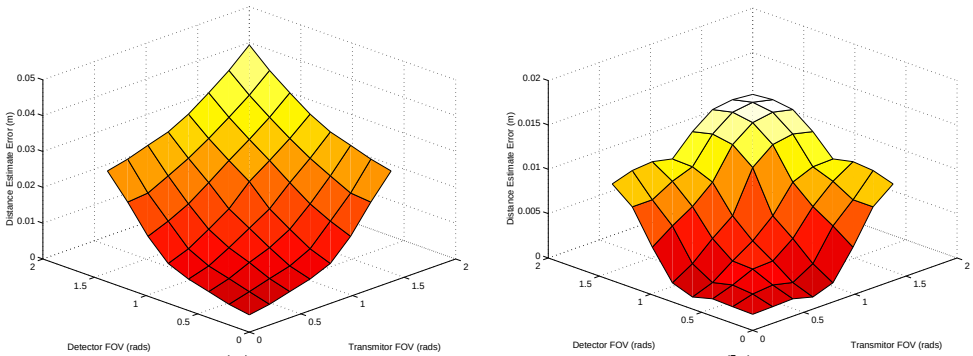


Figure 5.28: *(a) Mean error in the distance estimate calculated from a simple two A-Cell model over a range of transmitter and detector FOVs. (b) The difference between the 3D and 2D results (2D subtracted from 3D).*

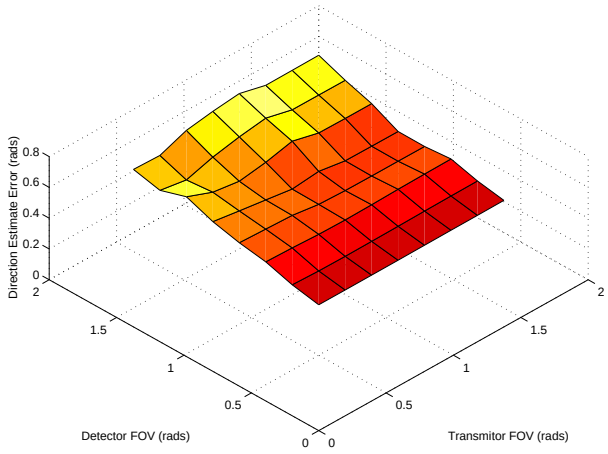


Figure 5.29: *Mean error in the direction estimate calculated from a simple two A-Cell model over a range of transmitter and detector FOVs.*

5.3 Sorting

In this section, reruns of the earlier 2D differential adhesion experiments (see section 4.2) are presented. As before, two sets of experiments were carried out investigating the model's response to ranges of both heterotypic and homotypic adhesions.

5.3.1 Heterotypic Adhesion

In the following experiments, the homotypic adhesions are fixed at the original values discussed in section 4.1.5:

- $\mathbf{b} \rightarrow \mathbf{b}$ – attraction at equilibrium of $0.15125N$
- $\mathbf{a} \rightarrow \mathbf{a}$ – attraction at equilibrium of $0.3025N$

In addition, the heterotypic adhesion is varied between $0N$ and $0.45375N$. Therefore, the key values in heterotypic adhesion which should border changes in the final configuration of the aggregate will remain at $0.15125N$, $0.226875N$, and $0.3025N$ (see section 4.2.1).

Figures 5.30 and 5.34 show the results that were obtained when starting the aggregates from either a sorted or intermixed lattice. It can be seen that in the extremes of the graph ($0 \rightarrow 0.1134375$ and $0.3403125 \rightarrow 0.45375$), where the aggregates should achieve either a partially sorted or completely intermixed configuration, the results quite closely match those that were obtained for the similar 2D simulations (see fig. 4.10 & 4.15). However, in the central section of the graph (see fig. 5.32), where the final configuration changes from sorted to intermixed, instead of the expected sudden change the results show a much more gradual transition.

Investigating this range in more detail (see fig. 5.32 & 5.37) reveals that in both instances only partial sorting is achieved. In addition, when starting from an intermixed lattice, the aggregates only achieve this when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion is lowest. However, when starting from the sorted lattice, the level of sorting slowly increases until the central boundary is reached and then intermixing starts to occur (see fig. 5.38).

One possible explanation for this behaviour is that the cubic close packing observed in the A-Cell aggregates (see section 5.4) is too rigid and limits local exploration, thus

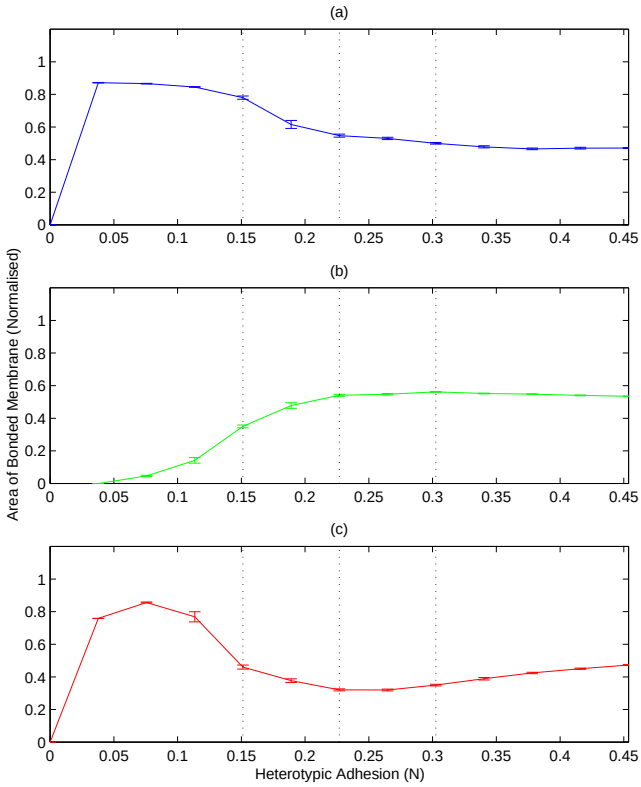


Figure 5.30: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed lattice. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

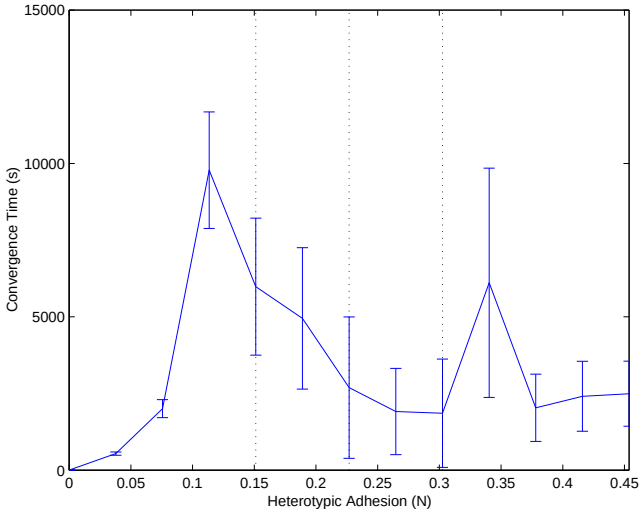


Figure 5.31: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed lattice*

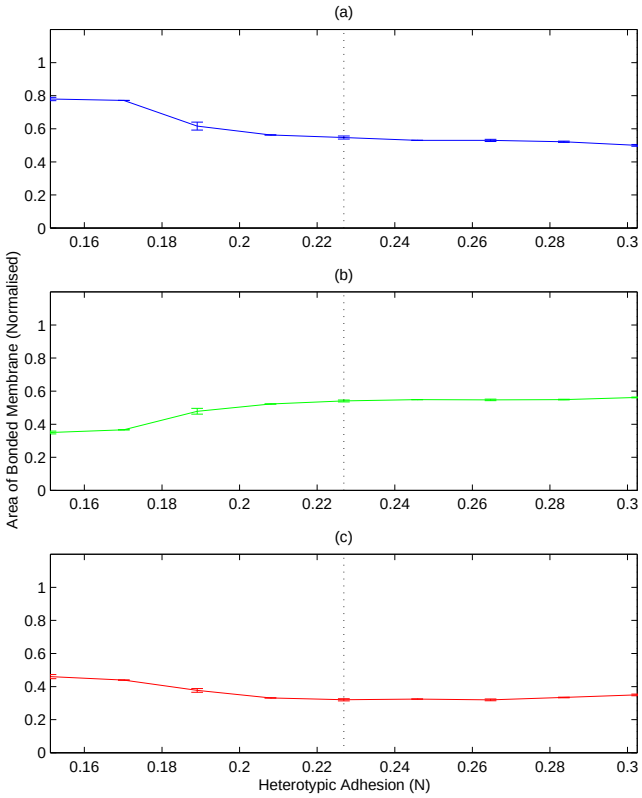


Figure 5.32: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed lattice. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

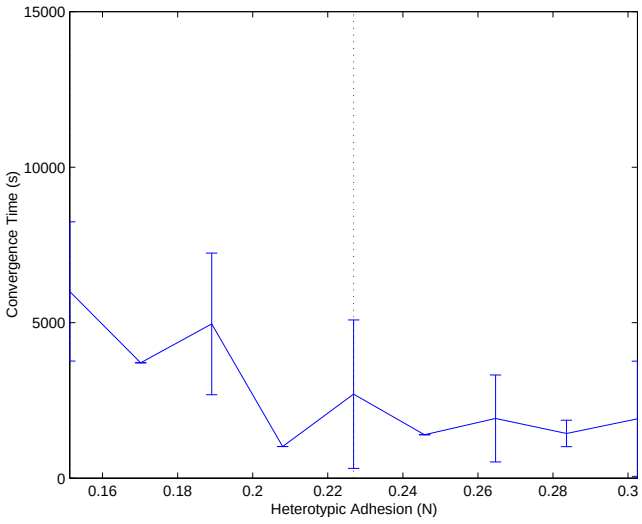


Figure 5.33: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed lattice*

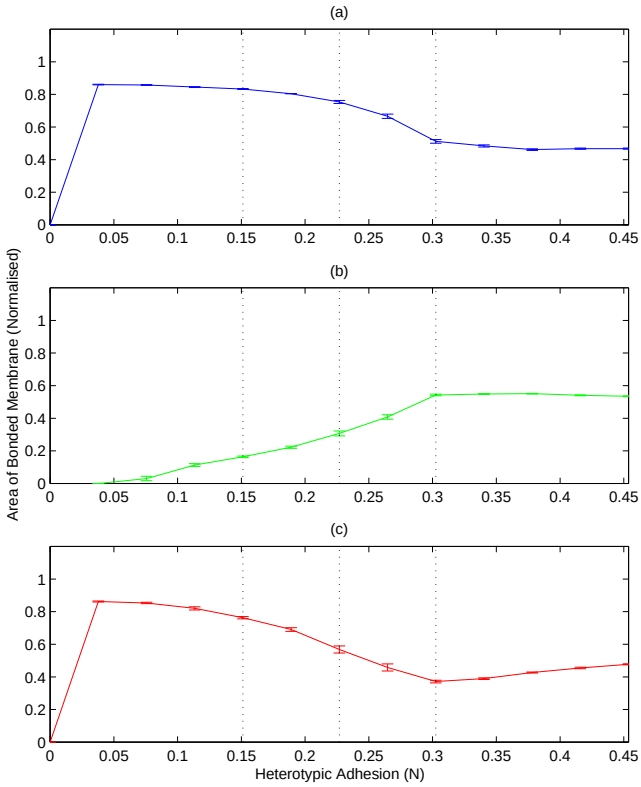


Figure 5.34: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted lattice. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

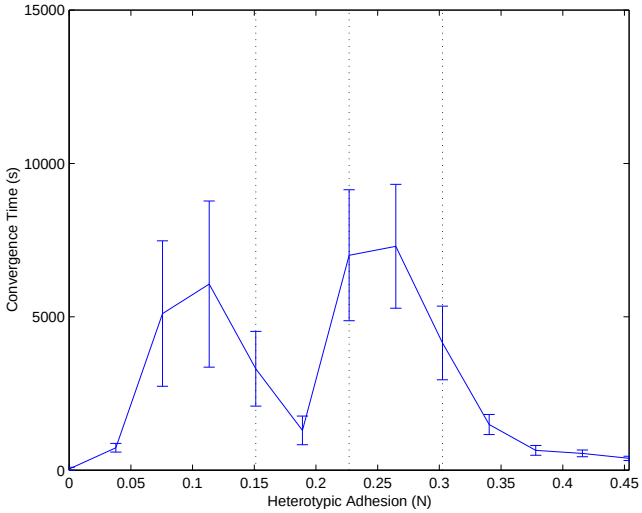


Figure 5.35: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted lattice*

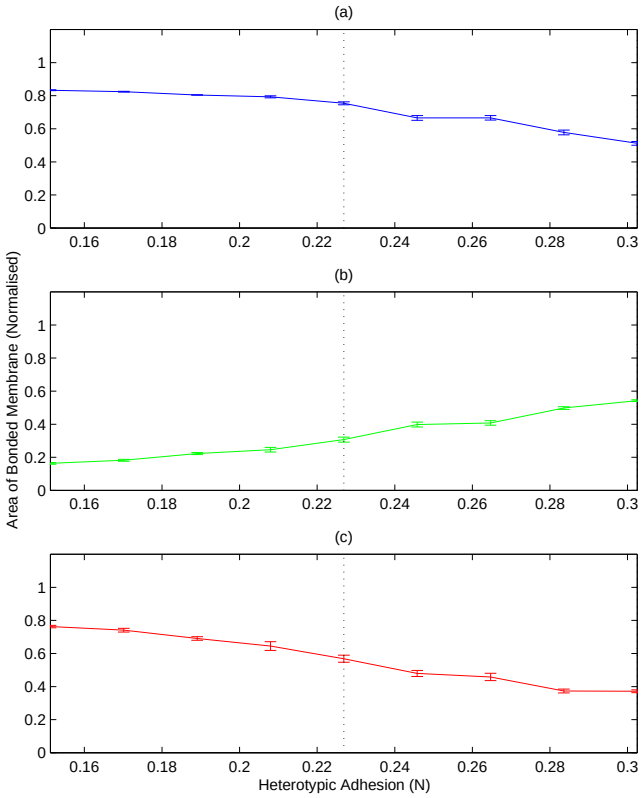


Figure 5.36: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted lattice. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

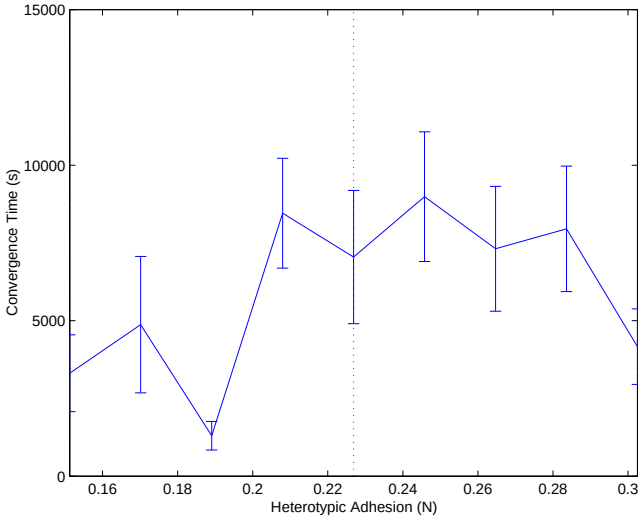


Figure 5.37: CT *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted lattice*

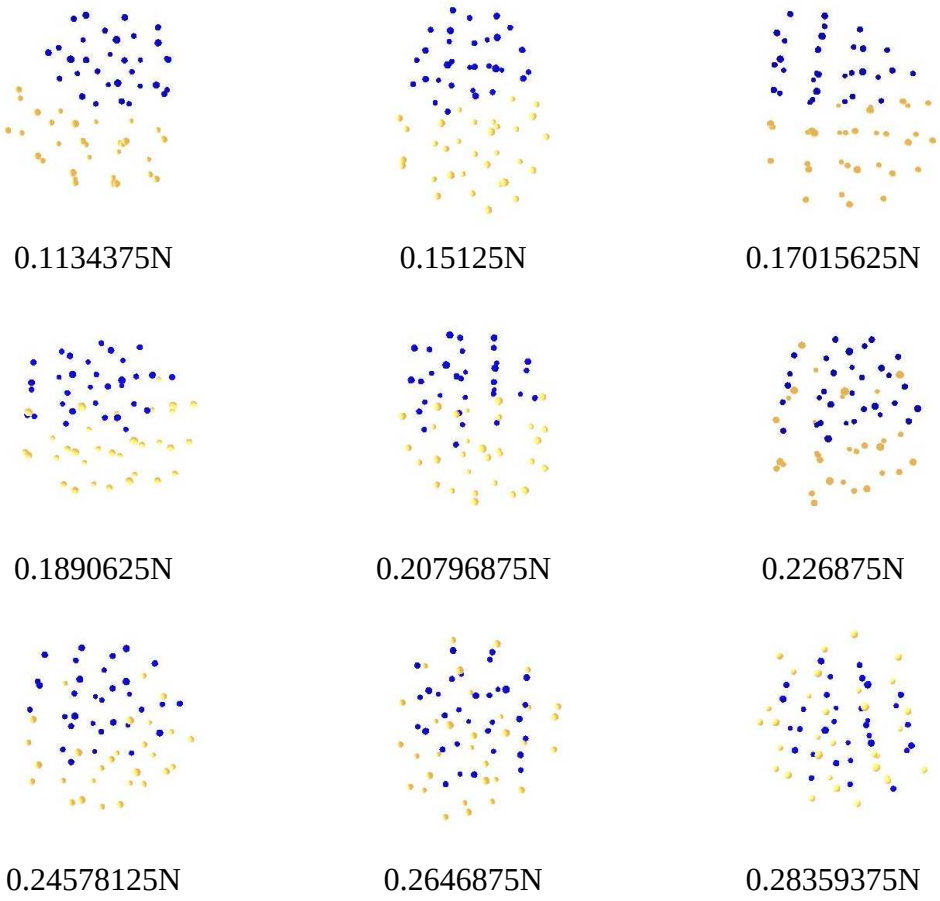


Figure 5.38: *Screen shots showing the final configuration of aggregates consisting of two A-Cell types with various heterotypic attraction at equilibrium values. The aggregates consisted of 64 A-Cells and started with the initial configuration of a sorted lattice.*

preventing the aggregates from achieving the final configurations. However, if this were the case, there should be also be some noticeable effect when the $\mathbf{a} \rightarrow \mathbf{b}$ adhesion is at either extreme and this is not observed in the actual results. Instead, it appears that the actual cause relates to level of difference between the homotypic adhesions. Some evidence for this is provided by the earlier 2D experiments presented in section 4.2.2. These experiments compared the effect of reducing or increasing the level of difference between the homotypic adhesions and demonstrated the existence of a minimum threshold below which the expected sorting can no longer take place. The results obtained for the simulations with magnitudes of difference in the homotypic adhesion below this threshold very closely match those presented in this section. In addition, the results presented in the following section (see section 5.3.2), which carries out a similar comparison of differences in homotypic adhesion, reveals that when the difference in homotypic adhesion is increased to $0.3025N$, the behaviour of the 3D aggregates more closely resembles that predicted by the DAH.

This raises the question: why does the required minimum difference in homotypic adhesion appear to increase when moving from two to three dimensional aggregates? A likely explanation simply relates to the number of neighbours each A-Cell has. The reason for this is that the model is partially driven by the A-Cells randomly discovering other A-Cells they can produce stronger bonds with. When such an encounter occurs, the new bond should produce some small force which will help pull the two A-Cells together. However, clearly this force will act against the existing bonds between the A-Cells and their neighbours. Therefore, if the number of neighbours the A-Cells have increases, the effect of any new bonds will be reduced.

5.3.2 Difference Between Homotypic Adhesions

The experiments presented in this section replicate the earlier 2D simulations presented in section 4.2.2. The same four magnitudes of difference in the homotypic adhesions were considered, again fixing the weaker $\mathbf{b} \rightarrow \mathbf{b}$ adhesion at $0.15125N$ while increasing the higher $\mathbf{a} \rightarrow \mathbf{a}$ adhesion as follows:

1. $\mathbf{a} \rightarrow \mathbf{a} = 0.1890625N$

2. $\mathbf{a} \rightarrow \mathbf{a} = 0.226875N$

3. $\mathbf{a} \rightarrow \mathbf{a} = 0.3025N$ (from section 5.3.1)

4. $\mathbf{a} \rightarrow \mathbf{a} = 0.45375N$

The results of these experiments are summarised in figures 5.39, 5.40, 5.41 and 5.42. It is clear from these graphs that, excluding the case where the aggregates should achieve an intermixed configuration starting from an intermixed lattice, only the homotypic adhesion pair with the highest magnitude of difference actually gets close to achieving the expected configurations. In each of the other three cases, the aggregates converge on states which deviate less from their original configurations (see fig. 5.45)[4]. More specifically, there is an approximately linear relationship between the deviation of the CBIs from those which are expected and the magnitude of difference in the homotypic adhesion.

Figures 5.43 and 5.44 show the full range of CBIs that were obtained for the pair with the highest magnitude of difference. Screenshots of the example final configurations are also shown in figure 5.46. These results demonstrate that in this case the behaviour of the model is a much closer approximation of DAH. The type *b*A-Cells still struggle to fully engulf the central aggregate when starting from a sorted lattice. However, it is possible that increasing the random movement forces from the low default value may allow full engulfment to take place.

Overall, these experiments confirm the earlier 2D results (see section 4.2.2) and demonstrate a strong link between the difference in the homotypic adhesion of the A-Cells and their ability to successfully sort or intermix. A discussion of the difference between the magnitude of difference required for 2D and 3D aggregates is included in the previous section 5.3.1.

[4]The convergence is not quite as rapid in some of the engulfment simulations. Therefore, it is possible that an extended time period may allow some slight additional improvement.

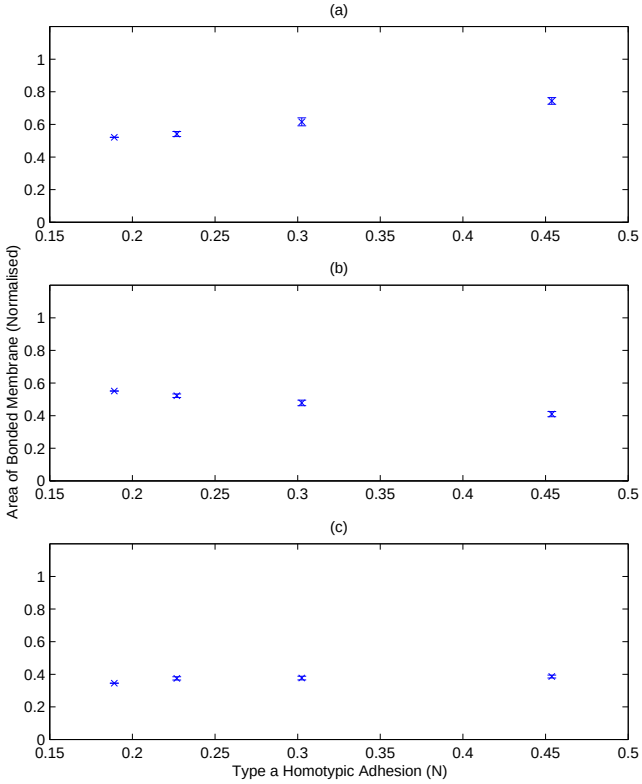


Figure 5.39: CBI *results for 64 A-Cell aggregates which should form the onion configuration from an initially mixed grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

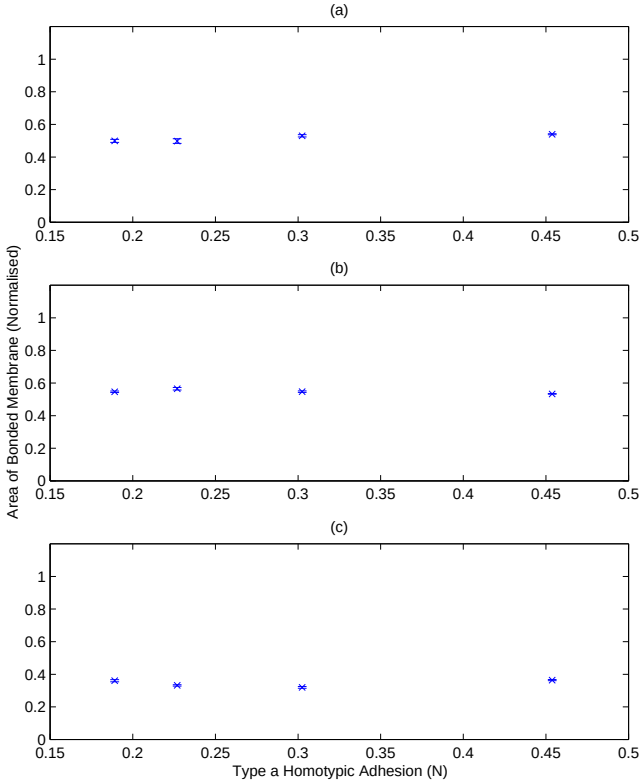


Figure 5.40: CBI *results for 64 A-Cell aggregates which should form an inter-mixed configuration from an initially mixed grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

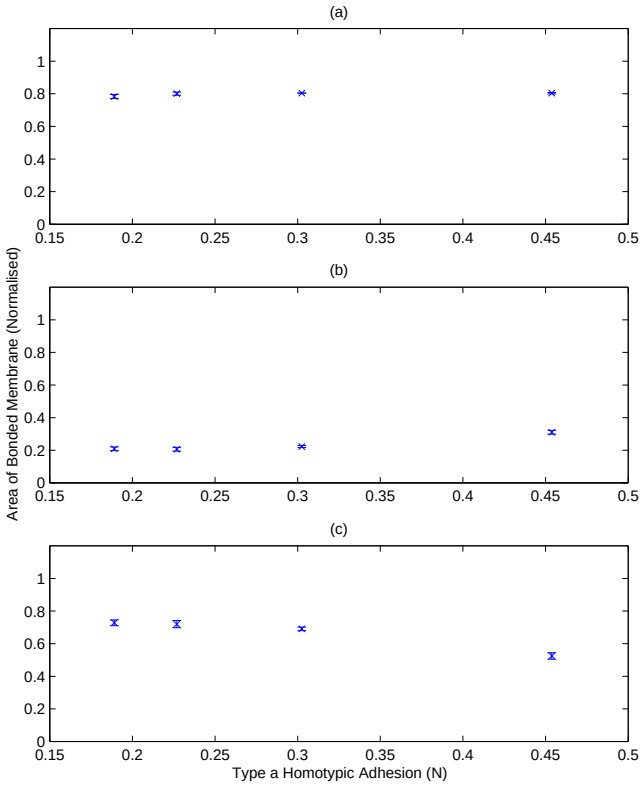


Figure 5.41: CBI *results for 64 A-Cell aggregates which should form the onion configuration from an initially sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

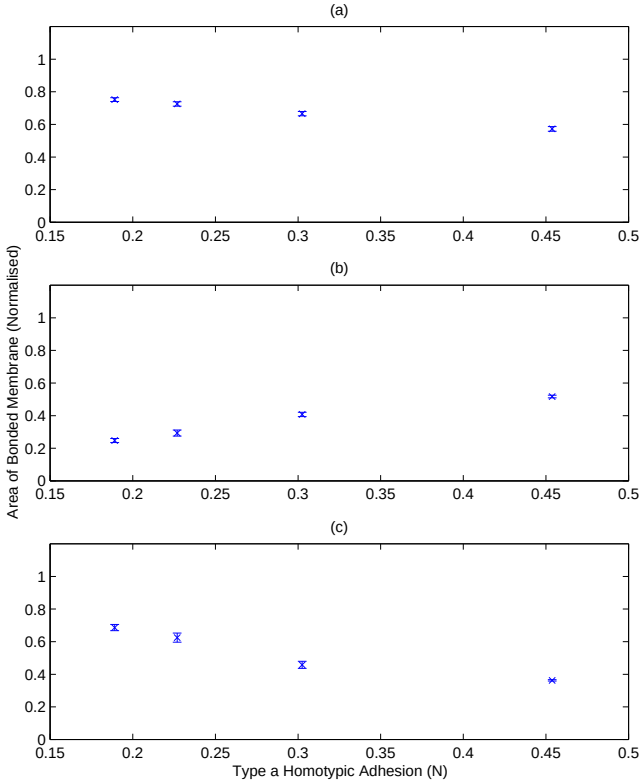


Figure 5.42: CBI *results for 64 A-Cell aggregates which should form an inter-mixed configuration from an initially sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

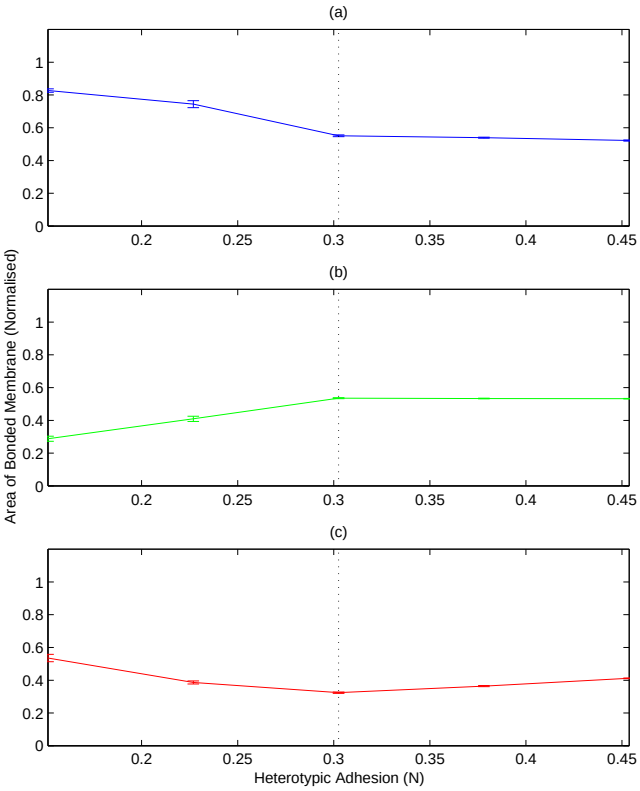


Figure 5.43: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed lattice. The homotypic adhesions of the two cell types were* $\mathbf{b} \rightarrow \mathbf{b}\ 0.15125N$ *and* $\mathbf{a} \rightarrow \mathbf{a}\ 0.45375N$*. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

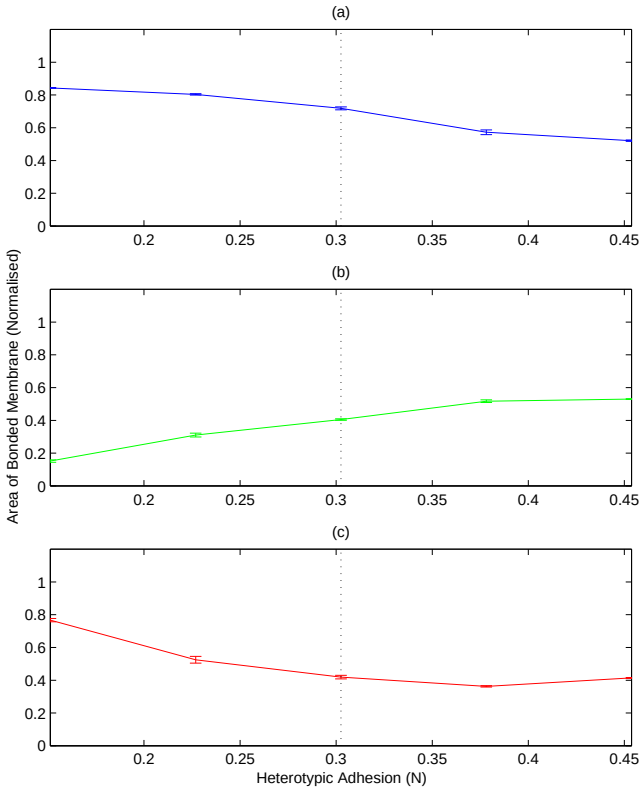


Figure 5.44: CBI *for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an sorted lattice. The homotypic adhesions of the two cell types were* $\mathbf{b} \rightarrow \mathbf{b}$ $0.15125N$ *and* $\mathbf{a} \rightarrow \mathbf{a}$ $0.45375N$*. (a)* $\mathbf{a} \rightarrow \mathbf{a}$ *(b)* $\mathbf{a} \rightarrow \mathbf{b}$ *(c)* $\mathbf{b} \rightarrow \mathbf{b}$

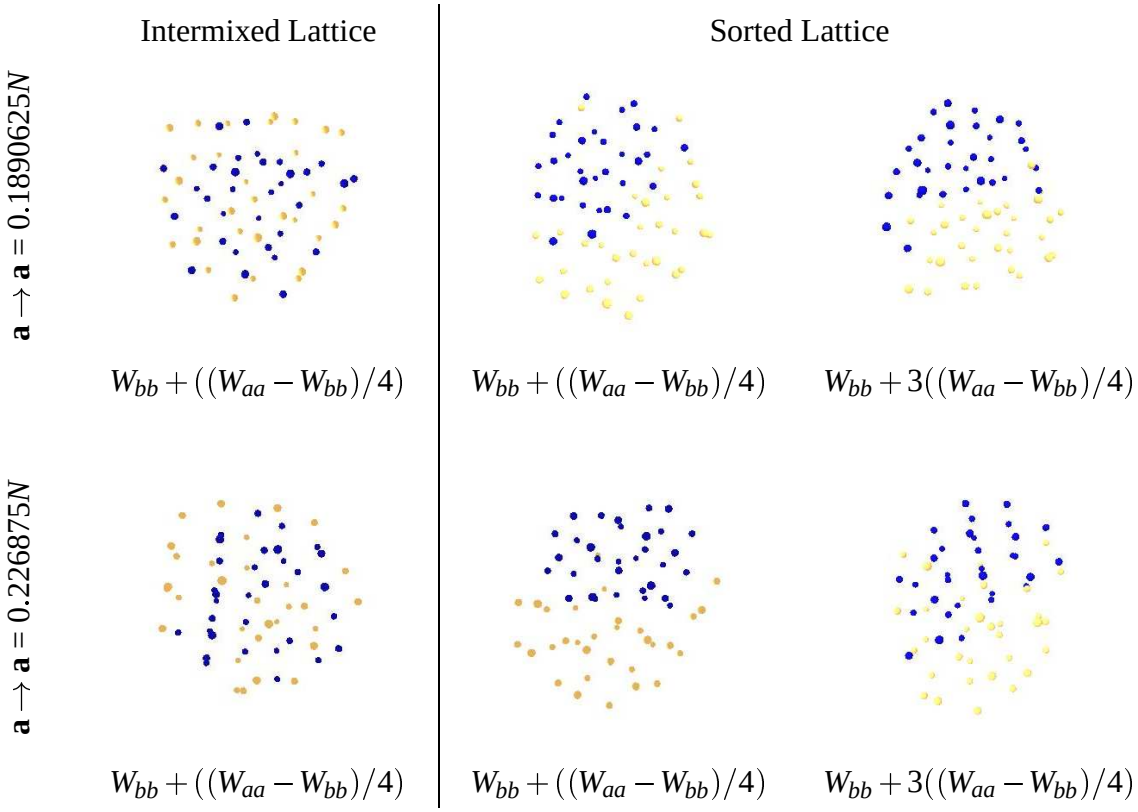


Figure 5.45: *Screen shots showing the final configuration of aggregates from several simulations. The aggregates consisted of 64 A-Cells and started from either a sorted lattice or mixed lattice initial configuration. In each case the* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion was* $0.15125N$.

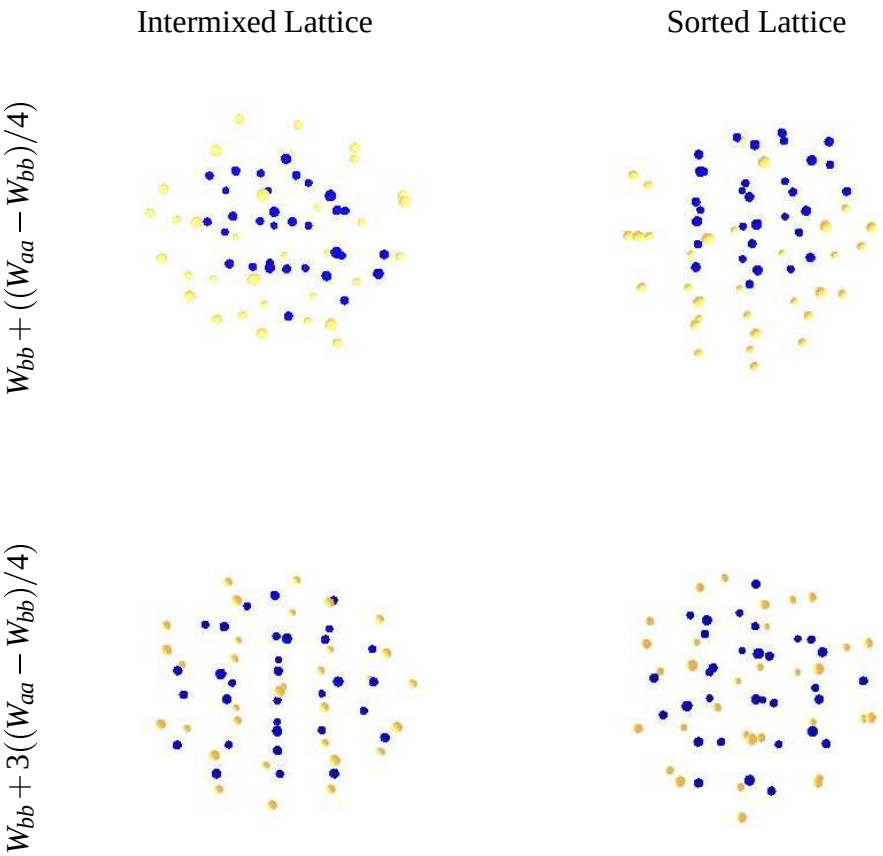


Figure 5.46: *Screen shots showing the final configuration of aggregates consisting of two A-Cell types with different heterotypic adhesion values and initial configurations.* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion was* $0.45375N$, $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion was* $0.15125N$.

5.4 Summary & Conclusions

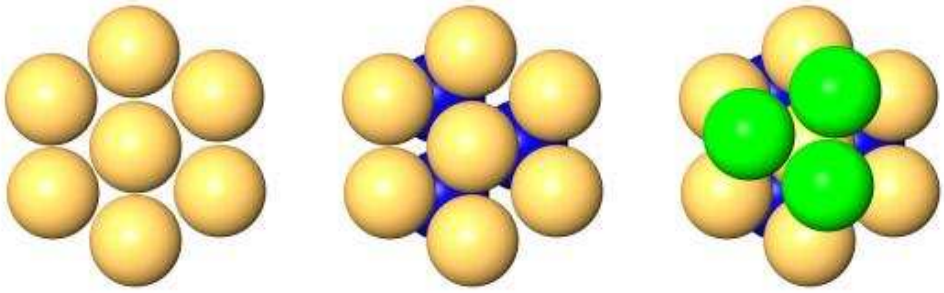
The work presented in this chapter demonstrates that the cell adhesion model translates well from 2D to 3D space. This translation results in near spherical aggregates which retain the same type of approximate close packing that was observed in the earlier 2D simulations (see fig. 5.47 & 5.48)[5].

More specifically, the rounding experiments demonstrated that, as one might expect, random forces are still fundamental to the performance of the model. However, the natural fluctuations in position generated by the rotation of the A-Cells' sensors are no longer sufficient to drive the aggregates to their equilibrium state and, instead, at least some random movement forces are required to allow sufficient exploration of the vertical space.

In addition, it was shown that the sensor FOV limits identified in section 3.5.3 still apply. Furthermore, the debilitating effects of divergent sensor coverage actually increase when the A-Cells are no longer fixed to the same plane.

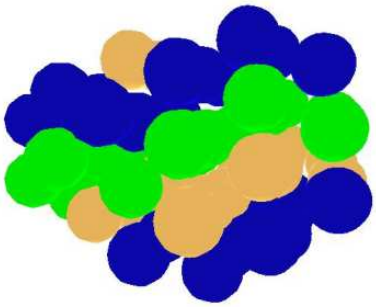
The sorting experiments which were carried out also revealed some interesting behaviours. In particular, it was shown that although the model still behaves as predicted by the DAH, the minimum threshold on the difference between the homotypic adhesions increases significantly. Therefore, it is not necessarily possible to use 2D simulations as a direct predictor of expected 3D behaviour.

[5]The cubic close packing which is observed only differers from hexagonal close packing in the positioning of the top layer 5.47.

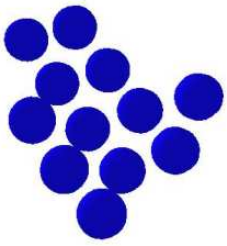


(a) (b) (c)

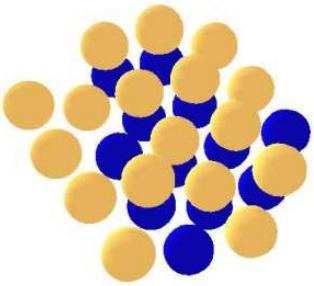
Figure 5.47: *An example of the layers found in cubic close pack.*



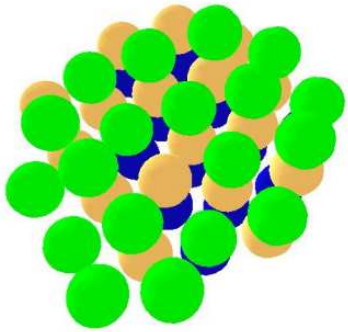
All Layers



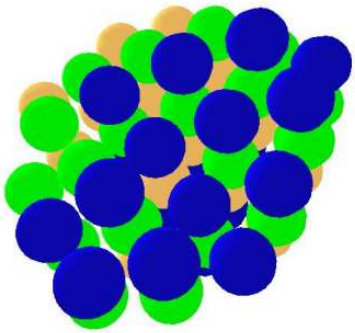
Layer 1



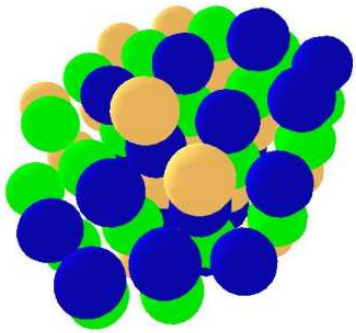
Layer 2



Layer 3



Layer 4



Layer 5

Figure 5.48: *An example of the cubic close pack layering found in a real A-Cell aggregate.*

Chapter 6

Dynamic Control

The previous chapters in this thesis demonstrate that the cell adhesion model can generate robust self-organisational dynamics. However, the A-Cells' membrane properties remain fixed and thus, more complex patterns of behaviour which require aggregates to change function over time or react to environmental stimuli are not possible.

In this section we present a hybrid system which combines the robustness of the cell adhesion model with the flexibility of a genetic regulatory network controller (GRN) to create a single robust controller which overcomes these limitations. The integration of these two systems and the accompanying experimentation and analysis were carried jointly with Taylor [61].

6.1 Genetic Regulatory Network Controller

Genetic regulation is the process responsible for the dynamic behaviour of living cells. In each cell, active genes cause the production of specific proteins in the cell's cytoplasm and in turn certain regulatory proteins control the activity of individual genes. Therefore, complex patterns of regulation can arise allowing the function of cells to change over time. In addition, some proteins diffuse between cells and the extra cellular matrix and thus, a cell's function may also be determined by its local environment.

The majority of artificial GRN systems use these dynamics to grow static structures [5, 15]. The major benefit of this approach is the ease with which suitable genomes can

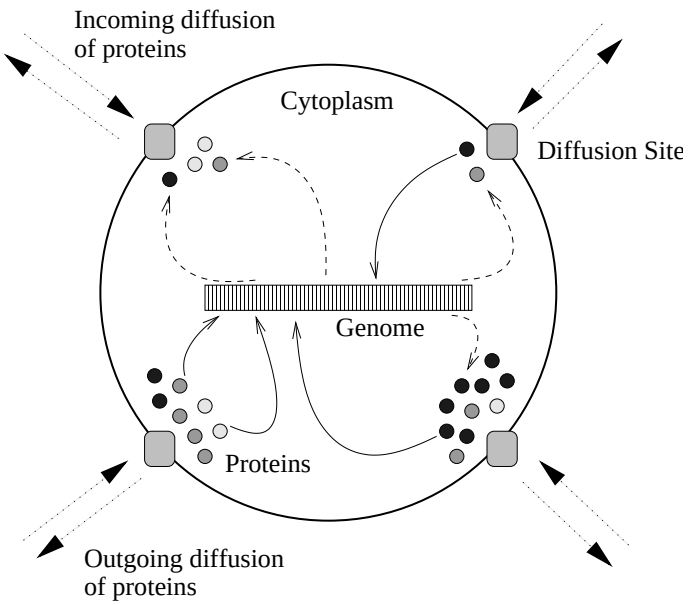


Figure 6.1: *Summary of the GRN Controller design. Figure originally published in [61].*

be evolved. However, more recent work has shown that GRNs can also be successfully adapted for use as control systems by coupling specific proteins with the sensors and actuators of the system [47]. In this section we present a brief outline of Taylor's GRN controller which develops this idea further, such that a single genome controls a group of agents to produce some desired collective behaviour [61].

In Taylor's system, the main components of the GRN controller for each agent consist of a unique region of cytoplasm and a copy of the common genome (see fig. 6.1). The cytoplasm represents the internal space of an agent and is divided into significant areas corresponding to the location of its sensors or actuators. The genome is a variable length string which may encode any number of individual genes and any inhibiting or enhancing proteins which regulate them.

When a gene becomes active, the protein it produces is deposited in a specific area of the cytoplasm. However, in addition to simply controlling further gene expression, some subsets of proteins can be assigned more specific roles. For example, some motility proteins can cause particular actuators to become active depending on the area of the cytoplasm they are produced in. Alternatively, sensory proteins may be

added to areas of the cytoplasm as a result of detected environmental conditions (e.g. light levels, distance to obstacles etc.). Finally, it is also possible for some proteins to diffuse between neighbouring agents if they are produced or detected at particular communication sites. This means that the expression of genes in each agent will be a result of both the neighbouring agents and the local environment.

To develop GRNs which are capable of producing specific behaviour, a genetic algorithm (GA) is used to evolve a population of genomes. At each cycle in the GA, every genome is evaluated by inserting it into a collection of simulated agents and comparing the resulting behaviour with the requirements of the given task. Crossover and mutation operators are then applied to generate the population for the following cycle.

6.2 Combination Controller

It is possible to combine the GRN controller with the cell adhesion model without greatly affecting the design of either system. The main change is that the motility proteins no longer interface directly with the actuation system and instead control the expression of A-CAMs on an A-Cell's membrane (see fig. 6.2). In this way, the dynamics of the cell adhesion model are still responsible for the low level A-Cell control and the GRN controller manipulates these dynamics to produce a desired collective behaviour.

To further combine the two systems, some of the sensory proteins of the GRN were also linked with the formation of particular A-CAM bonds, such that the number of bonds were directly proportional to the level of the proteins. This provides a crude method to allow the differentiation of gene expression in the A-Cells from both the number of membranes they are in contact with and the A-CAMs expressed on those membranes.

In addition, during these initial experiments, the desire was to keep the combined controller as simple as possible and thus the diffusion of proteins between A-Cells was not implemented. Therefore, the only interaction between the A-Cells was as a direct result of A-CAM expression.

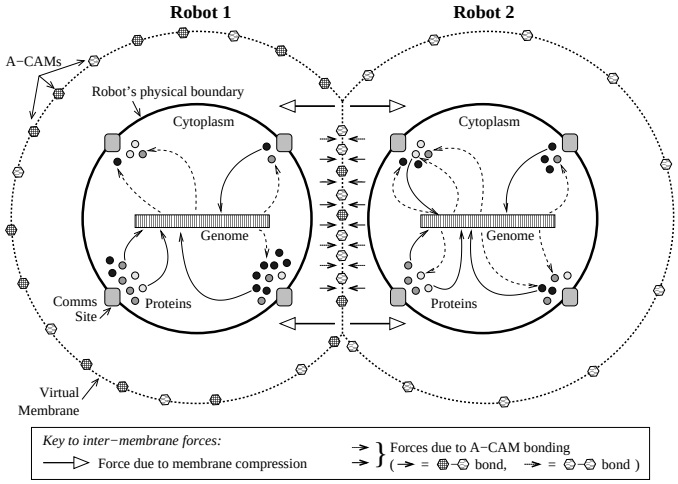


Figure 6.2: *Schematic of the combined GRN-CAM controller design. Figure produced by Taylor.*

6.3 Experiments

In this section we present a series of experiments which compare the performance of the GRN-only controller with that of the combined GRN and cell adhesion model controller. For consistency, both controllers were evaluated using the A-Cell simulator. In addition, for simplicity and speed, the movement forces of the A-Cells were restricted to the the horizontal plane.

The experiments consist of three different tasks of varying degrees of difficulty and are carried out in two sets. In the first set, the controllers are evaluated based on the ease with which they can be evolved and the quality of the final solutions. In the second set, the robustness of the evolved controllers is then evaluated by placing them in unseen environments.

6.3.1 Implementation Details

Sections 6.1 and 6.2 describe the general properties of the controllers. Here we present the more specific details of their implementation for the experiments presented in this section.

Many of the specific dynamics of the GRN match Taylor's original design.

6.3.1.1 Protein and Cytoplasm Dynamics

Both controllers provide each A-Cell in a simulation with a unique region of cytoplasm in which proteins from a predefined list may be expressed. Furthermore, within this cytoplasm four equally spaced 'diffusion sites' are specified (see fig. 6.1 & 6.2). These diffusion sites act as locations were proteins can be introduced into the cytoplasm and where the concentration of the proteins being expressed in the cytoplasm can be measured.

The proteins themselves are represented by a concentration value which ranges from 0 to 1 (giving a maximum value of 4 for the summed concentration of a specific protein across all diffusion sites). As genes in the GRN become active, they introduce more of their product protein into the cytoplasm at one or more diffusion sites (see section 6.1) causing an increase in the protein's concentration [1]. However, underlying dynamics of attenuation and diffusion are also modelled such that the protein concentrations naturally change over time.

The attenuation mechanism causes the concentration ($C_p(t)$) of any protein (p) at any diffusion site to decay over time:

$$C_p(t+1) = kC_p(t) \tag{6.1}$$

where k is the attenuation constant ($0 < k < 1$).

The diffusion process allows proteins to spread through the cytoplasm from the site were they are initially introduced (i). This causes a decrease in the concentration of the protein at i and an associated increase at each of the N neighbouring sites j:

$$C_p(i,t+1) = (1-d)C_p(i,t) \tag{6.2}$$

$$C_p(j,t+1) = C_p(j,t) + \frac{dC_p(i,t)}{N} \tag{6.3}$$

where d is the internal diffusion constant ($0 < d < 1$).

[1] The increase in concentration is directly linked to the level of activation of the gene. For example, a gene with an activation level of 0.5 will cause a 0.5 increase in the product protein's concentration value.

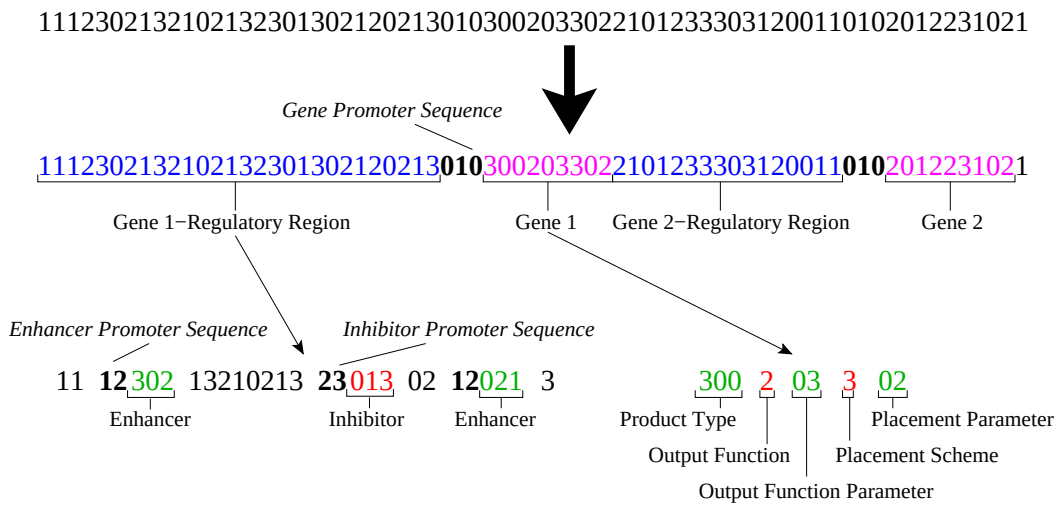


Figure 6.3: *A section of an example base 4 genome. In the example gene 1 is inhibited by protein 7 (013 in base 4) and enhanced by proteins 50 (302) and 9 (021). The specification of the gene shows that it produces protein 48 (300) when expressed. It has a Gradient Above Zero output function (2) with gradient 3 (03). In addition, the product is deposited at a specific diffusion site (3), which is site number 2 (02).*

6.3.1.2 Genome and Gene Structure

In both controllers the GRN which governs the A-Cell behaviour is encoded by a variable length string of base 4 integers. To construct the GRN the string is first divided into individual genes and their associated regulatory region. This is achieved by identifying a specific sequence in the genome known as the gene promoter sequence (see fig. 6.3). The gene itself is represented by a fixed number of digits immediately following the promoter sequence and its regulatory region extends from the end of the previous gene up to, and including, the digit preceding the promoter sequence.

A gene's regulatory region identifies those proteins which affect the activation of the gene. This includes enhancer proteins, whose presence in the cytoplasm increases the gene's activation, and inhibitor proteins, whose presence have the opposite effect. The proteins are encoded by a fixed number of digits which immediately follow either an enhancer or inhibitor promoter sequence. Therefore, by scanning the region from

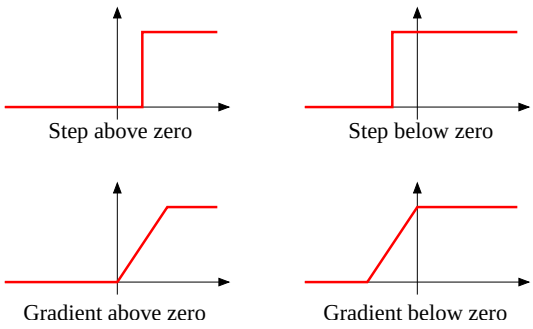


Figure 6.4: *The four possible activation functions.*

beginning to end a full list of the gene's regulatory proteins can be constructed.

The substring of digits representing the gene are used to encode its specific characteristics. This includes the relationship between the total concentration of the gene's regulators and its level of activation (the output function), the type of protein the gene manufactures, and where this product protein is introduced into the cytoplasm (the product placement scheme).

In this work the following four output functions are made available (see fig. 6.4) [2].

1. Step above zero

$$A_g \begin{cases} = 1, & \text{if } R_g > P_g \\ = 0, & \text{if } R_g <= P_g \end{cases} \tag{6.4}$$

2. Step below zero

$$A_g \begin{cases} = 1, & \text{if } R_g > -P_g \\ = 0, & \text{if } R_g <= -P_g \end{cases} \tag{6.5}$$

3. Gradient above zero

$$A_g \begin{cases} = 0, & \text{if } R_g <= 0 \\ = 1, & \text{if } R_g >= P_g \\ = \frac{R_g}{P_g}, & \text{if } 0 < R_g < P_g \end{cases} \tag{6.6}$$

[2] As the string consists of base 4 integers this allows the output function to be defined by a single digit.

4. Gradient below zero

$$A_g \begin{cases} = 0, & \text{if } R_g <= -P_g \\ = 1, & \text{if } R_g >= 0 \\ = 1 + \frac{R_g}{P_g}, & \text{if } -P_g < R_g < 0 \end{cases} \tag{6.7}$$

where A_g is the level of activation of gene g ($0 <= A_g <= 1$)and R_g is the concentration of its enhancer proteins ($\mathcal{E}_g$) minus the concentration of its inhibitor proteins (I_g), summed across all diffusion sites ($\mathcal{D}$).

$$R_g = \sum_{d \in \mathcal{D}} \sum_{e \in \mathcal{E}_g} Conc(e,d) - \sum_{i \in I_g} Conc(i,d) \tag{6.8}$$

The parameter P is the output function parameter which is used to refine the output function's behaviour. This parameter is encoded in the gene alongside the output function (see fig. 6.3).

Similarly, four placement placement schemes are also available:

1. place the product at the diffusion site with the highest concentration of a specific marker protein;
2. place the product at the diffusion site with the lowest concentration of a specific marker protein;
3. place the product at a specific diffusion site;
4. distribute the product evenly across all the diffusion sites.

Once more an additional parameter (the product placement parameter) (see fig. 6.3), is used to refine the overall behaviour, in this case specifying the specific site or marker protein as appropriate.

6.3.1.3 Protein Actions

To provide a rich environment for the GRNs to operate on 64 different proteins are specified[3]. Each of these proteins is free to participate in gene regulation. However,

[3]This allows each protein in the genome to be encoded with 3 digits.

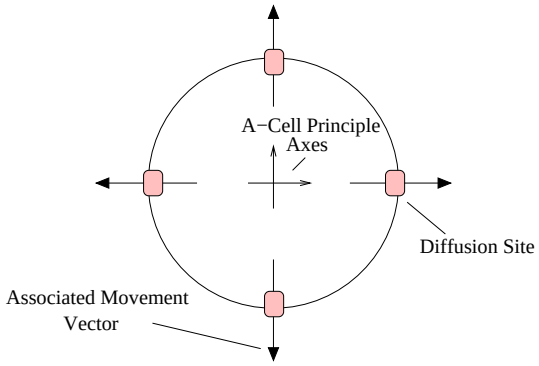


Figure 6.5: *Movement vectors associated with the diffusion sites in the GRN-only controller.*

as discussed in section 6.1, subsets of the proteins are also associated with the specific sensing and actuation mechanisms that are available to the controllers.

The GRN-only controller interfaces directly with the A-Cell sensors and actuators as in Taylor's original design [61]. For actuation, each diffusion site is linked with movement in a specific direction along the principle axes of the A-Cell (see fig. 6.5). These movement vectors (M) are then scaled by the summed concentration of 16 predefined actuation proteins ($\mathcal{A}$) at the respective diffusion site giving a desired movement in each direction. A final movement vector (F) is then calculated as the average of these 4 scaled values:

$$F = \sum_{d \in \mathcal{D}} M_d \sum_{a \in \mathcal{A}} Conc(a,d) \tag{6.9}$$

The sensing mechanisms made available to the GRN-only controller include external protein diffusion and a local neighbours sensor. The external diffusion mechanism provides crude information about the distance and direction of neighbouring A-Cells by using the local communication system to allow a specific protein (p) to diffuse between them. This is achieved by associating each diffusion site with one of the 4 transmitter/receiver pairs located on the A-Cell hull. When the externally diffusing protein is expressed at one of the sites (i), a signal is broadcast from the associated transmitter to indicate the amount of the protein that is diffusing from the site (E). The local concentration of the protein (C_p) is also adjusted accordingly:

$$C_p(i,t+1) = (1-e)C_p(i,t) \tag{6.10}$$

$$E_p(i,t) = eC_p(i,t) \tag{6.11}$$

where e is the external diffusion constant ($0 < e < 1$).

Likewise, when one of the sensors detects an incoming signal from a neighbouring A-Cell, the level of the protein is increased at the sensor's associated diffusion site (j). The increase is a function of the concentration value encoded in the signal and the distance (d) that the signal has travelled:

$$C_p(j,t+1) = C_p(j,t) + \frac{k_t}{d^2}E_p(i,t) \tag{6.12}$$

where k_t is the predefined transmission coefficient (0.0009 in these experiments).

The local neighbours sensor increases the concentration of a specific protein at all the diffusion sites, by an amount proportional to the number of A-Cells (n) within communication range:

$$C_p(t+1) = C_p(t) + \frac{n}{N} \tag{6.13}$$

where N scales the value appropriately[4]. This sensor is included to help approximate the sensing abilities of the combined controller such that both controllers have reasonably equivalent abilities.

As discussed in section 6.2 in the combined GRN and cell adhesion model controller the GRN is not responsible for the low level A-Cell control. Instead, 12 proteins are associated with the level of expression of unique A-CAMs on the A-Cell's membrane. Therefore, the GRN actually varies the adhesion between neighbouring A-Cells and any real actuation is generated as a result of modelling the adhesions. The bonding relationships between the A-CAMs are predefined such that 4 of them are capable of homotypic bonding while the remaining 8 are capable of heterotypic bonding. This provides 8 different bonding relationships for the GRN to exploit.

[4]This value can be manually tuned to adjust the sensitivity. A value of 4 was used in these experiments.

GRN-Only Controller		Combined Controller	
Proteins	*Description*	*Proteins*	*Description*
0	External diffusion protein		
32-47	Actuation proteins	36-47	A-CAM expression proteins
48	Local neighbours sensor	48-59	A-CAM bonding proteins

Figure 6.6: *Summary of the specific protein actions in each controller.*

To map each protein (p) to some level of A-CAM expression (A_p), its average concentration is calculated across all 4 diffusion sites (as A-CAMs are distributed evenly across the whole membrane). This gives a value between 0 and 1 which is then multiplied by the maximum level of A-CAM expression that is allowed (m)[5].

$$L(A_p,t) = m\frac{1}{|D|}\sum_{d\in\mathcal{D}} C_p(t) \tag{6.14}$$

In a similar fashion, sensing in the combined model is achieved by linking the 12 A-CAM types to the concentration of 12 unique proteins. When one of the A-CAM types bonds with an A-CAM expressed on another A-Cell's membrane, the protein's concentration is increased by a level proportional to the number of bonds:

$$C_p(t+1) = C_p(t) + \frac{B(A_p)}{m} \tag{6.15}$$

where $B(A_p)$ is the number of bonds formed by the A-CAM linked to protein p.

The specific protein actions for each controller have been summarised in table 6.6.

6.3.1.4 Genetic Algorithm

The controllers were evolved using a standard generational GA. The initial populations consisted of 60 random genomes with a string size of 1000 and subsequent generations were generated by using tournament selection and elitism with group sizes of 3 and 1

[5]Given two A-Cells each expressing a single type of A-CAM, such that the two A-Cells adhere, the maximum level of expression was calculated as the lowest value which, if expressed on both A-Cells, would cause them to physically collide.

respectively. The offspring genomes were formed by using both crossover and mutation. The crossover operator was simple two-point crossover and different crossover points were selected on each parent to allow the size of the genomes to evolve over time. The mutation operator was applied to each digit in the resulting genomes with some fixed probability P_{TM} ($= 0.002$ in the reported experiments), replacing the digit with a random value. However, instead of simply relying on chance mutations of gene regulators and output functions, an additional form of smart mutation was also used[6]. After traditional crossover and mutation this new operator was applied to each genome with a fixed probability P_{SM}^{Genome} ($= 0.75$).

The main principle of the smart mutations is that by using knowledge about the GRN (recorded during each simulation), mutations can be made which are more likely to have some observable effect. In total, three different types of smart mutation were applied:

1. **mutation of the regulators** to ensure each regulator could potentially affect the gene's activation: if the observed maximum concentration of a regulator was zero, the regulator is replaced with another chosen at random from a list of those known to have non-zero concentration during the controller's operation.

2. **mutation of the output function** to align the gene's range of sensitivity with the sum of the regulators: if the gene has an "above zero" output function (Step above zero or Gradient above zero) and the maximum observed sum of the regulators R_{obs}^{Max} is less than zero, the output function is switched to the corresponding "below zero" function. However, if R_{obs}^{Max} is greater than the threshold specified in the gene, it is replaced with a random value between zero and R_{obs}^{Max}. The reverse procedure is used for genes with "below zero" output functions.

3. **mutation of marker proteins (used in product placement)** to ensure that marker proteins which achieve non-zero concentrations are used: if the product placement scheme specifies that the product protein should be placed at a diffusion site specified by another protein's concentration and that protein's maximum concentration does not exceed zero, the signal protein is replaced with

[6]This idea was developed by Tim Taylor who is currently investigating it further.

one randomly chosen from a list of those known to have non-zero concentration during the controller's operation.

When the smart mutation operator is applied to a genome each of its genes are considered in turn and the different forms of mutation are applied with probability P_{SM}^{Gene} $(= 0.75)$.

6.3.2 Tasks

The tasks outlined below are presented in order of difficulty. In each case, a sample of five evolutionary runs were carried out for both the GRN-only controller and the combined controller. Unfortunately, the evaluation time required for each individual run (several days when running in parallel on a dedicated cluster) meant that a larger sample was not feasible. However, once more, bootstrapping was applied to improve the error estimates.

During the execution of each task it is clearly desirable that, at the very least, the A-Cells maintain a single connected aggregate [7]. Therefore, the fitness value awarded to each controller (F_t) was divided into two separate components. The first of these (F_1), simply reflects the proportion of the simulation during which the connected aggregate was maintained:

$$F_1 = \frac{T_M}{T_E} \tag{6.16}$$

where T_E is the total evaluation time and T_M is the time the connected aggregate was maintained for. The second (F_2), was the fitness score achieved in relation to the specific task, and therefore was only awarded when the aggregate was maintained for the entire simulation period. These components were combined with the relative weightings of 0.1 and 0.9 respectively, giving the two following scenarios.

1. Aggregate not maintained:

$$F_t = 0.1F_1 \tag{6.17}$$

[7] In the combined controller a connection is achieved through membrane contact whereas in the GRN-only controller it implies that A-Cells are within communication range. During these experiments the two distances are equivalent.

2. Aggregate maintained for the duration of the simulation ($F_1 = 1$):

$$F_t = 0.1 + 0.9F_2 \tag{6.18}$$

In addition, to determine a more general fitness value, each controller was evaluated from three different initial configurations. Each of these was selected from a population of randomly generated configurations in which the aggregates were fully connected (see appendix A.2), such that they had compactness values 0.4, 0.5 and 0.6. Once all three fitness values were calculated, the average result was taken.

6.3.2.1 Clustering

The first task tests the controllers' ability to perform simple aggregation or rounding. As in previous chapters (see section 3.1.1.1), this is measured using the compactness index C (see equation 3.2)[8]. However, to encourage the controllers to achieve maximum compactness as soon as possible, the final fitness is calculated as the mean compactness of the aggregate over the entire $100s$ evaluation period (T_E).

$$F_2 = \frac{1}{T_E}\sum_{i=0}^{T_E} C_i \tag{6.19}$$

6.3.2.2 Response to External Signal

This task was designed to investigate whether the controllers could easily switch between different functions in response to some environmental trigger. In practise, such a trigger could take the form of a change in any environmental condition that the A-Cells are able to detect and would result in a corresponding change in the their relative sensory proteins. Therefore, this was represented in the simulations by increasing the concentration of a small number of proteins in each A-Cell's cytoplasm at some random time T_S, within a predefined period. The fitness value for the task was based on the A-Cell aggregate's ability to demonstrate different behaviours before and after the

[8] In the case of the GRN-only controller, the concave hull is calculated using the communication range as a measure of connectivity.

signal was introduced. In this case, initially the aggregate was rewarded for maintaining a constant minimum enclosed area. However, once the signal was introduced, it was instead rewarded for maintaining a constant maximum enclosed area[9]. The fitness function therefore had two components relating to the difference between the area values D and the variance V of each period of the simulation.

$$m_a = \frac{1}{T_S}\sum_{i=0}^{T_S} a_i \qquad (6.20) \qquad m_b = \frac{1}{T_E - T_S}\sum_{i=T_S+1}^{T_E} a_i \qquad (6.21)$$

$$D = \frac{m_b}{m_a + m_b} \qquad (6.22)$$

where a_i is the area of the aggregate's concave hull at time i and T_E is the full evaluation time of the simulation.

$$v_a^2 = \frac{1}{T_S}\sum_{i=0}^{T_S}(a_i - m_a)^2 \qquad (6.23) \qquad v_b^2 = \frac{1}{T_E - T_S}\sum_{i=T_S+1}^{T_E}(a_i - m_b)^2 \qquad (6.24)$$

$$V = \frac{min(v_a v_b, k_v m_a m_b)}{k_v m_a m_b} \qquad (6.25)$$

where the constant k_v is used to control the maximum level of variance that impacts on the reward (0.25 in these experiments). As both D and V have different levels of importance the final fitness is defined as:

$$F_2 = k_f D + (1 - k_f)(1 - V) \qquad (6.26)$$

where k_f determines the relative weighting.

In the simulations presented in this work the evaluation period was set to 100s and the signal was introduced at a random time in the period $50 \pm 20s$. In addition, the control parameter k_f was set to 0.75 to add extra weight to the level of change in the aggregate's behaviour.

[9] If there is no notion of maintaining a constant configuration, the controllers tend to exploit the dynamics of the simulation to produce solutions where the behaviour is fine tuned to the evaluation period T_E.

6.3.2.3 Differentiation of Function

Differentiation is the process which allows a collection of initially homogeneous agents to divide into a number of distinct subsets, each of which is capable of performing a more specific function. This aim of this task was to determine whether the controllers could achieve this by splitting an aggregate of A-Cells into two equally sized sets ($\mathcal{A}$ and $\mathcal{B}$), each exhibiting a unique pattern of protein expression. In the actual simulations, the patterns that were selected involved two ranges of 16 proteins α and β, such that one pattern was achieved by maximising the concentration of the α proteins while minimising the concentration of the β proteins and the second pattern was achieved by performing the converse. To encourage the population to differentiate as quickly as possible, the mean difference between these concentrations during the entire evaluation period was used to calculate the final fitness value. Therefore, it was possible to determine the degree to which any single A-Cell was a member of either $\mathcal{A}$ or $\mathcal{B}$ using the following d_r values:

$$g_a = \frac{1}{U|\alpha|}\sum_{t=0}^{T_E}\sum_{i\in\alpha} c_i^t \qquad (6.27) \qquad\qquad g_b = \frac{1}{U|\beta|}\sum_{t=0}^{T_E}\sum_{i\in\beta} c_i^t \qquad (6.28)$$

$$d_r = \frac{1}{2}\left(g_a + (1 - g_b)\right) \qquad (6.29)$$

where c_i^t is the cytoplasmic concentration of protein i at time t, and U is a normalisation constant representing the maximum summed concentration of a single protein over the entire evaluation period. Therefore, both g_a and g_b and thus d_r will be in the range $\{0 \rightarrow 1\}$, with:

$$d_r \begin{cases} > 0.5, & \text{if there are more set } \alpha \text{ proteins than set } \beta \\ = 0.5, & \text{if both sets of proteins are equal} \\ < 0.5, & \text{if there are more set } \beta \text{ proteins than set } \alpha \end{cases} \qquad (6.30)$$

Thus, the A-Cells can be assigned to $\mathcal{A}$ and $\mathcal{B}$, such that $\mathcal{A}$ consists of those with the highest 50% of the d_r values. The overall fitness is then defined as:

$$F_2 = \frac{1}{n}\left(\sum_{i\in\mathcal{A}} d_i + \sum_{j\in\mathcal{B}} (1-d_j)\right) \tag{6.31}$$

Therefore, if the A-Cells are completely undifferentiated (all the d_r values are the same), $F_2 = 0.5$. However, if as desired the A-Cells in $\mathcal{A}$ express more of the set α proteins than set β proteins and the reverse is true for the A-Cells in $\mathcal{B}$, $F_2 > 0.5$ indicating that the A-Cells have partially differentiated.

6.3.3 Controller Evolution

The controller evolution results for both controllers are shown in figure 6.7. In each case, the mean and maximum fitness values are given for each generation of a full evolutionary run. Most significantly, the results show that the combined controller achieved better final fitness values for each of the three tasks. In addition, it also appears far easier to find solutions for the combined controller which often starts at generation 0 with a higher fitness than the best controllers evolved for the GRN by itself. Additionally, the performance of the combined controller is more consistent, as indicated by the smaller error bars.

The difference in performance is most clearly marked in the initial clustering task. In this case, the combined controller can solve the task, to some degree, by simply expressing any two A-CAMs which adhere. The only real challenge for the GA is to find a GRN which sets the levels of A-CAM expression that generate the most appropriate level of adhesion for the environment[10]. This is particularly visible in the results which show that even at generation 0 the best individual achieved a very high fitness which was then increased only slightly over the remainder of the run.

In the signal response task, the combined controller again finds a solution relatively quickly. However, the benefit of combining the controllers is more obvious as the GA is able to find progressively better solutions during the course of the run. The introduction of the signal can clearly be identified on the protein trace (see fig. 6.8)

[10] In practise, as the combined controller does not require any dynamic behaviour to achieve this task, the GRN actually adds unnecessary complexity. Instead, a GA could simply be used to directly set the levels of A-CAM expression which could remain fixed for the duration of the simulation.

as the sudden production of the signal proteins (protein numbers 1-15) 40 seconds into the simulation. This signal causes a change in gene expression and the GRN quickly settles into a new stable state. In this new state, there is a clear change in the expression of proteins, particularly those proteins involved in the production of A-CAMs (protein numbers 36-47). Figure 6.10 shows screen shots taken at regular intervals during the simulation the traces were taken from. Viewing a simulation of the best solution for the GRN-only controller, no noticeable change in behaviour is apparent and in many cases the aggregate even fails to remain connected. The traces do show a momentary change in behaviour as a result of the signal (see fig. 6.9). However, the GRN quickly reverts to its earlier pattern of expression.

Finally, in the differentiation task, it can be seen that the combined controller rapidly achieves differentiation, as indicated by the fitness in excess of 0.55 [11]. However, the GRN-only controller fails to find any solution. In addition, the results show that although the combined controller does continue to improve, the overall increase in fitness is actually very small. This is mainly due to the very large upper bound (U) used in the fitness calculation (see section 6.3.2.3). In actual fact, due to an error, the U value that was used while the results were recorded was double the size of the true maximum. Therefore, the maximum achievable fitness in the task was actually 0.775.

6.3.4 Controller Quality

To evaluate the robustness of the evolved controllers, the best combined controller and the best GRN-only controller obtained for each of the previous tasks (see section 6.3.3) were then subjected to the following three scenarios. For each scenario, each of the controllers was again evaluated for its respective task while the simulation setup was modified as described. A sample of 20 runs were carried out for each of the parameter values.

1. **Scalability**: To determine whether the evolved controllers were scalable, they were tested with a number of aggregate sizes consisting of 12-60 A-Cells. In

[11] Using only the differentiation fitness function a value of > 0.5 indicates differentiation. However, in the final fitness value this increases to 0.55 (e.g. $(0.5 * 0.9) + 0.1$ (see equation 6.18)).

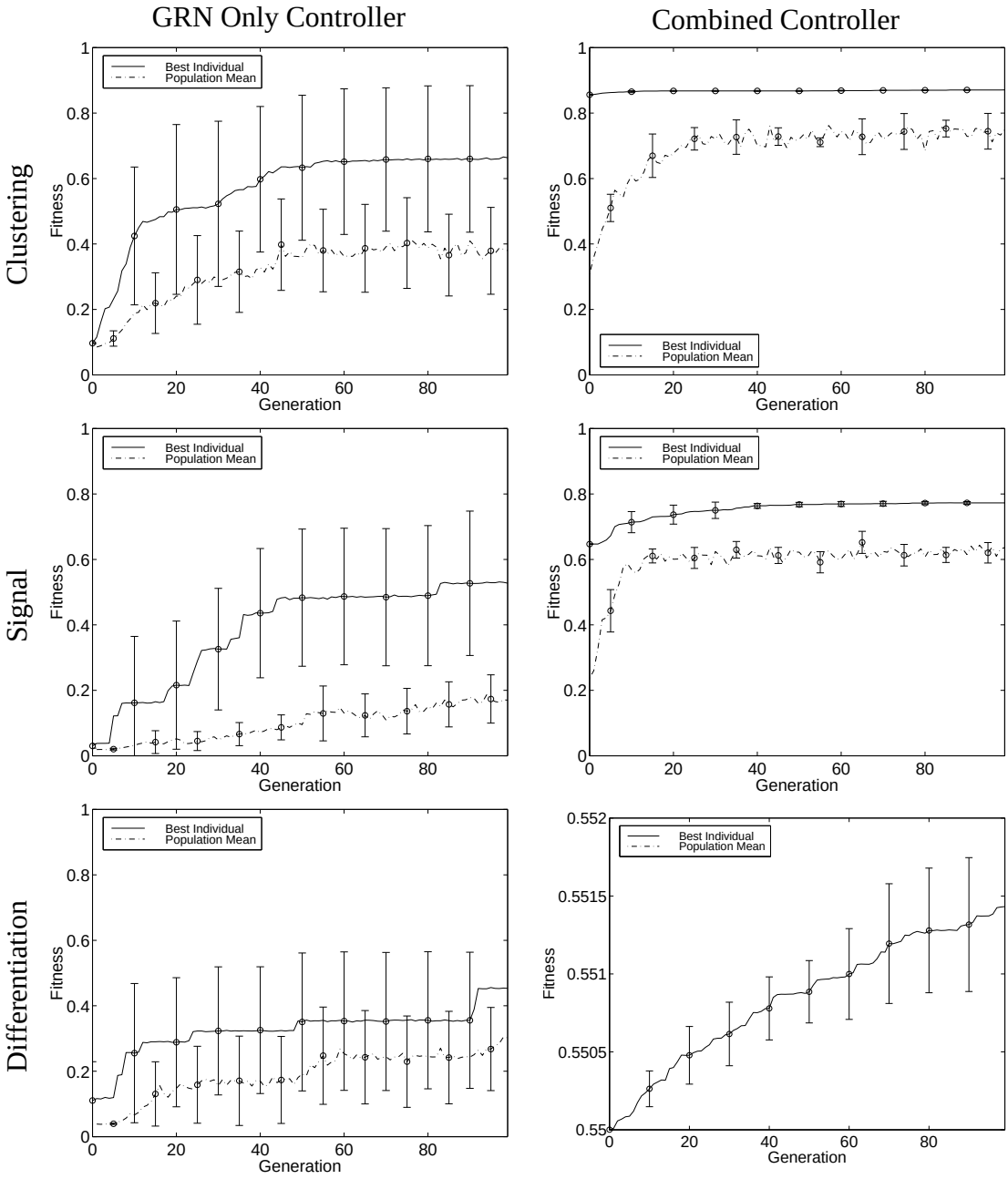
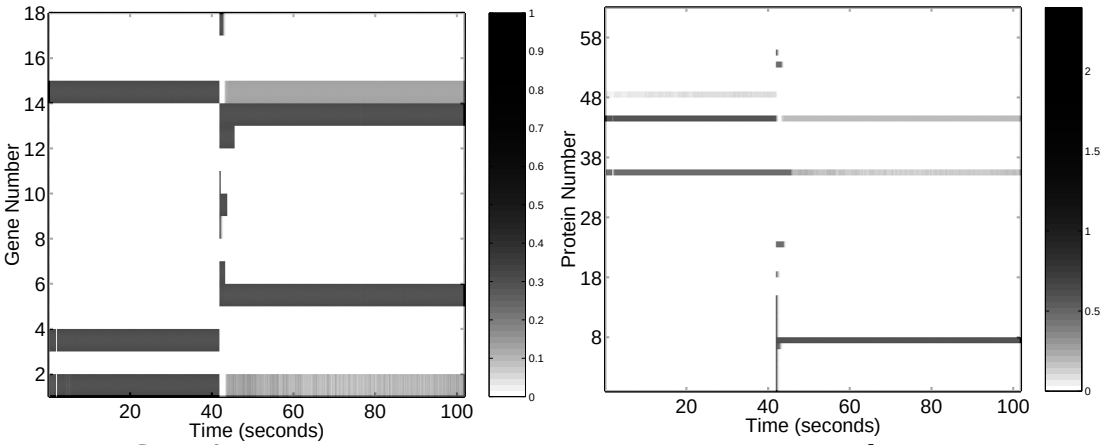
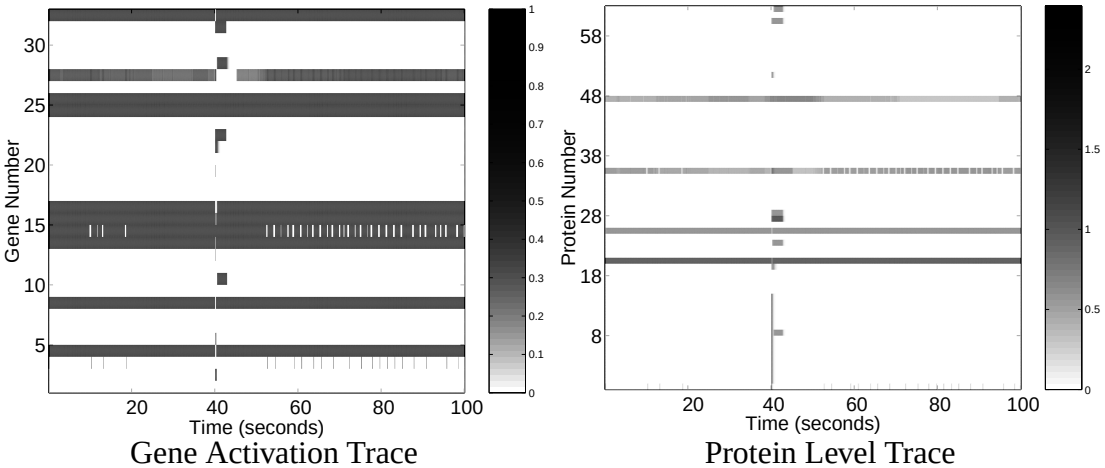


Figure 6.7: Mean and maximum fitness over time for each of the six evolution conditions.



Gene Activation Trace Protein Level Trace

Figure 6.8: *Traces obtained from a genetic regulatory network which was evolved to carry out the signal response task with the combined controller.*



Gene Activation Trace Protein Level Trace

Figure 6.9: *Traces obtained from a genetic regulatory network which was evolved to carry out the signal response task with the GRN-only controller.*

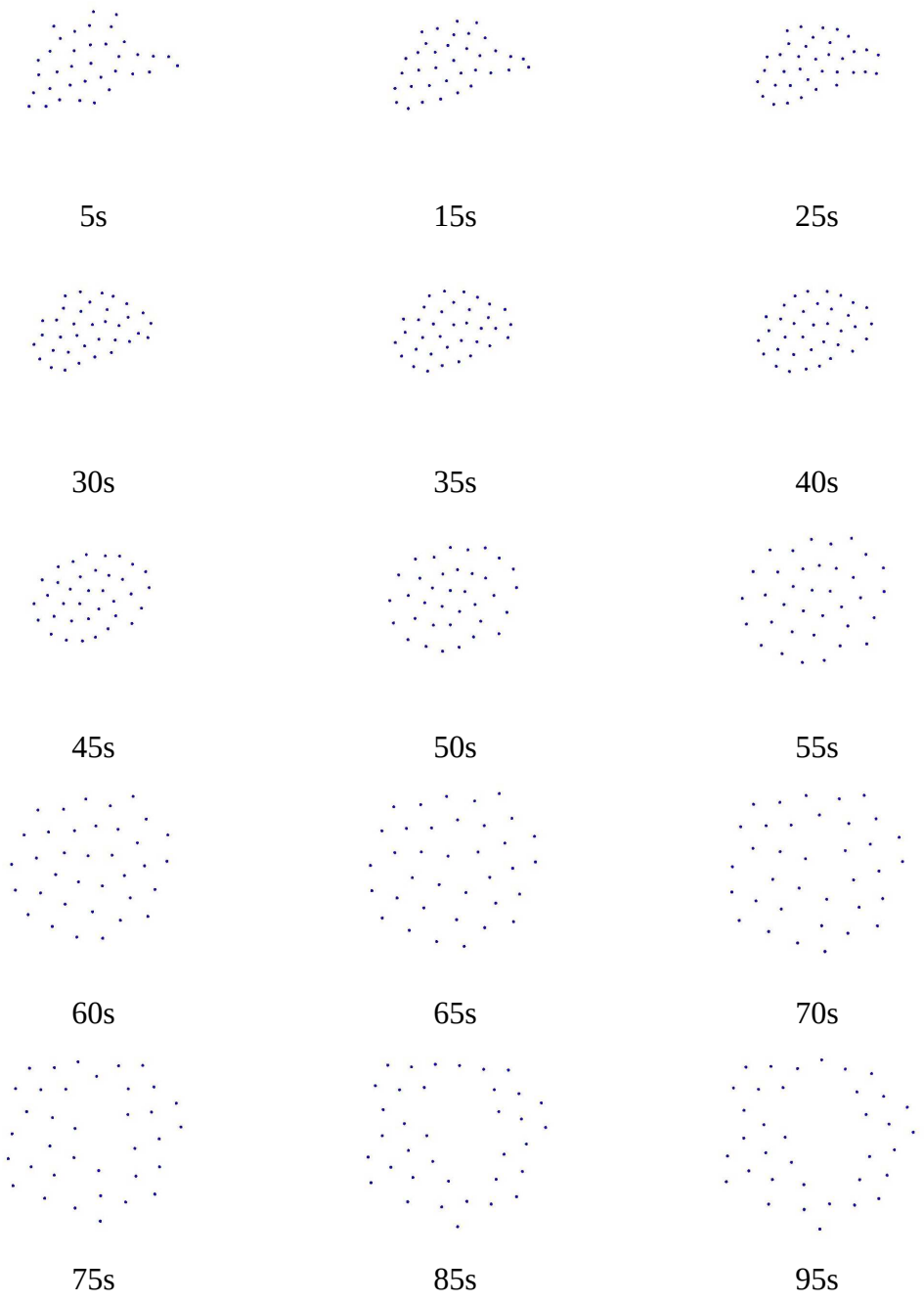


Figure 6.10: *Screen shots of the combined controller carrying out the signal response task. The signal was introduced at exactly* 40*s.*

each case, a random initial configuration was selected such that the compactness was equal to 0.5 ± 0.01 (the mid-range of those used during evolution).

2. **Specificity of Initial Configuration**: This scenario evaluated the controllers using unseen initial configurations with a wide range of compactness values $0 \rightarrow 0.8$, to determine if their performance was specific to the configurations used during evolution.

3. **Random Forces**: In this final scenario, the controllers were tested in a variety of environmental conditions which differed from those used during evolution. As in the earlier cell adhesion model experiments (see section 3.3), three different tests were carried out: (1) Small random movement forces only, (2) small random torque forces only, (3) small random movement forces in the presence of a fixed random torque.

The results of these robustness experiments are shown in figures 6.11, 6.12 and 6.13. Again, the combined controller outperformed the GRN-only controller and in most cases it is clear that the scenarios had a significantly smaller effect on its performance. In addition, the combined controller still shows far greater consistency, significantly so for both the clustering and differentiation tasks.

The results show most clearly that the significant difference between the models appears to be their ability to cope with environmental changes. The different conditions have very little effect on the combined controller. However, when the level of the random movement forces increases above those that were present during evolution, the GRN-only controller's performance begins to deteriorate. In addition, the presence of any torque appears to reduce the performance of the GRN-only controller (as can be seen by the difference between the simulations with only random movement forces and random movement forces with a small fixed torque) and an increase in torque has a disastrous effect. This is possibly due to a time lag in the GRN between the detection of externally diffusing proteins and the production of actuation proteins. For example, if an A-Cell detects a neighbour in one direction and it rotates before the appropriate actuation proteins are produced, it will head off in the wrong direction.

There is also a significant effect on both types of clustering controller when the initial configuration is changed. However, this is understandable given the task as the robots will require a longer time to cluster from an initially less compact configuration.

6.3.5 Discussion

It has been demonstrated that by combining the GRN controller with the cell adhesion model, it is possible to improve upon each individual system's performance. By adding the cell adhesion model (which already implements a robust self-organising behaviour) to the GRN controller, it removes the need for the GRN to interface directly with the robot's sensors and actuators. This appears to make the evolution of effective controllers easier and improves the overall robustness of the controllers that are evolved (in particular their ability to cope with environmental perturbations). In theory this would also make it possible to transfer the same controller to any robot which implements the cell adhesion model.

Additionally, adding the GRN controller to the CAM controller allows both temporal and spatial differentiation of A-CAM expression across the population of robots. This allows more complex tasks to be performed with a population of (initially) purely homogeneous robots.

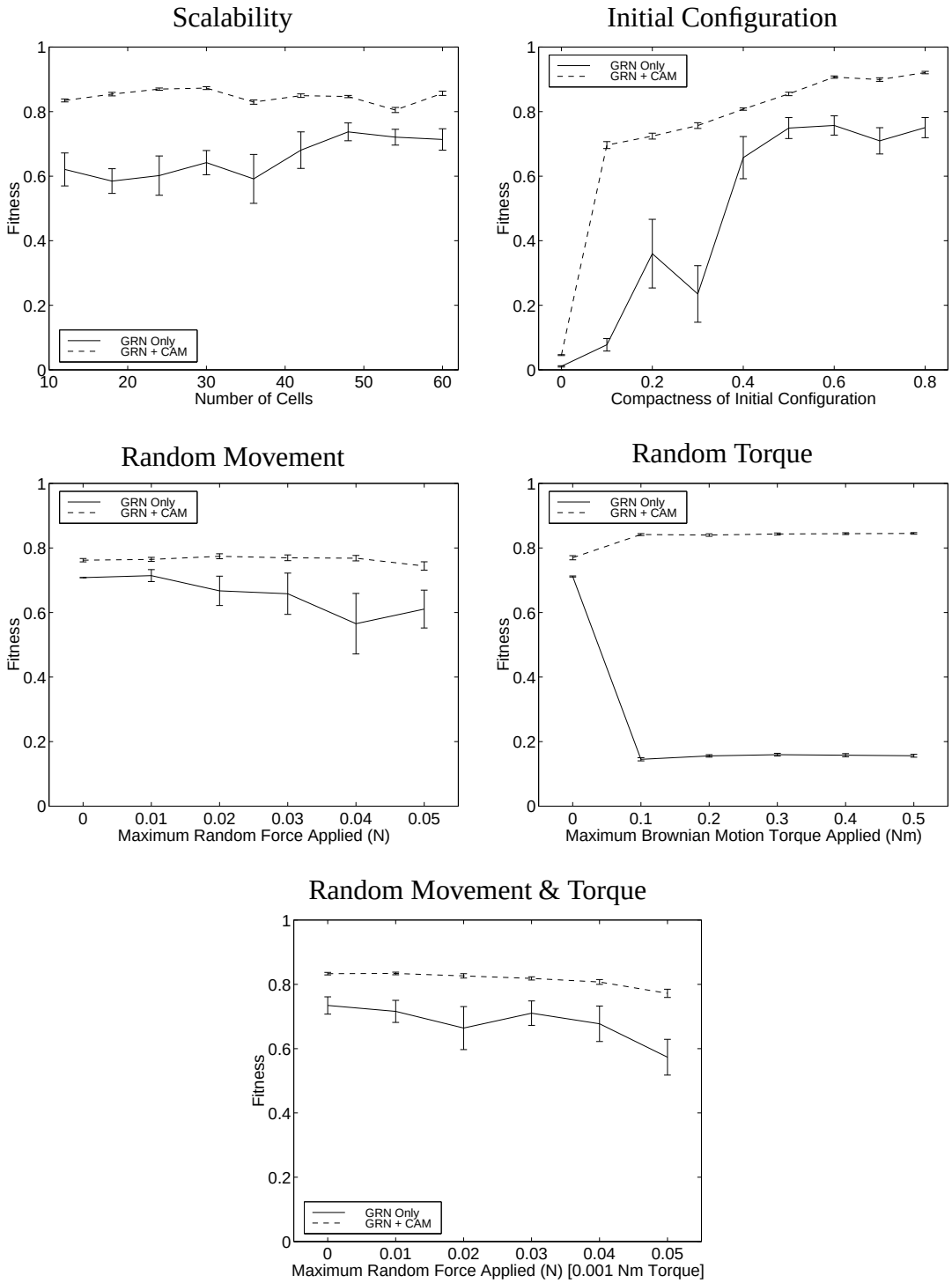


Figure 6.11: Robustness experiments carried out with the clustering controllers.

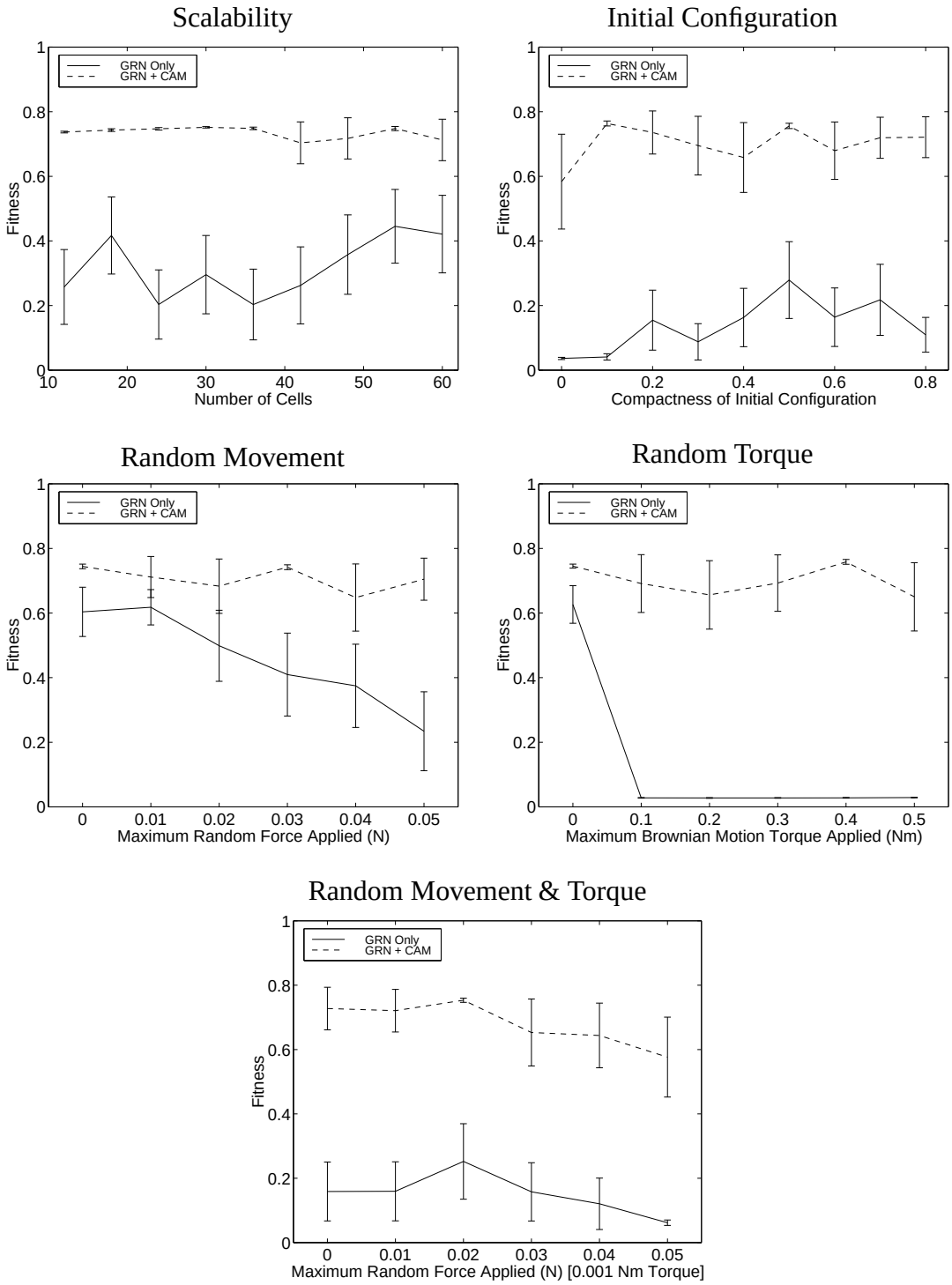


Figure 6.12: Robustness experiments carried out with the signal response controllers.

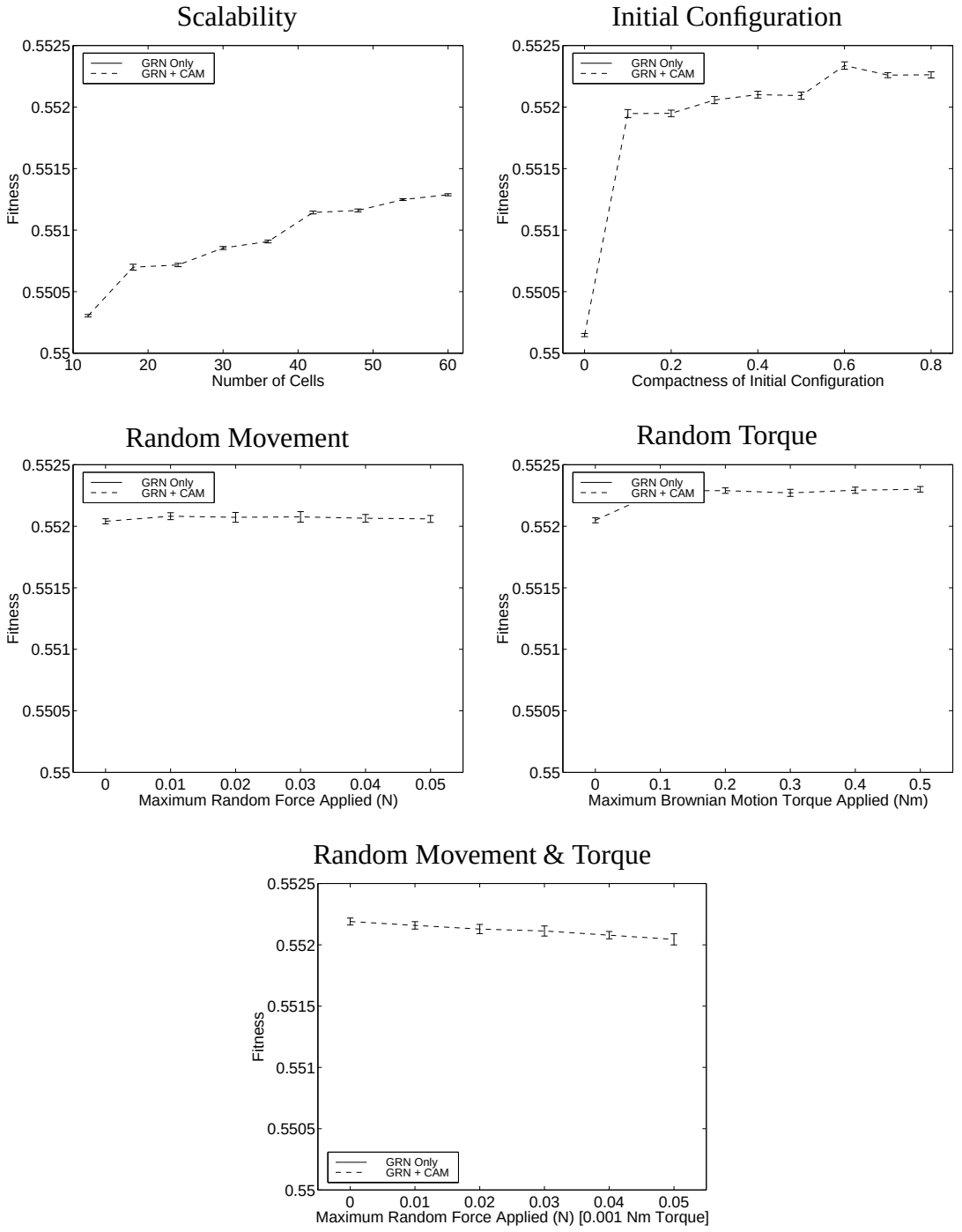


Figure 6.13: Robustness experiments carried out with the differentiation controllers. The results for the GRN-only controller cannot be seen as the mean values were generally below the lower bound of the scale .

Chapter 7

Summary

7.1 Achievements

This thesis has demonstrated the possibility of harnessing the behaviours exhibited by multicellular biological systems to create robust self-organising dynamics in populations of mobile robots. In this section some of the main achievements of this work are presented and discussed.

- **Use of cell adhesion mechanisms to generate interesting global behaviour in a population of mobile robots:**

  The main achievement of this work has been the novel use of abstract cell adhesion mechanisms to robustly control a population of mobile robots. More specifically, this work has demonstrated how this distributed control mechanism can be use to achieve interesting global behaviour where the population move toward some stable configuration.

- **Abstract model of cell adhesion mechanisms:**

  A significant achievement of this work was the development of an abstracted model of the cell adhesion mechanisms which give aggregates of cells their basic structure and layering. The model was developed to accommodate the physical limitations imposed by current technology and was implemented on realistically

simulated robotic modules to demonstrate the potential for physical embodiment and to provide a testbed for experimentation.

- **Distributed self-organisation and self-repair:**

  The proposed model requires only local communication and sensing to generate an extremely robust level of self-organisation in modules with either two or three dimensional movement. In the simplest case, the model can generate effective aggregation behaviour where the modules actively minimise their exposed surface area and form approximate disc or sphere configurations with the mean separation between the modules determined by the system parameters. However, it is also possible to define more complex hierarchical relationships between preselected subsets of the population, causing varying degrees of sorting and mixing in the final configuration. Both these behaviours have demonstrated robustness to environmental perturbations which actively move or rotate the individual modules. Furthermore, it has been shown that this type of turbulence actually helps drive the system toward its final equilibrium state. In addition, as the individual modules are constantly driven toward the same equilibrium state, the model has inherent self-repair allowing it to recover from even large disruptions which may reorganise the whole population.

- **Behaviour consistent with the DAH:**

  The simulation work has shown that the model behaves as predicted by the DAH which inspired its development. This allows the final configuration of any aggregate to be predicted from the adhesive relationships between the subsets of modules in a population. More importantly, the reverse is also true, and given a final configuration, the appropriate adhesions can be calculated.

- **Integration of a powerful dynamic control mechanism:**

  As a final exercise, the basic model's main weakness in its lack of dynamic control was addressed through a collaborative work where the cell adhesion model was combined with a biologically inspired genetic regulatory network (GRN) controller. This introduces a mechanism for dynamically increasing or decreasing the levels of expression of particular A-CAMs (which control the adhesions

between the modules) in response to environmental stimuli. It was demonstrated that the GRN can successfully manipulate the dynamics of the cell adhesion model to achieve a number of more complex tasks. In addition, it is shown that this hybrid model easily outperforms a pure GRN controller in every case.

- **Framework for further development:**

  The final hybrid model provides a basic cellular system which affords the inclusion of additional cellular behaviours. These behaviours would provide additional dynamics for the evolved controllers to exploit, thus allowing increasingly complex patterns of behaviour to be produced.

7.2 Future Work

7.2.1 Examination of Scalability

The distributed nature of the model which has been presented in this work means that each module in the system runs the same algorithm regardless of the total population size. Therefore, in a physical implementation, no additional overhead is generated when extra modules are added. However, in simulation, the interactions between all the modules must be considered. Therefore, larger population sizes have far greater computational requirements.

To allow a thorough exploration of the model parameters the population sizes which could be effectively investigated in this work had to be limited. Therefore, an obvious continuation of this work could examine the true scalability of the model. Larger population sizes would also allow the investigation of more complex hierarchical relationships and should help reveal any limitations of this approach.

7.2.2 Expansion of the Combined controller

7.2.2.1 Morphogen Gradients & Chemotaxis

Morphogen gradients are a key element of morphogenesis and provide a robust mechanism for differentiating cells and breaking the symmetry of developing organisms

(see section 1.4.1.2). In addition, through the phenomenon known as chemotaxis, cells are able to follow even weak gradients toward their source and successfully migrate over long distances. Section 1.4.1.2 introduces the existing work in collective robotics where virtual gradients are used to create simple 2D geometric structures. In the most promising work, presented by Mamei et al. [39], the agents generate the gradients using only local communication by propagating signals through the population.

This elegant approach could easily be integrated into the GRN-controlled cell adhesion model by simply adding an extra category of diffusing proteins which the GRN can detect and produce. These values could then be propagated using the existing A-Cell communication mechanism. As the A-Cells have the additional ability to sense the direction the signals originate from, they could also generate additional attractive forces which drive them up or down the gradient to generate chemotaxis-like behaviour. However, unlike the pure gradient approach proposed by Mamei et al. [39], the underlying adhesive forces should prevent the A-Cells from simply wandering away from the main aggregate and thus increase the overall robustness.

7.2.2.2 Cell Division & Cell Death

Cell division and cell death both play a crucial role during morphogenesis and control the growth and development of every structure. In addition, in developed organisms these functions are critical for both general maintenance and repair. Therefore, it is clear that these functions could also be beneficial in the creation and maintenance of complex self-organising structures.

It is possible to envisage a potential method where the A-Cells presented in this thesis could mimic these complex cellular behaviours by using the inherent space between the modules to transport units to and from the structure. Such an approach would require each structure to maintain a reservoir of unused A-Cells. Therefore, a single cell division would require the recruitment of one of these free A-Cells, while a single cell death (retirement) would assign an A-Cell back to the reservoir. The complex task is to successfully transport these A-Cells to their destination without disrupting the existing structure. It is this task which could exploit the design of the A-Cell model. If an A-Cell fails to broadcast a communication signal it is effectively invisible to the

other A-Cells in the aggregate. However, it can still detect the signals being broadcast and thus avoid any collisions. Therefore, it is theoretically possible for A-Cells to navigate through the empty space of an existing structure, whilst having minimal effect on the A-Cells that they pass. This leaves the problem of successfully navigating to the destination. However, gradients like those proposed in section 7.2.2.1 could provide a solution. In a similar approach to the system employed by Stoy and Nagpal [56], dividing A-Cells could generate a diffusing gradient which would guide a recruit from the reservoir to the desired position. Once there the A-Cell could start broadcasting a communication signal and slowly increase the size of its virtual membrane. In this way, it would appear as though the A-Cell had slowly grown at that location. Similarly, through a reverse procedure, A-Cells could die and follow a gradient back to the reservoir.

Clearly the described process omits many of the details that such behaviours would require. However, it is simply intended to provide a basic illustration of a possible approach that could be employed when considering this problem.

Appendix A

Complementary Algorithms

A.1 Concave Hull Algorithm

This appendix describes the details of the two algorithms that were developed for calculating the concave hull of a connected set of points in two and three dimensions[1]. Considering the points and connections as vertices and edges, the concave hull is defined as the smallest polygon/polyhedron enclosing all the vertices which can be constructed from the available edges.

A.1.1 2D Concave Hull

The convex hull of a set of points is often described as the boundary which would be created if a rubber band was stretched around the points and released. More specifically this is the smallest polygon which encloses all the points. From this definition it is clear that the convex hull is simply a less constrained polygon which must enclose any concave hull which exists.

This concept forms the basis of the algorithm presented here. In summary, starting with a convex hull, edges are simply pushed into the polygon splitting the edge at some previously enclosed point until the desired concave hull is formed.

As mentioned above, the first step of the algorithm involves calculating the convex

[1] For the purposes of this work, where the points represent A-Cells, a connection exists between the two A-Cells if their membranes were in contact.

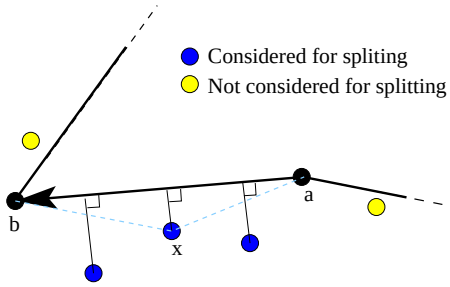


Figure A.1: *The subdivision of edge E_{ab} to form two new edges E_{ax} and E_{xb}.*

hull of the set of points. There are many standard algorithms which can be used for this purpose [24, 26, 33]. However, their individual characteristics vary greatly. In this work the Jarvis's March approach[33] was used as it can be implemented in such a way that it is more tolerant of near-linear datasets[2]. Once the convex hull has been determined, the algorithm proceeds by recursively subdividing the edges of the hull whose vertices are not connected. An edge (E_{ab}) is subdivided by forming two new edges, E_{ax} and E_{xb}, with the vertex (x) which is on the left hand side of E_{ab} and has the shortest perpendicular distance to it (see fig. A.1). If an edge between two unconnected points cannot be subdivided (i.e. there are no vertices which form a perpendicular line with the edge), then it is not possible to form a single concave hull around the points. Once all the edges of the polygon are formed by two connected points, the algorithm ends. Figure A.2 shows the algorithm in action.

A.1.2 3D Concave Hull

In 3D, concave hull construction is unsurprisingly more complex. In this case the problem is similar to a more common image processing problem – explicit surface reconstruction from unorganised points. Typically, this involves the reconstruction of a surface given a collection of points obtained by laser range scanning the original. However, critically, the final surface contains only points from the original data. In the alternative approach, implicit reconstruction, the surface is approximated by a func-

[2]Jarvis's March only requires a simple check to determine which side of a line a point lies on. Other approaches require points to be ordered by the angle they form with some reference point[24]. Therefore, near linear datasets require very high precision in the angle measurement.

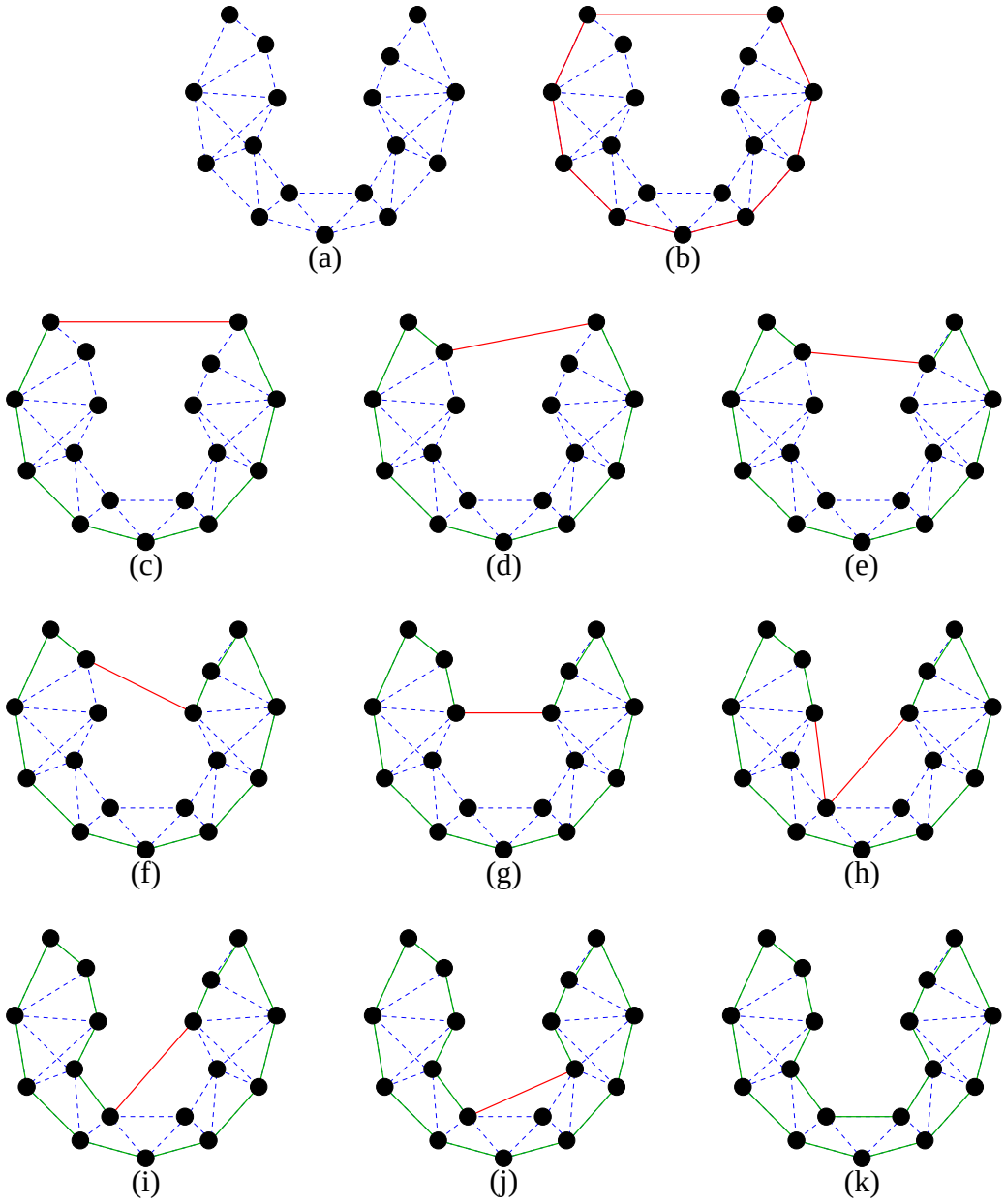


Figure A.2: *The 2D concave hull algorithm applied to a test dataset. (a) the points and the connections between them (b) the convex hull of the dataset (c)–(k) the intermediate stages in the algorithm.*

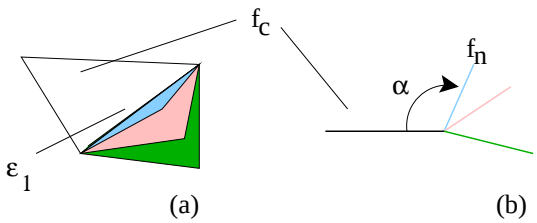


Figure A.3: *Selecting a new face f_n given the current face f_c.*

tion and a polygon mesh is fitted to the approximate surface using some optimisation criteria.

Our concave hull construction problem has two key differences from this traditional problem:

1. The dataset contains internal points in addition to those which lie on the surface.
2. There is additional information in the form of the connections which exist between the vertices.

As it was not possible to find an existing technique which could cope with this type of data the novel algorithm presented in the section was developed. Initially, some attempts were made to modify the existing 2D algorithm. However, it did not adapt well to 3D.

Instead of modifying an existing watertight hull[3], the 3D algorithm starts with only a seed face, which is known to be on the hull, and progressively builds the hull one face at at time. The main assumption this progressive construction is based on is:

*G*iven an edge (e_1) of a face (f_c) known to be on the concave hull and a set of faces ($\mathcal{F}$) which are connected to e_1, the face $f_n \in \mathcal{F}$ which makes the smallest clockwise angle α with f_c is also on the concave hull (see fig. A.3).

The clear problem with this assumption is that it requires knowledge of all the valid faces which can be constructed from a given edge. Therefore, for the progressive stage to operate efficiently, the algorithm must begin by first calculating all the valid faces which can be constructed. This is a very expensive step involving three nested loops (see fig. A.4) which, despite modifications to improve efficiency, will still be On^3,

[3] A fully connected hull with no gaps or holes.

greatly limiting the scalability of the algorithm. However, for datasets of the size used in this work, this is still manageable.

```
for(i=0; i < n-2; ++i)

     for(j=i+1; j < n-1; ++j)

          if( connected(point[i], point[j]) ) then
             for(k=j+1; j < n; ++j)

              if( connected(point[i], point[k]) and connected(point[j],
                 point[k]) ) then
                 face(point[i], point[j], point[k]) is valid.
```

Figure A.4: *Pseudo-code showing the nested loops required for generating the list of valid faces, where n is the number of points in the dataset.*

Having calculated the valid faces, the algorithm proceeds by finding a seed face known to be on the concave hull. This can be determined by finding the face furthest along some random direction vector. The algorithm then proceeds by placing the three edges which form the seed face on to a stack ($\mathcal{E}$) along with details of the face f_{e_i} they were associated with. The seed is then added to a list of final faces ($\mathcal{H}$) which form the concave hull. At this point, the main loop of the algorithm begins by popping the first edge (e_t) off the top of $\mathcal{E}$. The valid faces connected to this edge are then retrieved and the one which makes the smallest clockwise angle with f_{e_t} (i.e. the face associated with the edge), is selected. If this new face (f_n) is already contained in $\mathcal{H}$ (i.e. it has been processed previously) nothing more is done and the loop returns. However, if f_n is not contained in $\mathcal{H}$, it is added and the remaining edges of the face, not e_t, are pushed onto $\mathcal{E}$. In this way the hull slowly grows over the dataset until $\mathcal{E}$ is empty then $\mathcal{H}$ can simply be returned.

Unfortunately, although the above algorithm has an appealing simplicity there are certain cases where it breaks down. These cases relate to the instances where growing sections of the hull intermesh. Figure A.5 shows a clear example of the problem. At this point, some of the faces belonging to the hull have already been determined by the algorithm, including f_e. Therefore, the face f_1, which makes the smallest clockwise

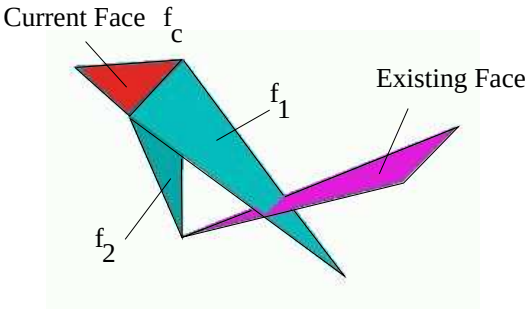


Figure A.5: *The intermeshing faces problem.*

angle with f_c, cannot also be selected as a face on the hull, as it is known to intersect f_e[4]. The solution to this problem is to select f_2, the face which makes the second smallest clockwise angle with f_c instead. To summarise; instead of selecting the face which simply makes the smallest clockwise angle, it is necessary to select the face which makes the smallest clockwise angle and additionally does not intersect any of the faces already determined to be on the hull ($\mathcal{H}$).

Even using efficient triangle intersection code[28], this is another expensive operation, with the number of possible checks growing as faces are added to the hull. However, again this overhead is sufficiently small for the datasets used in this work.

Figure A.6 shows various stages of the algorithm's progress for an example dataset taken from a simulation of the cell adhesion model. In this example, all the vertices are well connected. However, in some cases it could be that two sections of the dataset are connected by only a single vertex (see fig. A.7). It is also equally possible that an isolated edge (i.e. not part of a valid face) might join two sections of the dataset or alternatively link a single vertex with the remainder of the dataset (see fig. A.8). These features would cause problems for the concave hull algorithm as it only moves between faces which share an edge. Therefore, in the situations shown in figures A.7 and A.8, the concave hull would only be built over one section of the dataset. To avoid this complication, the datasets can very easily be segmented into clusters of points which are connected to each other by valid faces. Once the algorithm has been applied to each of these clusters, the individual hulls can then be combined (see fig. A.9).

[4]Clearly, faces on a valid watertight hull cannot intersect.

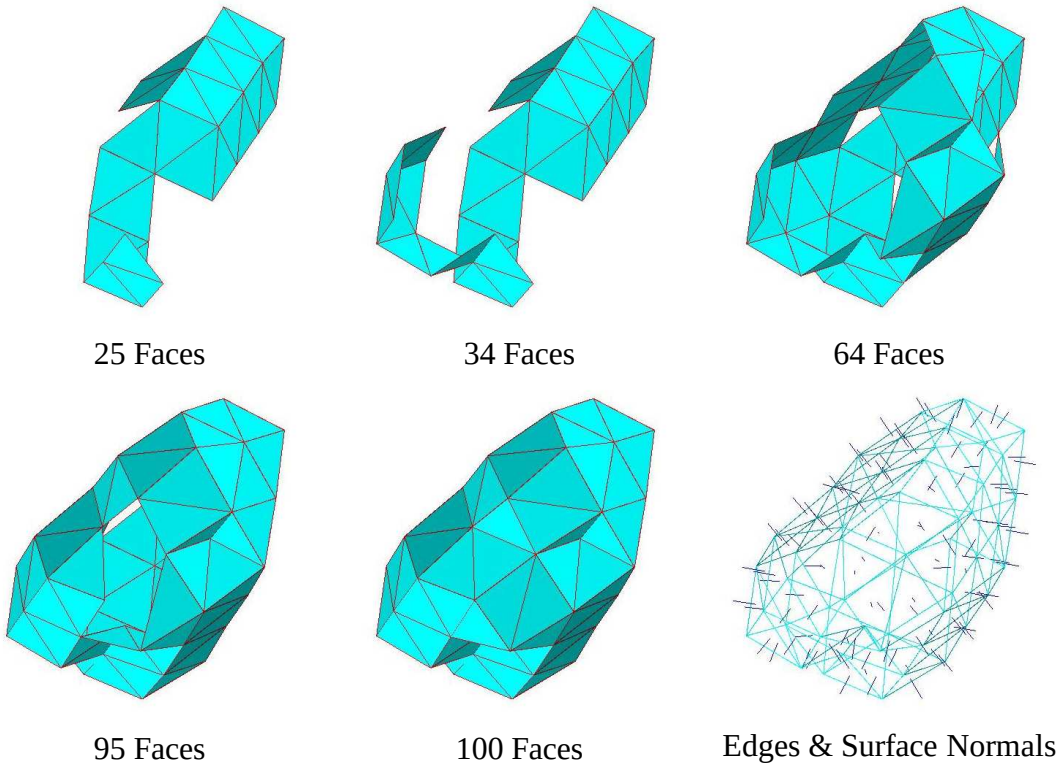


Figure A.6: *The 3D concave hull algorithm applied to a test dataset.*

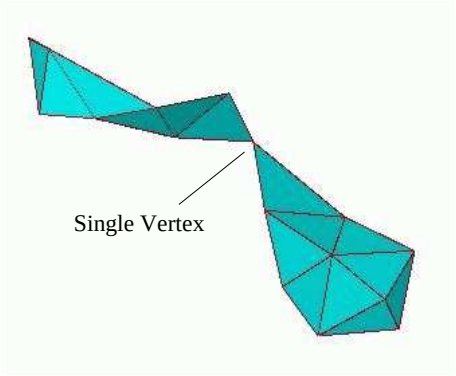


Figure A.7: *A single vertex joining two sections of a dataset.*

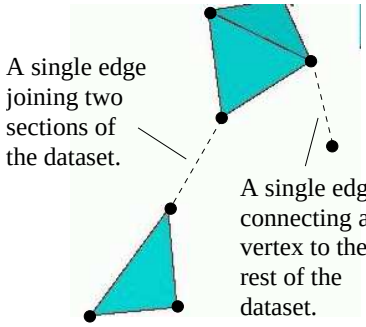


Figure A.8: *Isolated edges in a dataset.*

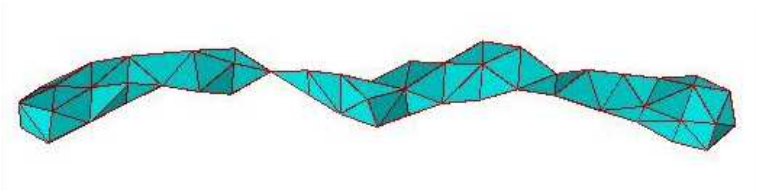


Figure A.9: *The combination of the two hulls calculated for the separate segments of the dataset.*

The design of the concave hull algorithm means that it can be applied to any connected point cloud. Therefore, it can also be used to generate hulls for datasets containing only points from an objects surface (including range data) (see fig. A.10). However, this involves assigning some arbitrary connection distance to the points which is likely to result in a much worse approximation of the actual surface than could be achieved with current image processing techniques [13].

A.2 Generating Random Configurations

This section proposes a method for generating random initial configurations for A-Cell aggregates which require a specific separation distance between the A-Cells. It is assumed that each A-Cell must be the required distance from at least one of its neighbours, to ensure the aggregate is well connected. However, to prevent repulsive forces between the A-Cells simply causing the aggregate to immediately disintegrate, no two A-Cells are placed closer together than this distance.

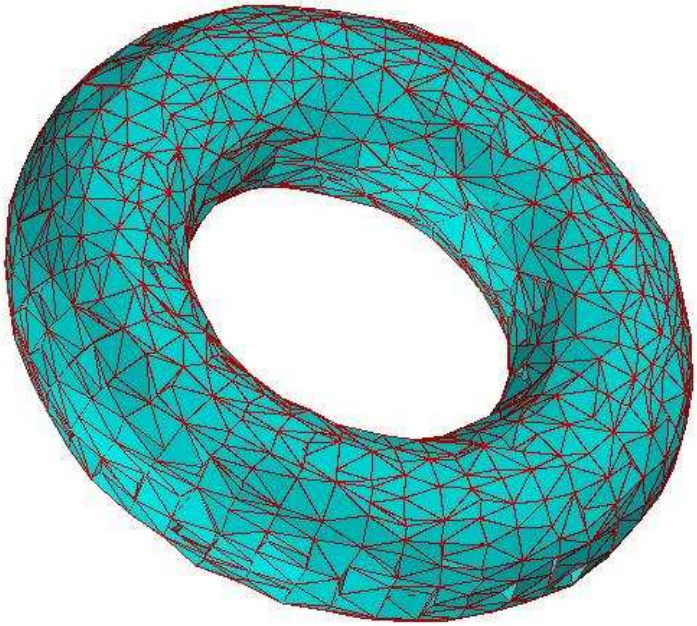


Figure A.10: *The concave hull calculated from surface points of a noisy torus.*

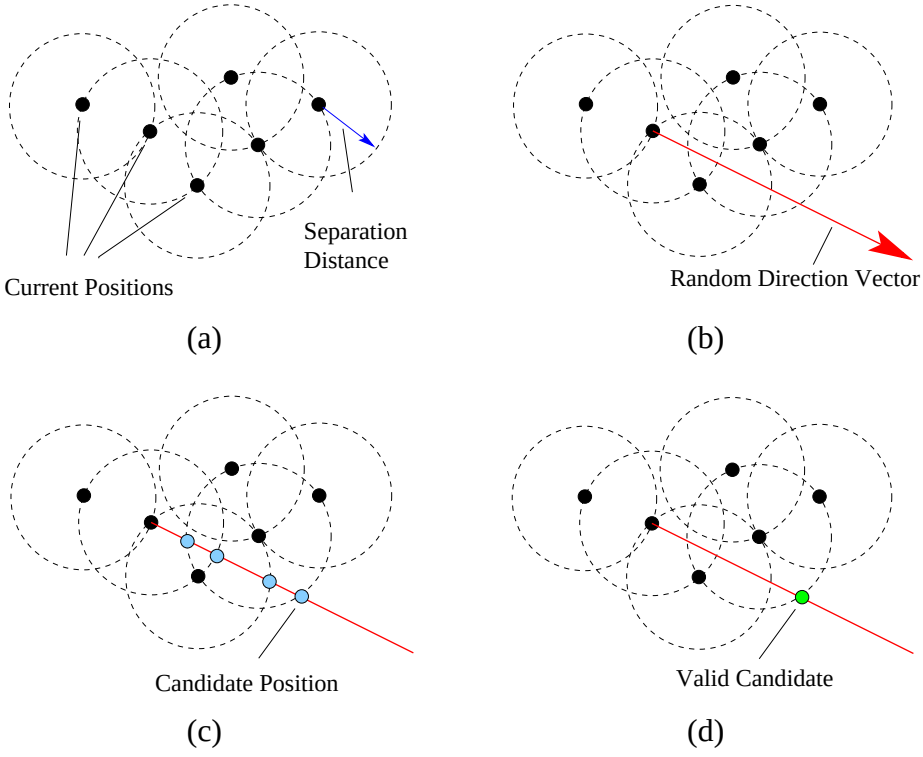


Figure A.11: *Selecting a new position for a random configuration.*

The algorithm starts with the origin as an initial seed position. New positions are then added one at a time until the desired aggregate size is achieved. To create a new position, a random direction vector (d) is generated. This can be either 2D or 3D, depending on the desired dimensionality of the random configuration (see fig. A.11 - b). A line is then created from the origin in direction d to some point well outside the bounding box of the current set of positions. A list of new candidate positions can then be generated by calculating the intersection between this line and a sphere of radius m, where m is the separation distance, around each of the current positions (see fig. A.11 - c). These are then processed to remove those which are within distance m of any of the other current positions (see fig. A.11 - d) and one of the remaining candidates is selected.

Figure A.12 shows an example random configuration generated for a 64 A-Cell aggregate using the algorithm outlined above.

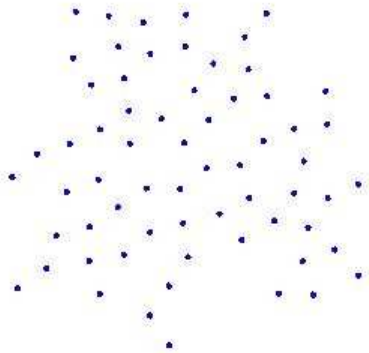


Figure A.12: *An example 2D random configuration generated for a 64 A-Cell aggregate.*

Appendix B

Additional Results from the 2D Sorting Experiments

This appendix contains full results from some of the experiments summarised in chapter 4.

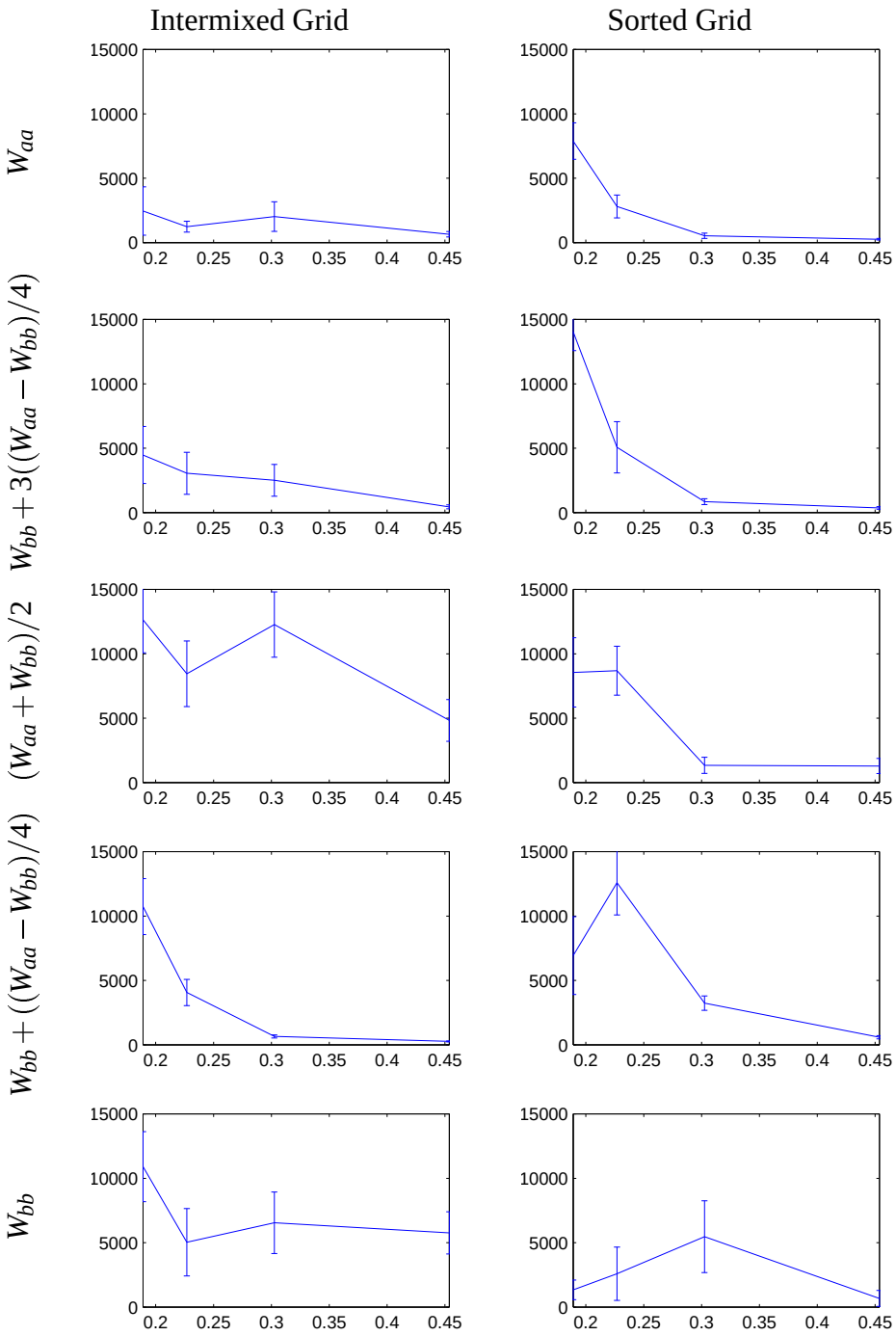


Figure B.1: CT *results for 36 A-Cell aggregates starting from both an initially mixed grid and a sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

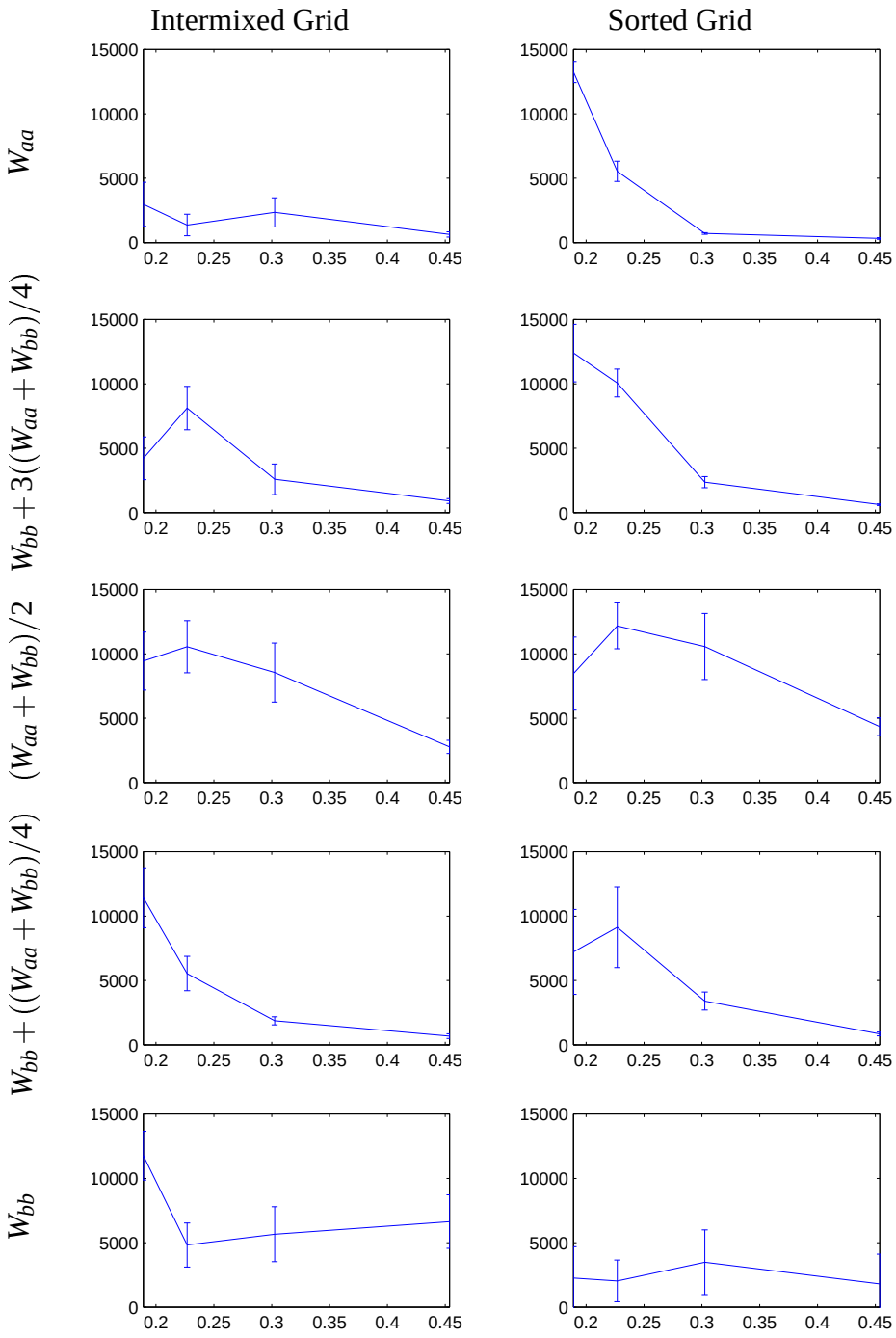
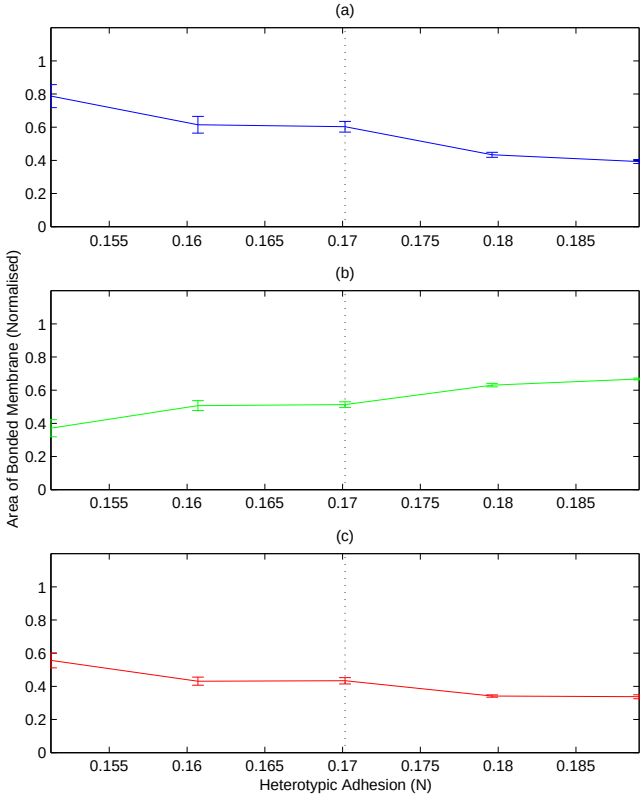


Figure B.2: CT *results for 64 A-Cell aggregates starting from both an initially mixed grid and a sorted grid. In each graph 4 homotypic adhesion pairs are shown. The* $\mathbf{b} \rightarrow \mathbf{b}$ *adhesion of each pair is fixed at* $0.15125N$ *while the* $\mathbf{a} \rightarrow \mathbf{a}$ *adhesion increases along the x axis. Therefore, the magnitude of difference between the homotypic adhesions increases from left to right.*

CBI Results: (a) **a** → **a** (b) **a** → **b** (c) **b** → **b**



CT Results.

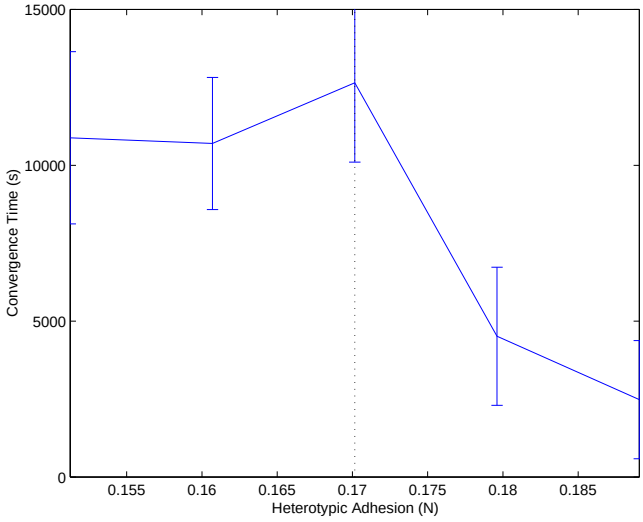


Figure B.3: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid. **a** → **a** adhesion = $0.1890625N$ and **b** → **b** adhesion = $0.15125N$

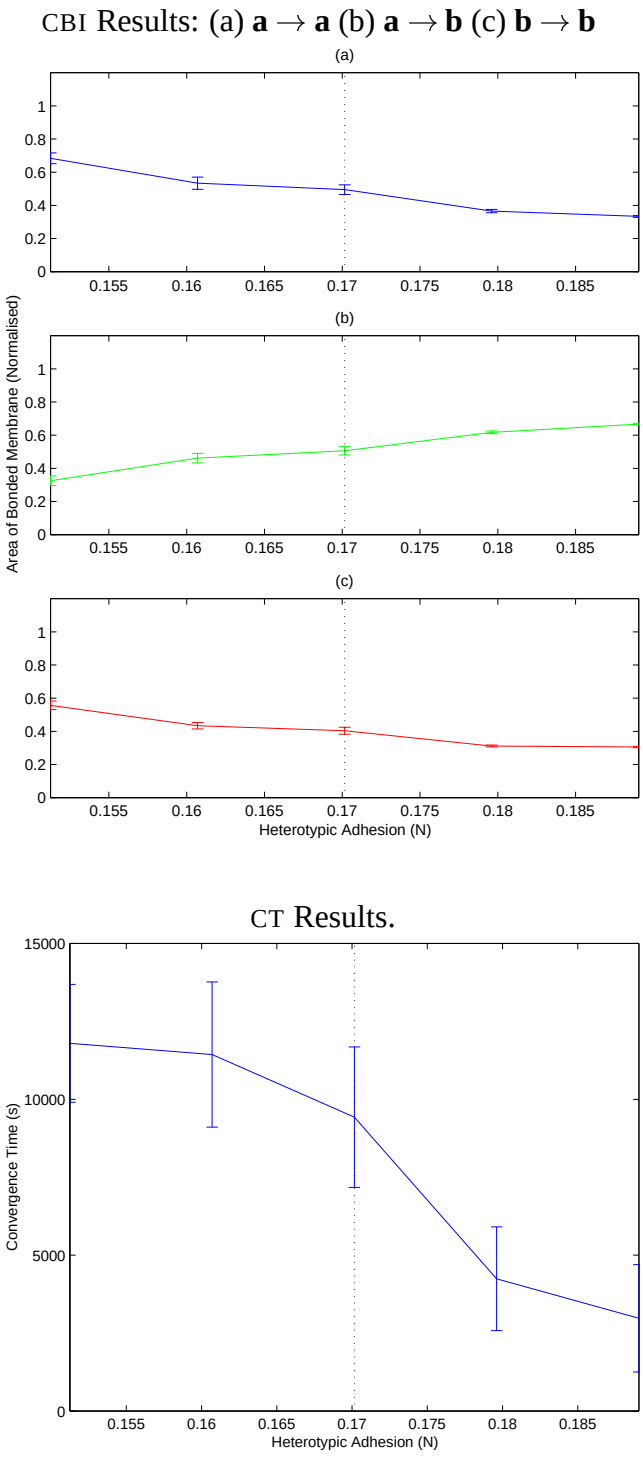
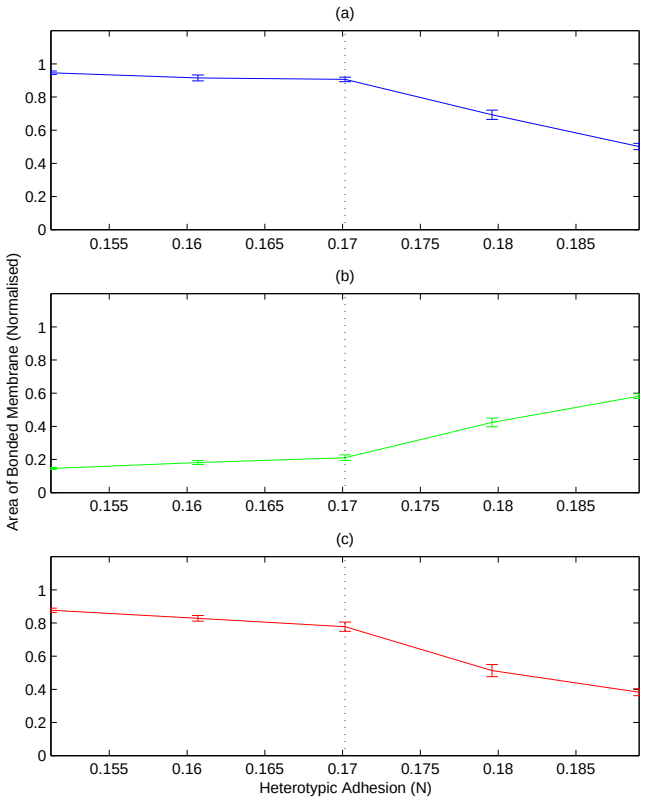


Figure B.4: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed grid. **a** $\rightarrow$ **a** adhesion = $0.1890625N$ and **b** $\rightarrow$ **b** adhesion = $0.15125N$

CBI Results: (a) **a** → **a** (b) **a** → **b** (c) **b** → **b**



CT Results.

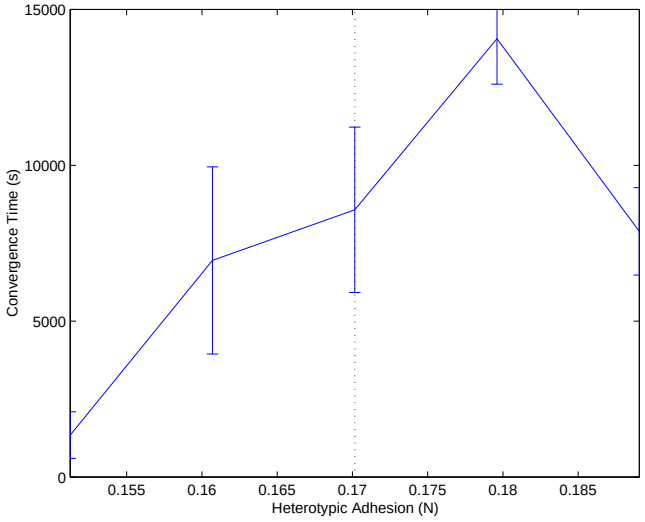


Figure B.5: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid. **a** → **a** adhesion = $0.1890625N$ and **b** → **b** adhesion = $0.15125N$

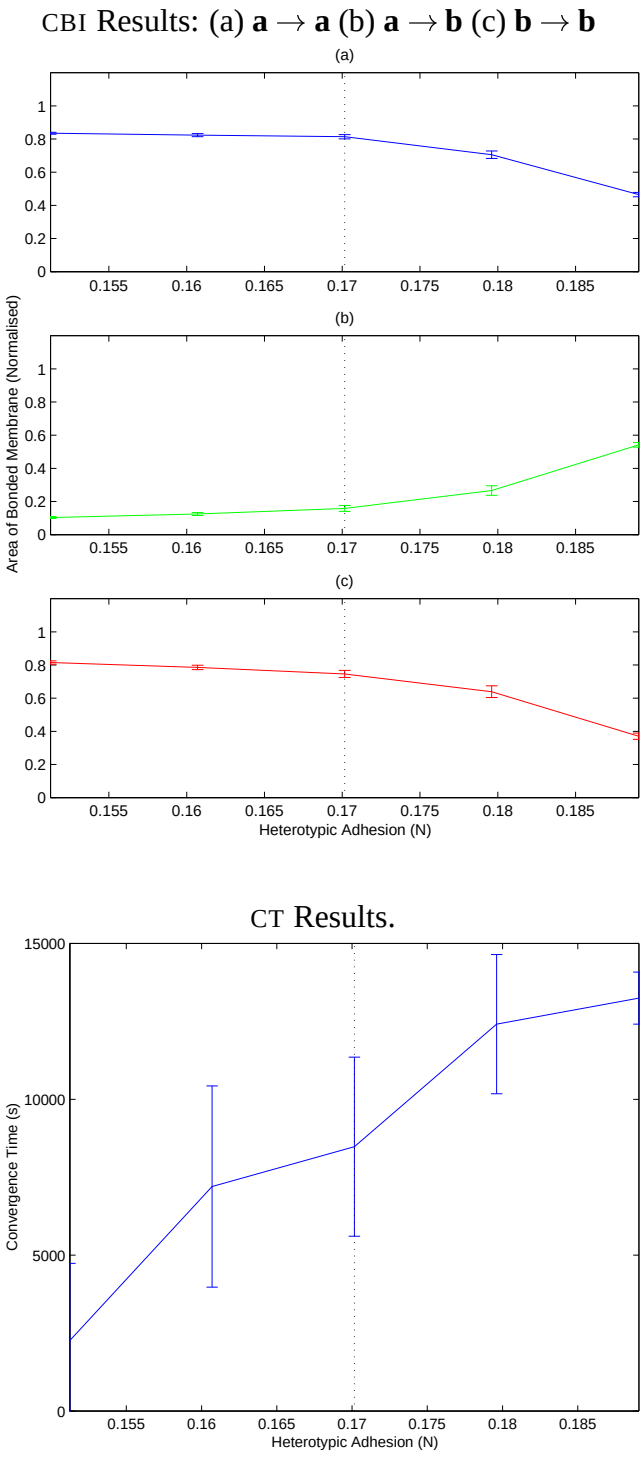


Figure B.6: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted grid. **a** → **a** adhesion = $0.1890625N$ and **b** → **b** adhesion = $0.15125N$

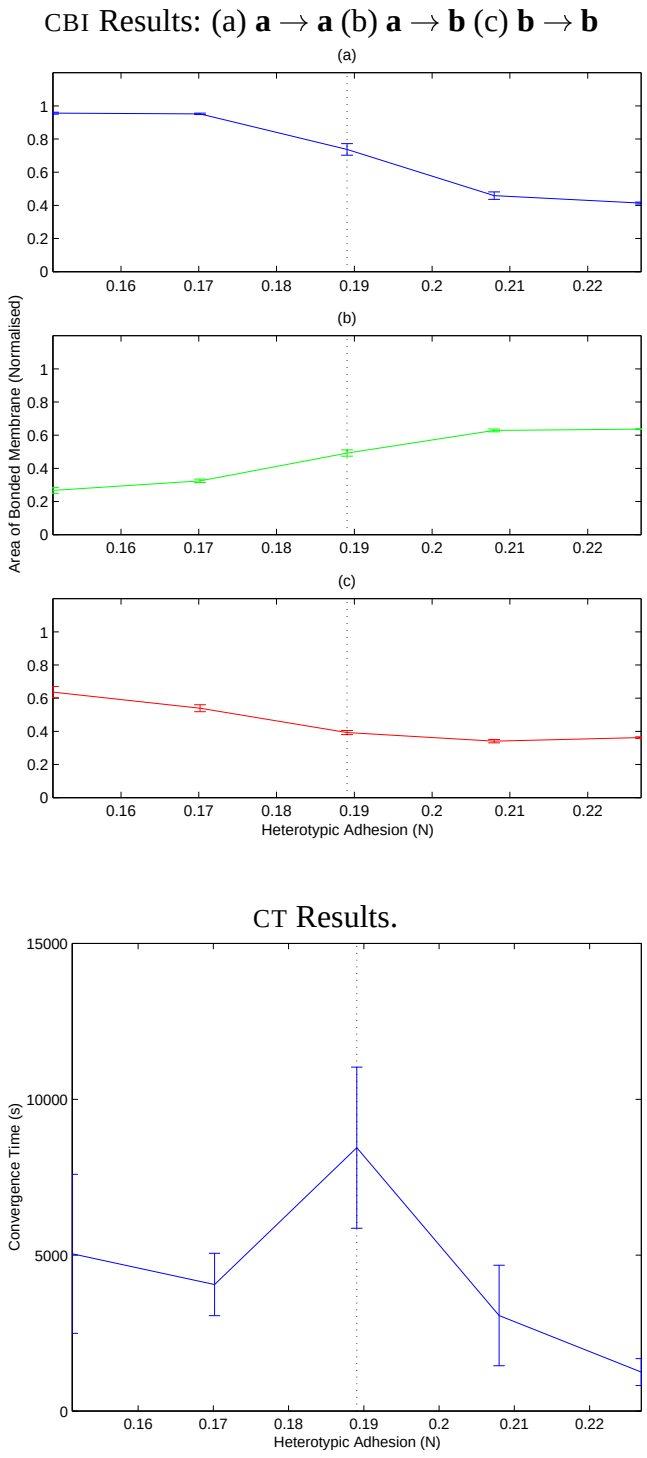


Figure B.7: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid. **a** → **a** adhesion = $0.226875N$ and **b** → **b** adhesion = $0.15125N$

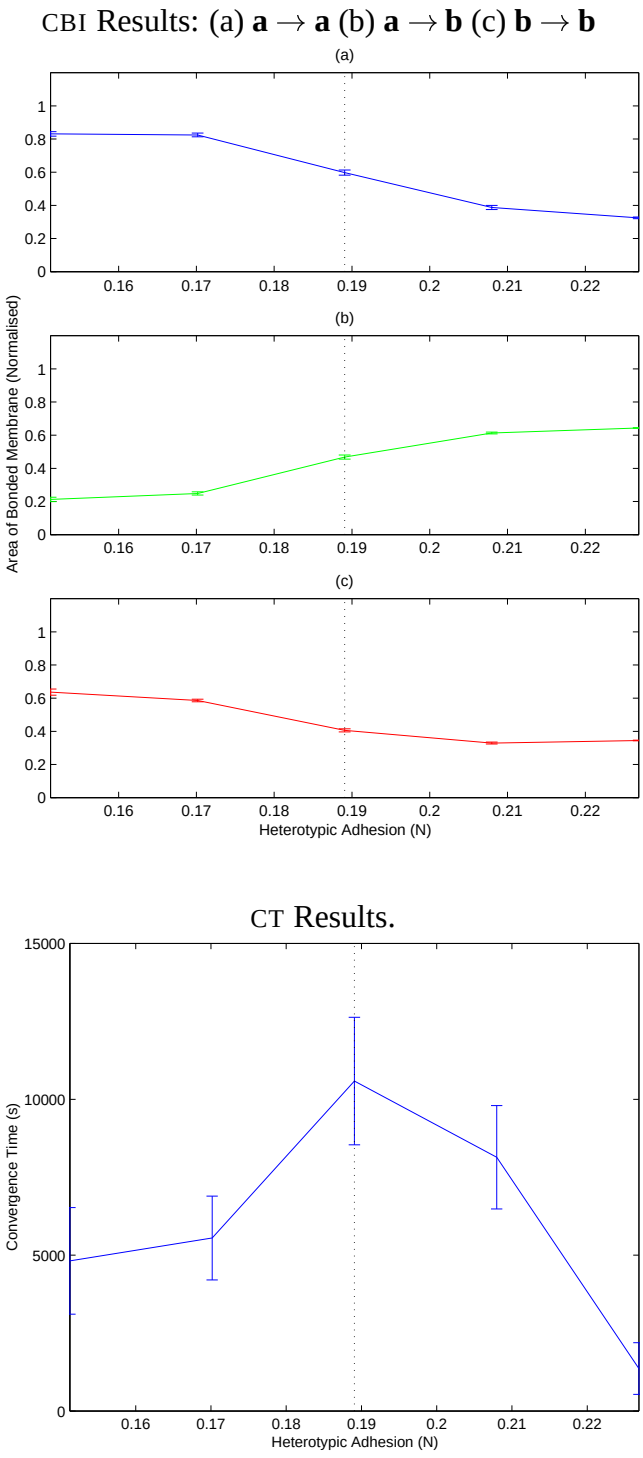
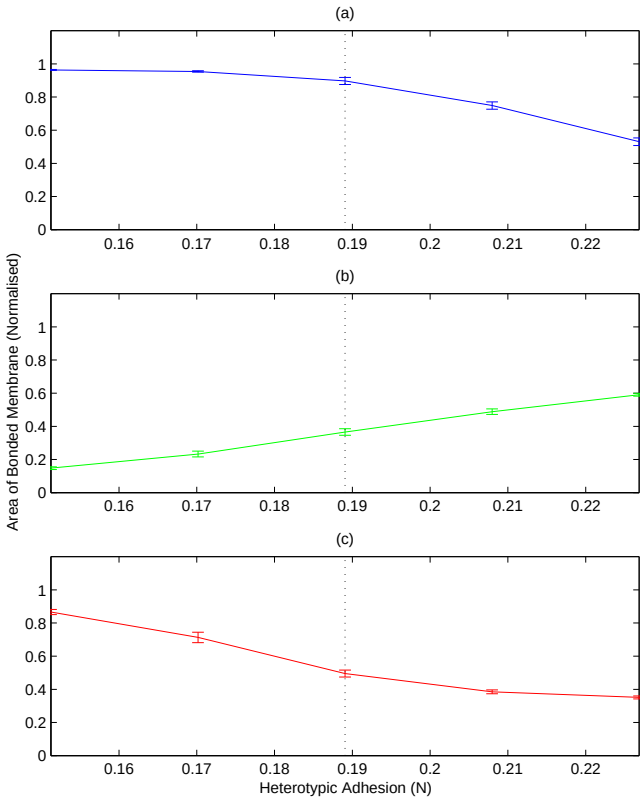


Figure B.8: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed grid. **a** $\rightarrow$ **a** adhesion = $0.226875N$ and **b** $\rightarrow$ **b** adhesion = $0.15125N$

CBI Results: (a) **a** → **a** (b) **a** → **b** (c) **b** → **b**



CT Results.

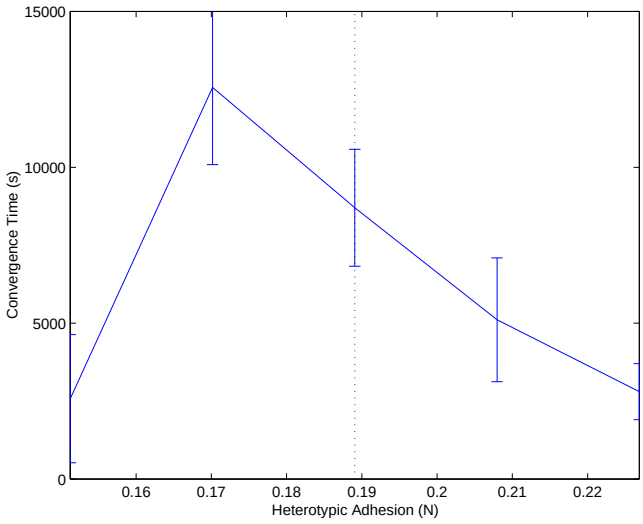


Figure B.9: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid. **a** → **a** adhesion = $0.226875N$ and **b** → **b** adhesion = $0.15125N$

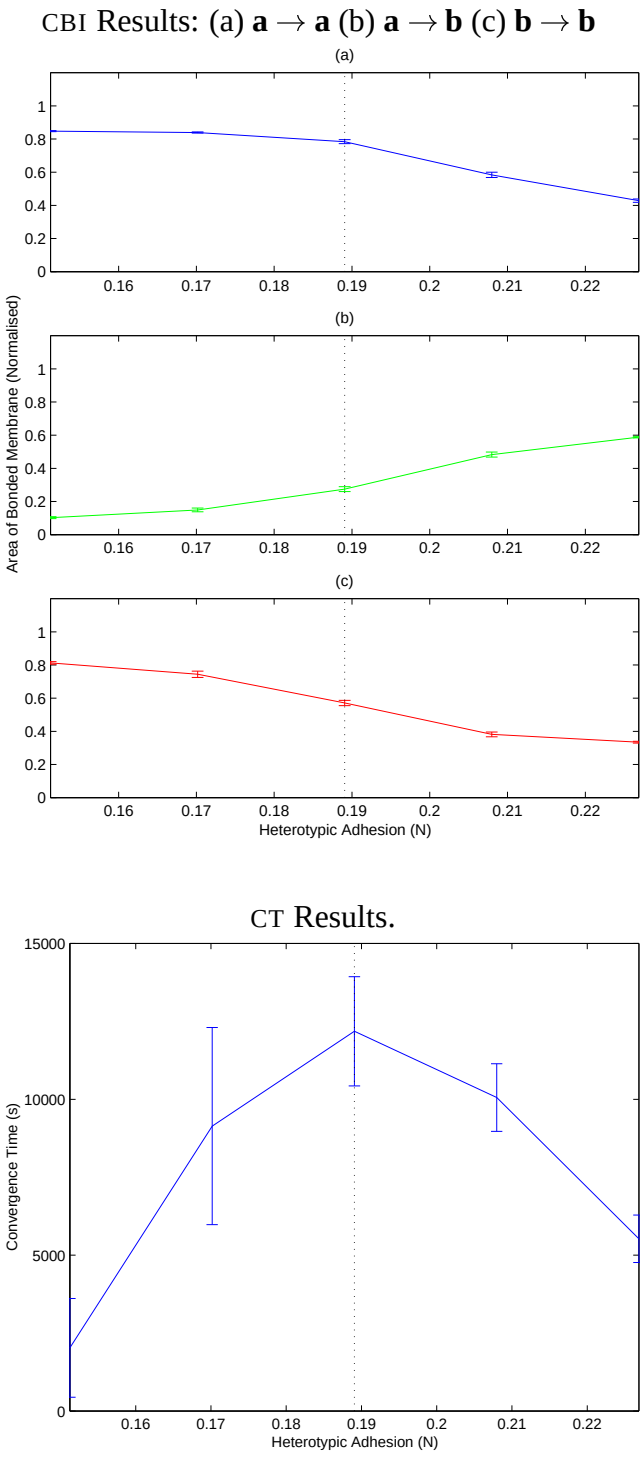
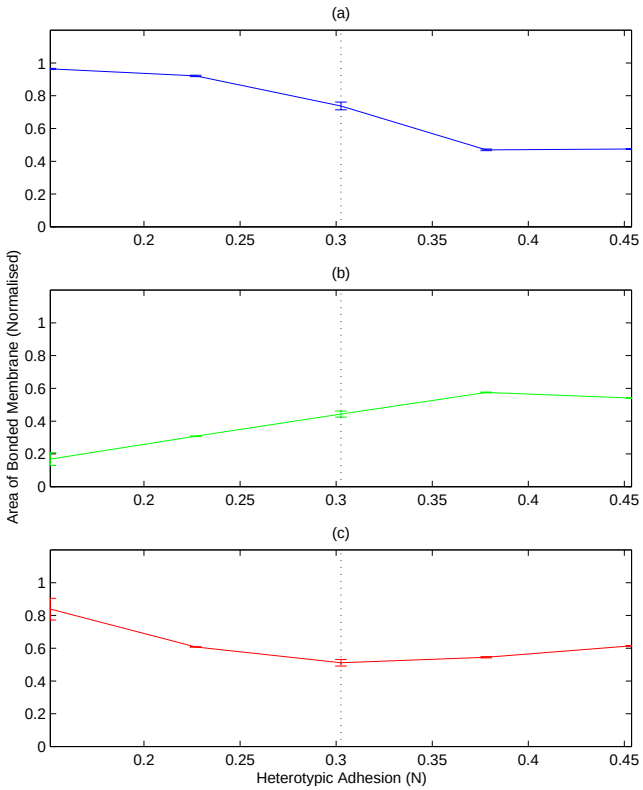


Figure B.10: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted grid. **a** → **a** adhesion = $0.226875N$ and **b** → **b** adhesion = $0.15125N$

CBI Results: (a) **a** → **a** (b) **a** → **b** (c) **b** → **b**



CT Results.

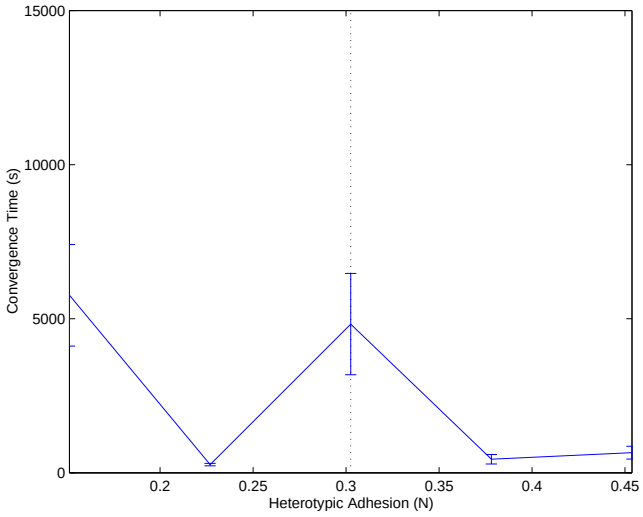


Figure B.11: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of an intermixed grid. **a** → **a** adhesion = $0.45375N$ and **b** → **b** adhesion = $0.15125N$

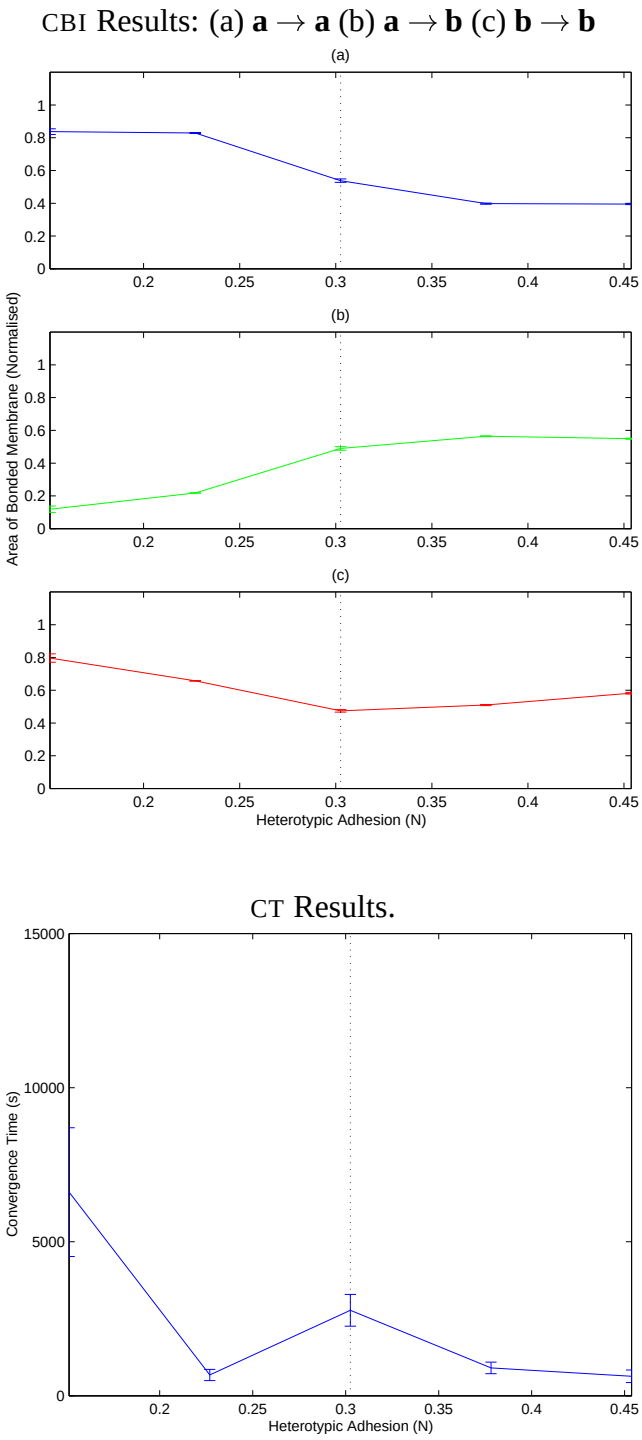
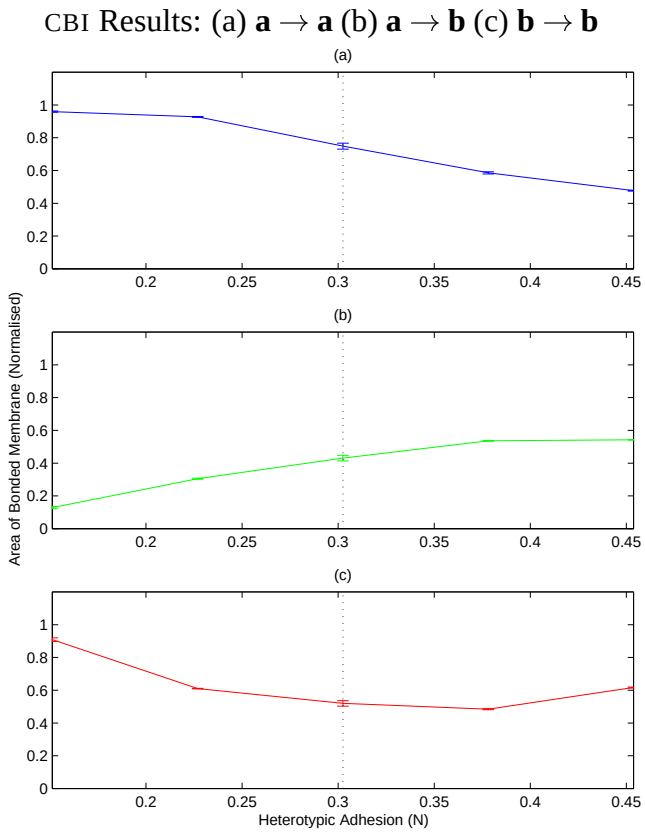


Figure B.12: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of an intermixed grid. **a** → **a** adhesion = $0.45375N$ and **b** → **b** adhesion = $0.15125N$



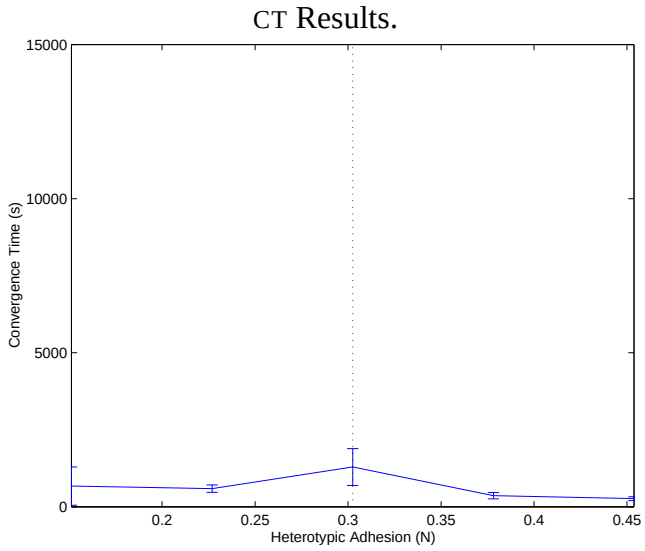


Figure B.13: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 36 A-Cell aggregates with the initial configuration of a sorted grid. **a** → **a** adhesion = $0.45375N$ and **b** → **b** adhesion = $0.15125N$

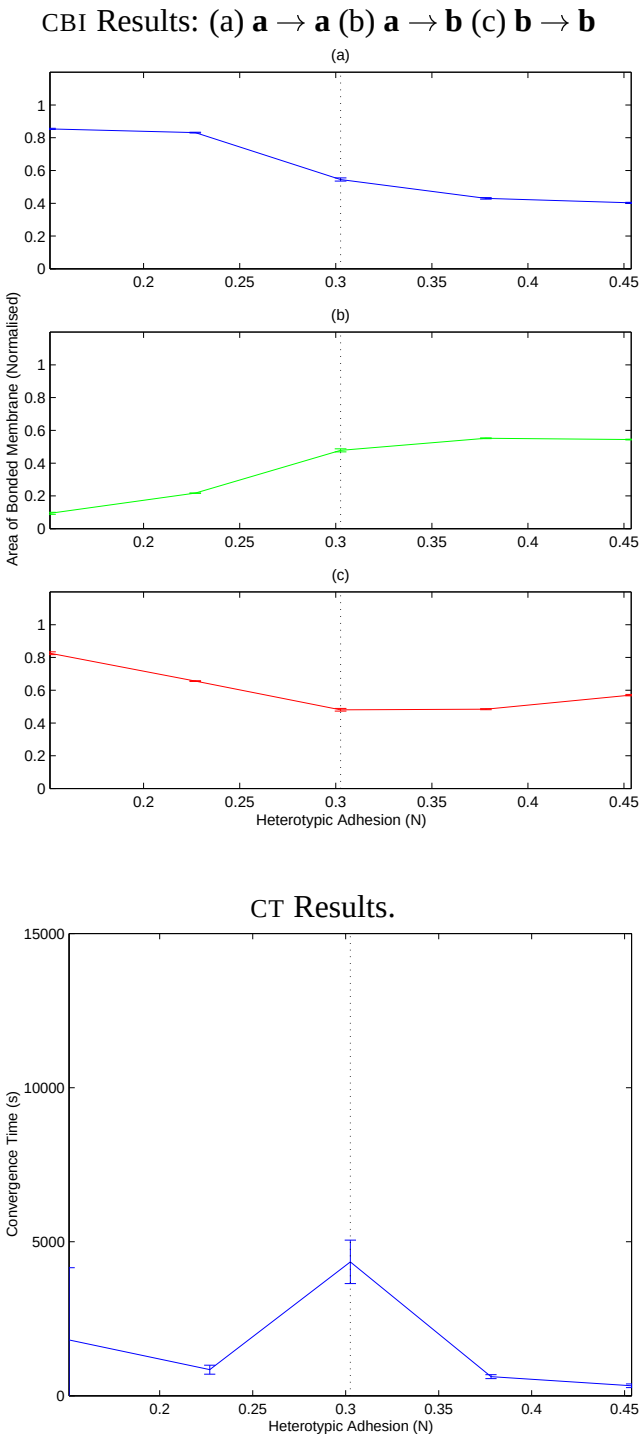


Figure B.14: Results for a range of heterotypic attraction at equilibrium values. The results were obtained from 64 A-Cell aggregates with the initial configuration of a sorted grid. **a** → **a** adhesion = $0.45375N$ and **b** → **b** adhesion = $0.15125N$

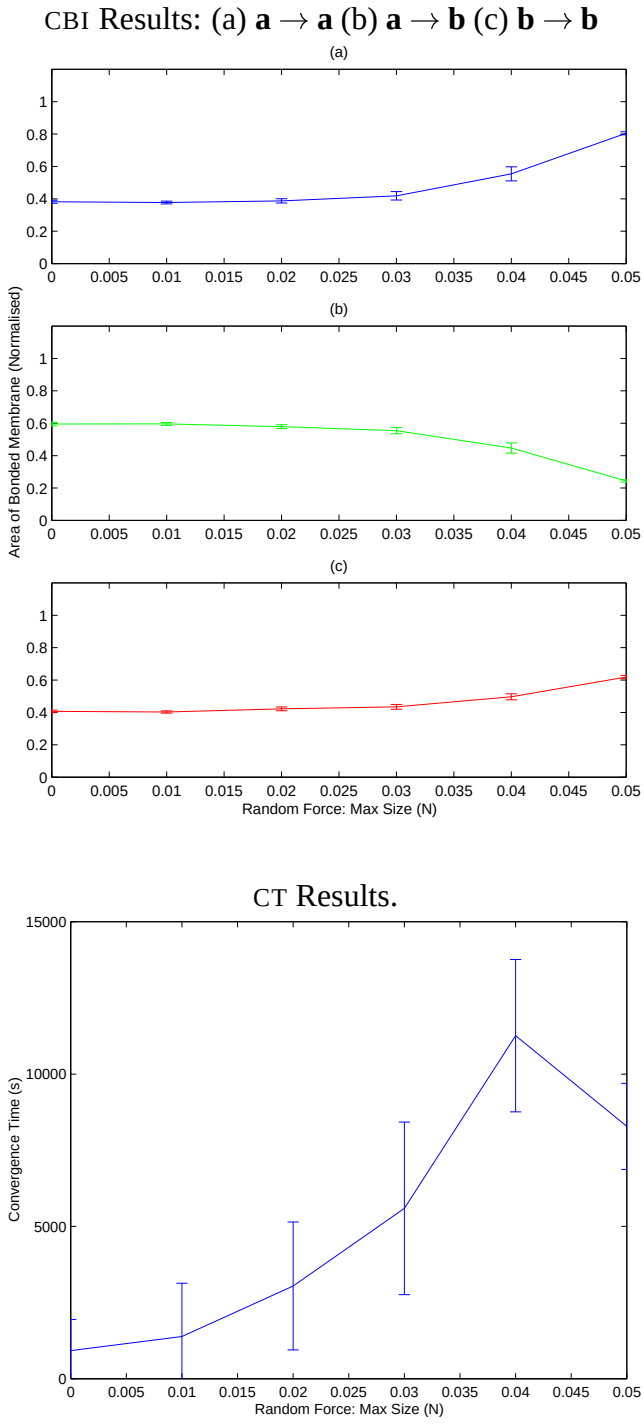


Figure B.15: *Results for 64 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the onion equilibrium state.*

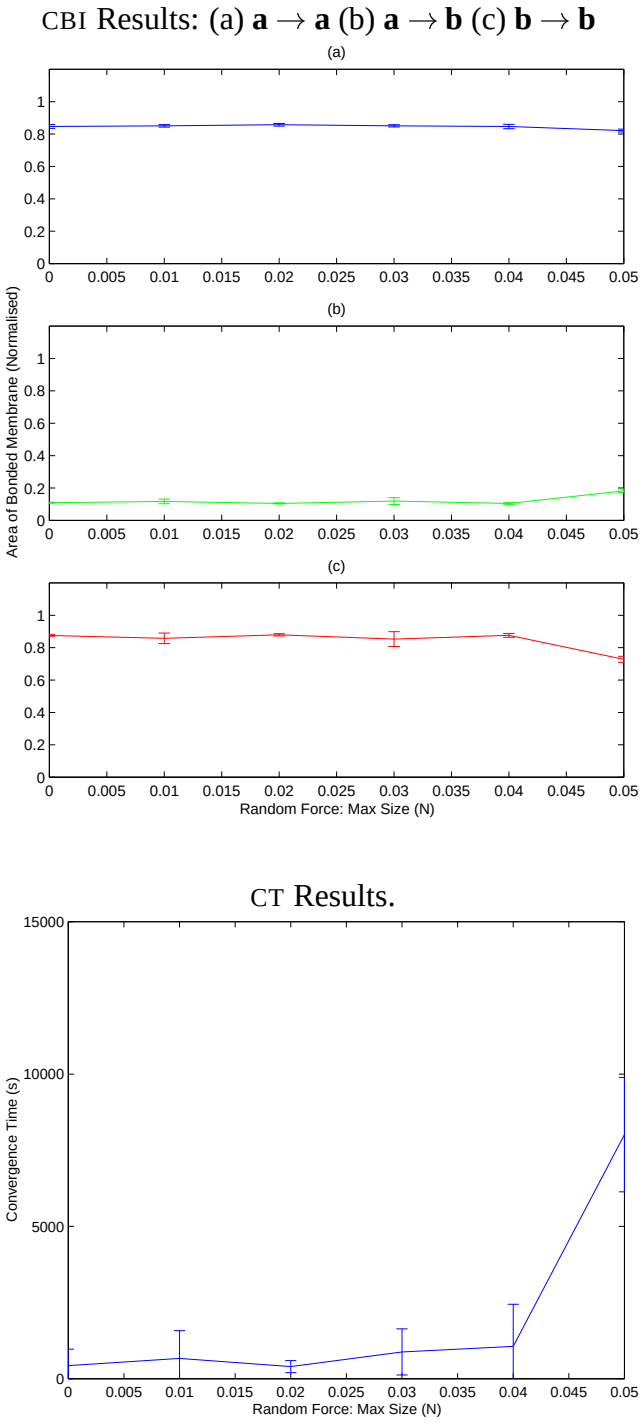


Figure B.16: *Results for 64 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the onion equilibrium state.*

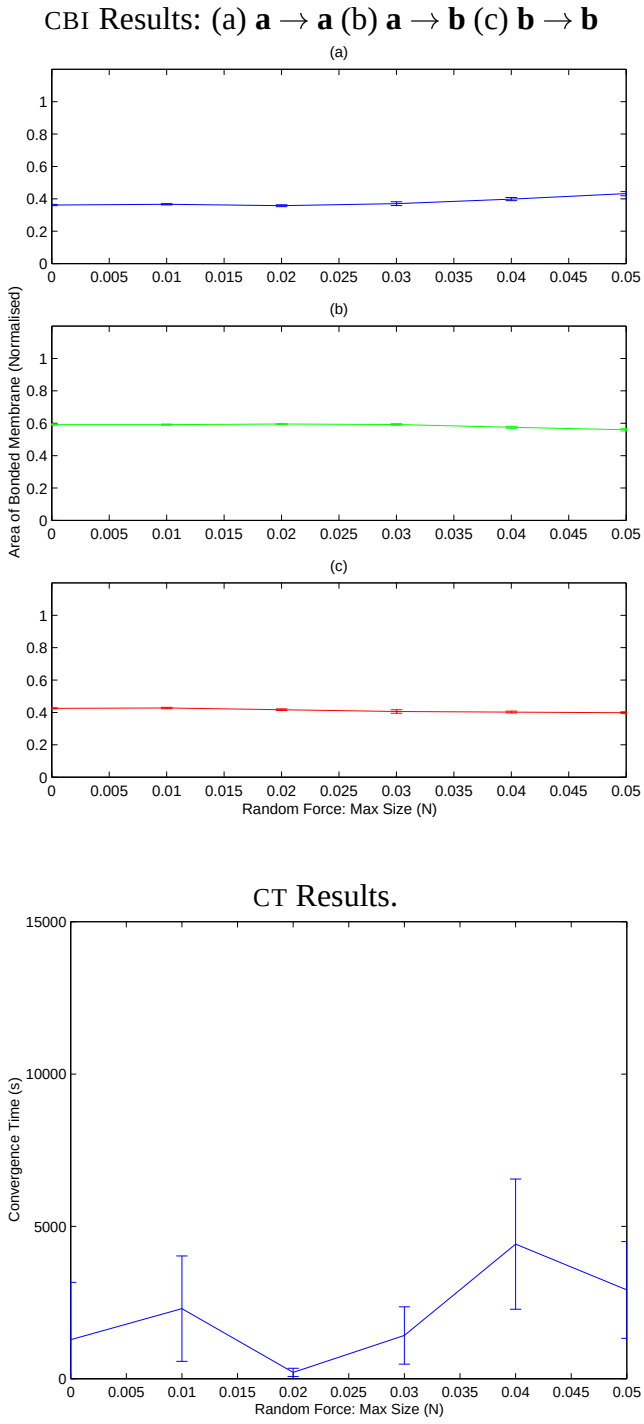


Figure B.17: *Results for 64 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the intermixed equilibrium state.*

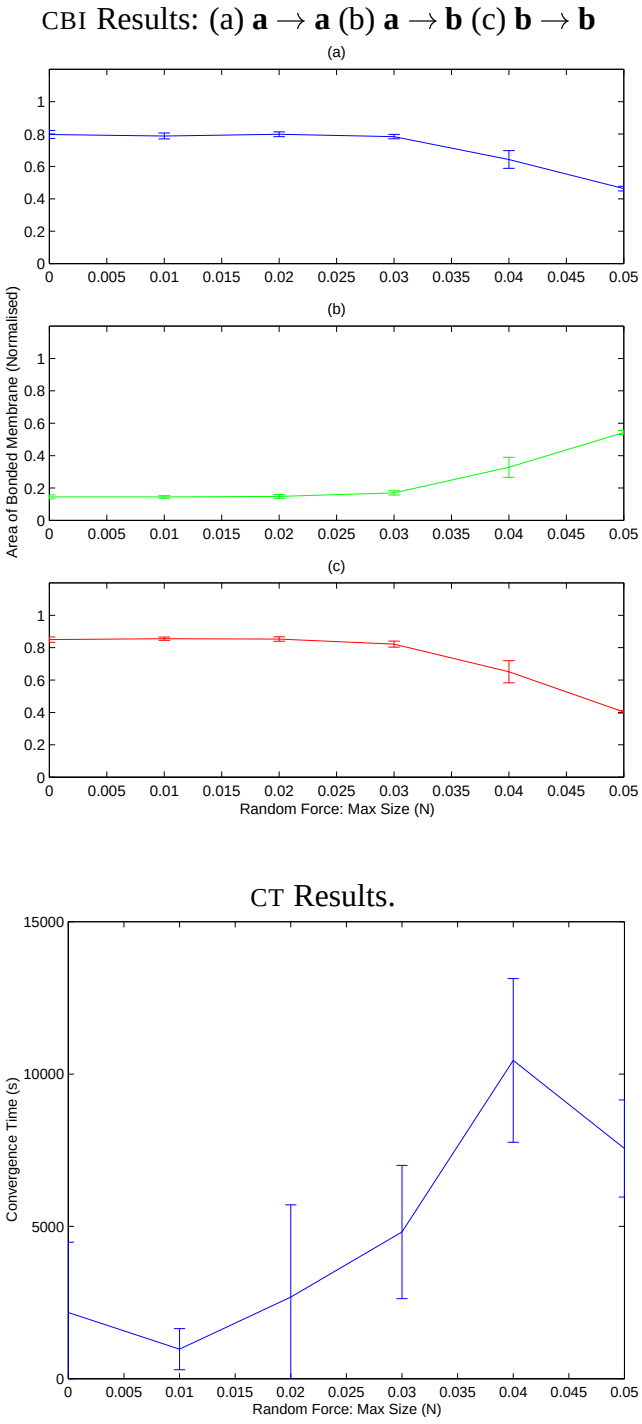


Figure B.18: *Results for 64 A-Cell aggregates over a range of random movement force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the intermixed equilibrium state.*

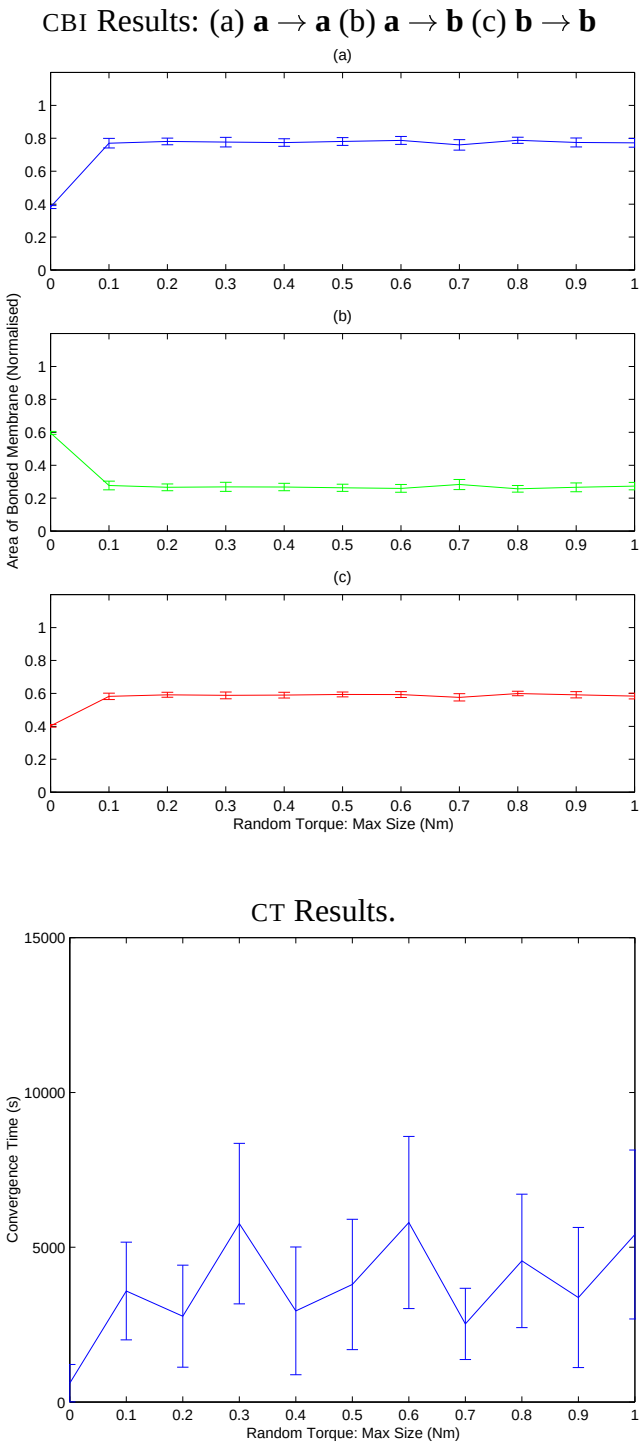


Figure B.19: *Results for 64 A-Cell aggregates over a range of random movement torque limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the onion equilibrium state.*

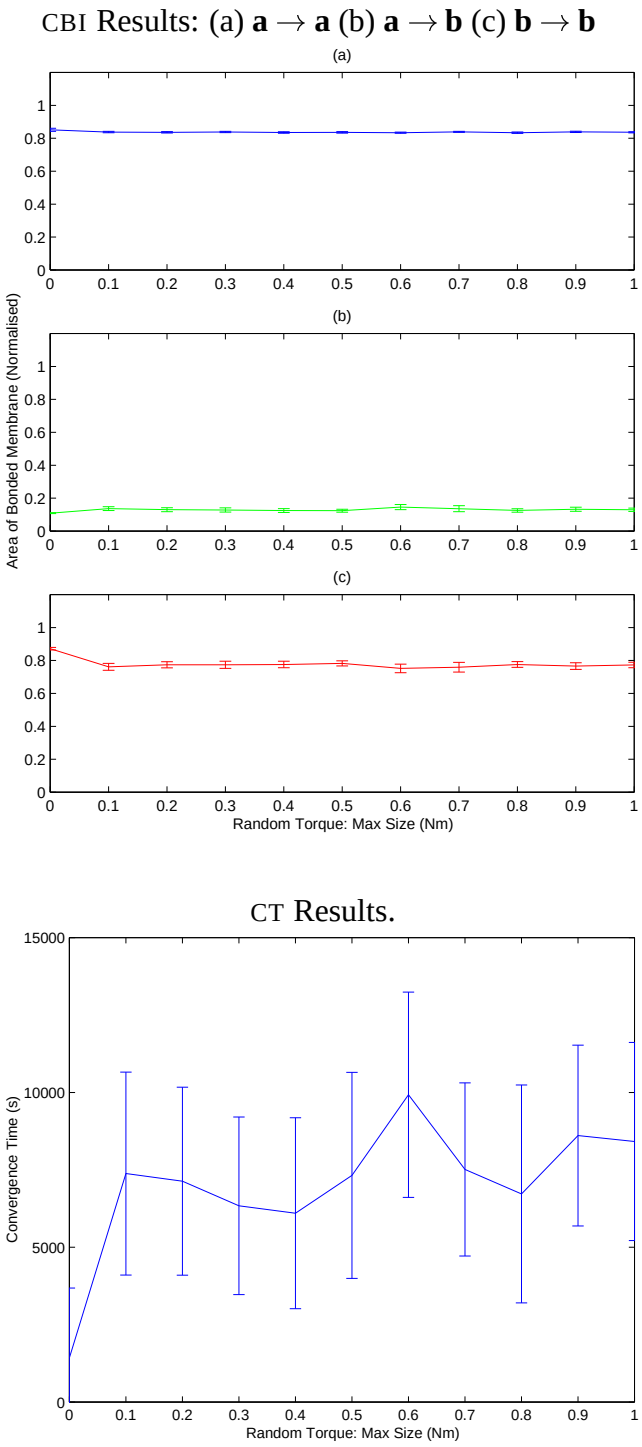


Figure B.20: *Results for 64 A-Cell aggregates over a range of random movement torque limits. The aggregates started from the initial configuration of a sorted grid and should move toward the onion equilibrium state.*

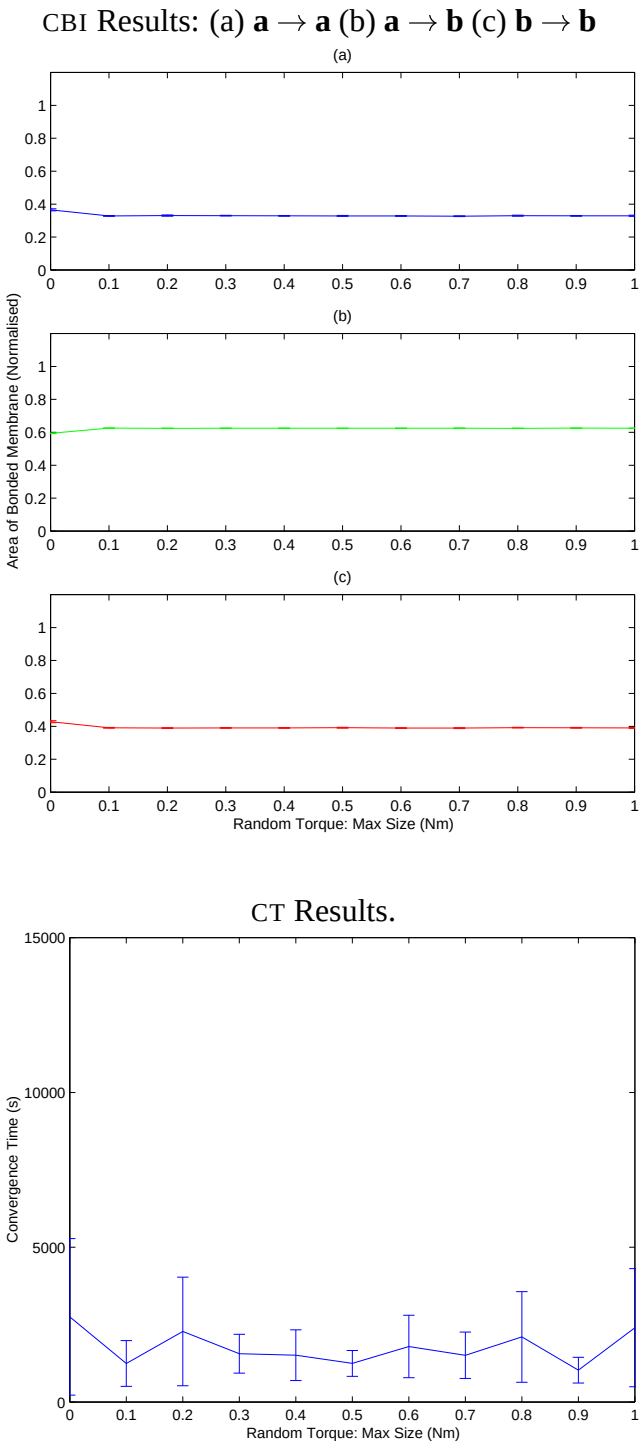


Figure B.21: *Results for 64 A-Cell aggregates over a range of random movement torque limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the intermixed equilibrium state.*

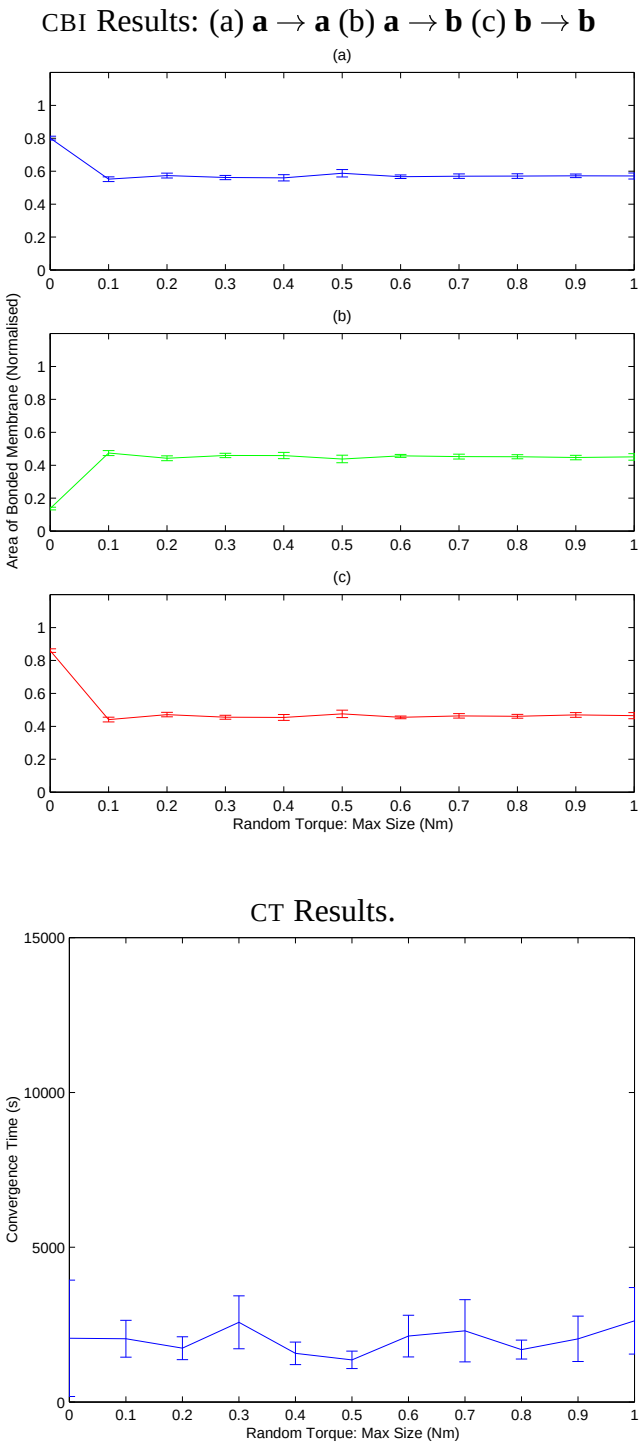
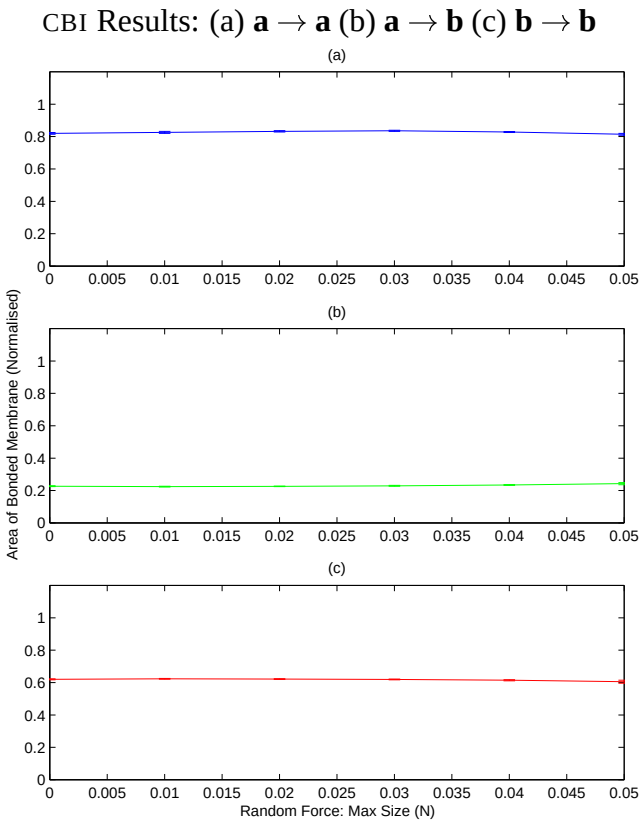


Figure B.22: *Results for 64 A-Cell aggregates over a range of random movement torque limits. The aggregates started from the initial configuration of a sorted grid and should move toward the intermixed equilibrium state.*



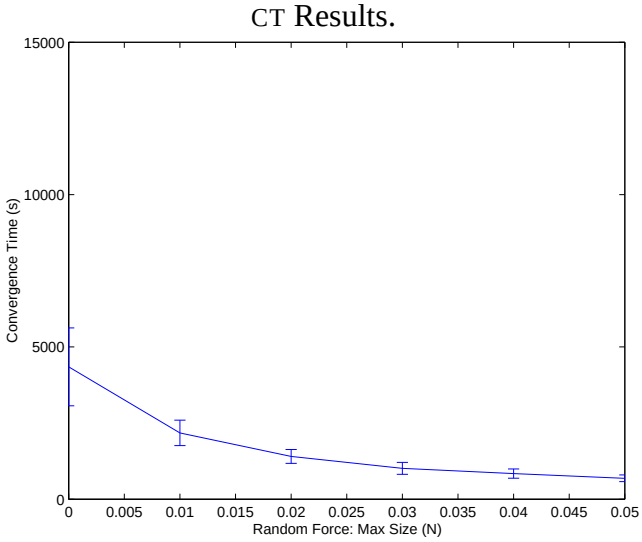
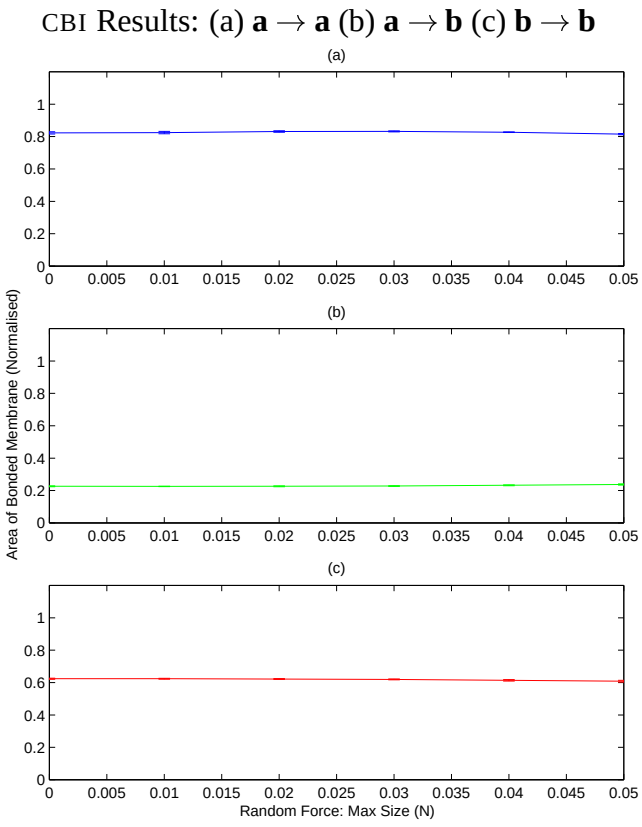


Figure B.23: *Results for 64 A-Cell aggregates with a fixed random torque limit over a range of random force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the onion equilibrium state.*



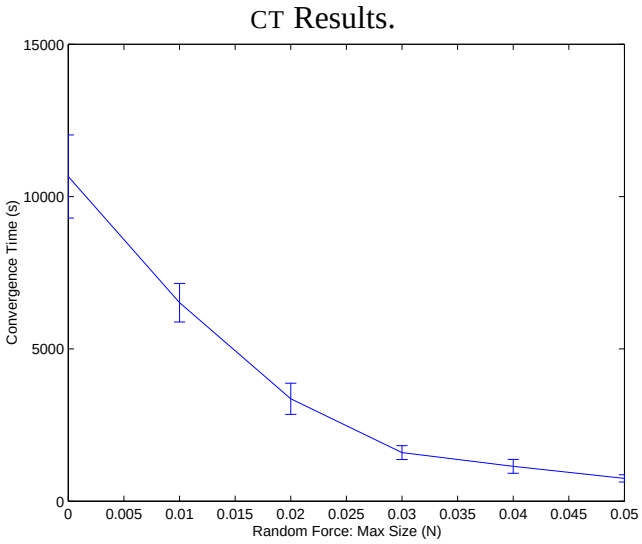


Figure B.24: *Results for 64 A-Cell aggregates with a fixed random torque limit over a range of random force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the onion equilibrium state.*

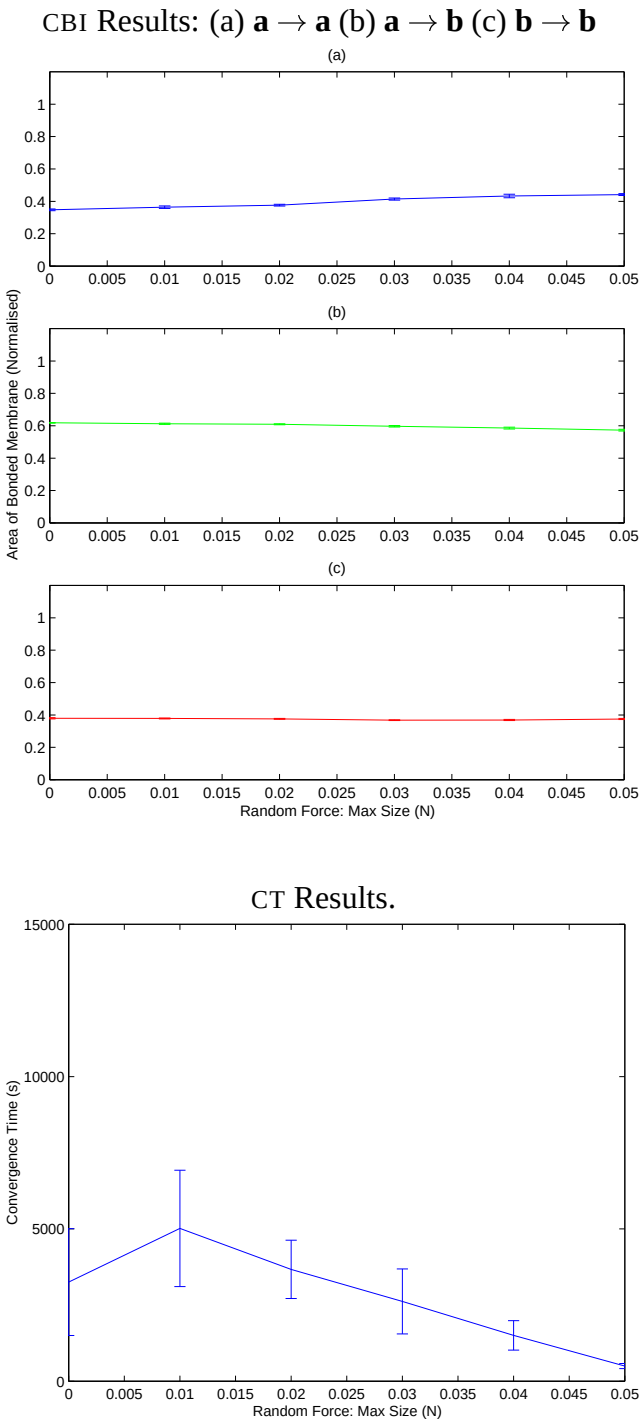


Figure B.25: *Results for 64 A-Cell aggregates with a fixed random torque limit over a range of random force limits. The aggregates started from the initial configuration of an intermixed grid and should move toward the inter-mixed equilibrium state.*

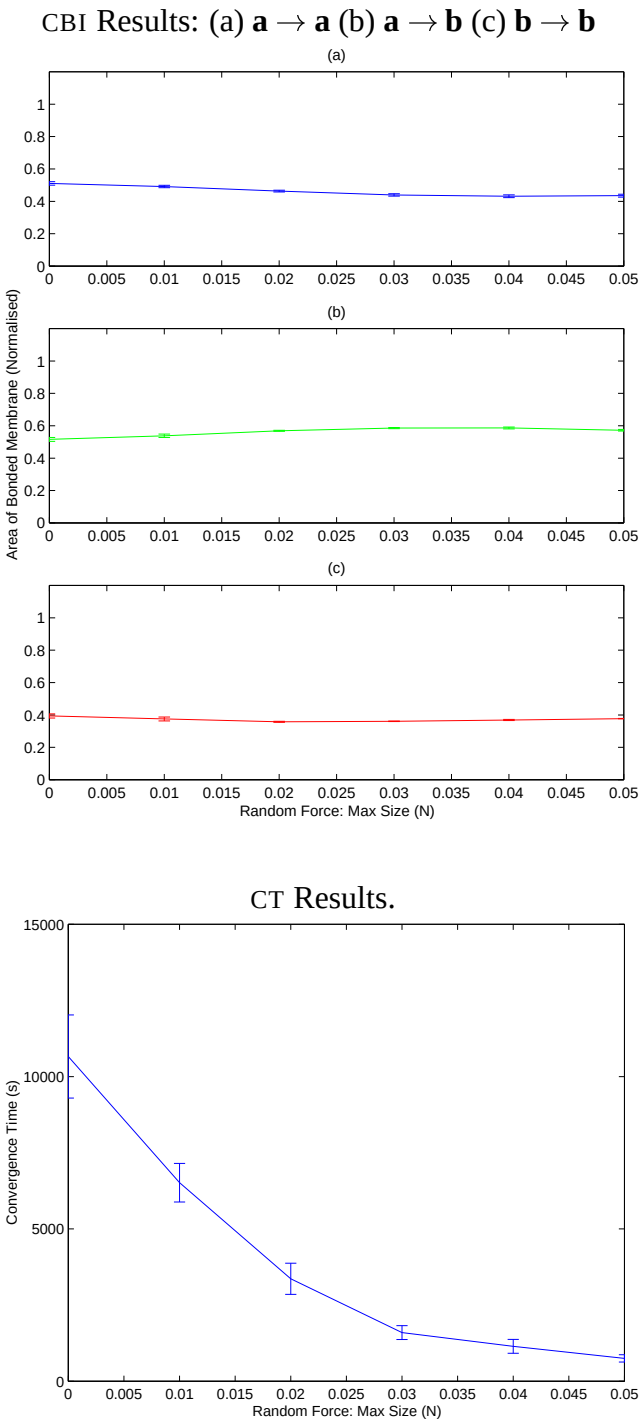


Figure B.26: *Results for 64 A-Cell aggregates with a fixed random torque limit over a range of random force limits. The aggregates started from the initial configuration of a sorted grid and should move toward the intermixed equilibrium state.*

Appendix C

Publications

- Peter Ottery and John Hallam. Steps Toward Self-reconfigurable Robot Systems by Modelling Cellular Adhesion Mechanisms. In Proceedings of the The 8th Conference on Intelligent Autonomous Systems (IAS-8), Pages 720-728, Amsterdam 2004

Steps Toward Self-reconfigurable Robot Systems by Modelling Cellular Adhesion Mechanisms

Peter Ottery & John Hallam
IPAB, School of Informatics, The University of Edinburgh
peter.ottery@ed.ac.uk jhallam@inf.ed.ac.uk

Abstract.
This paper presents a distributed control method for a collection of homogeneous robotic modules. We propose that by modelling simple cellular adhesion mechanisms, collections of these modules can self-assemble into virtually bonded aggregates which behave as predicted by Steinberg's 'differential adhesion hypothesis'. This allows simple self-repairing hierarchical structures to be formed, the exact properties of which can be controlled by varying the adhesive properties of the individual modules. Finally we present the results of some initial simulations which demonstrate that this model responds positively to small amounts of turbulence.

1 Introduction

Self-reconfigurable robotic systems have the ability to adapt to the needs of various tasks or operating environments. They consist of many independent modules which can reconfigure to change the overall geometric structure of the system. This shape changing ability allows these systems to perform many complex, yet desirable functions such as self-repair and self-assembly.

Many of the existing systems [1, 2, 3, 4] are composed of *homogeneous* units, making it possible for any unit in the system to replace any other unit. However, in some cases the control systems that guide the units are still based on centralised planning. If instead, the control is distributed then no individual unit is critical to the system, adding a higher level of robustness.

Much of the inspiration for these fully distributed systems has been derived from naturally occurring phenomena. At the lowest level the forces between individual molecules are modelled allowing the "artificial physics" to drive the units into simple geometric structures [5]. At the other end of the spectrum systems are based on the simple interactions between social insects which allow the formation of various connected and unconnected structures [6].

Our approach to self-reconfiguration, which is also inspired by nature, is to attempt to model some of the organisational properties of biological cells. These cells rearrange themselves in extremely complex ways to produce remarkably complicated organisms and they achieve this with an extremely high degree of accuracy.

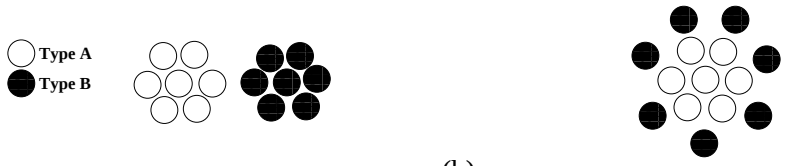
The most basic organisational behaviour exhibited by cells is due to the adhesion forces generated between them. This behaviour can adequately be explained by the 'differential adhesion hypothesis' (DAH) [7] which shows how simple hierarchical patterns can be formed simply due to the differences in the adhesion forces between different cell types.

In this paper we present a model of some of the interaction properties of biological cells. The model we propose provides a mechanism for controlling a collection of robotic modules in a complex fluid environment, allowing them to create and maintain simple hierarchical structures. This is achieved by making the modules behave like simple artificial cells which attempt to model cellular membranes and adhesions between those membranes, such that they behave as predicted by the differential adhesion hypothesis.

2 Differential Cellular Adhesion

The 'differential adhesion hypothesis', proposed by Steinberg[7], suggests that cells rearrange to minimise their free energy and thus form the most thermodynamically stable configuration.

Considering an aggregate of cells from the same tissue, held in suspension, this minimisation of free energy is demonstrated by the rounding that occurs. The cells attempt to maximise the adhesions they make and thus form the most compact structure - a sphere.



(a) (b)

Figure 1: Two equilibrium states proposed by the DAH.

The more interesting case occurs when two or more tissues are intermixed. For example, if two cell types (A and B) are mixed, initially the bonding between the cells will be random. However, if there exists a difference in the strength of the bonds then a gradual selection of the strongest bonds will cause the cells to rearrange into a more stable configuration. For example if homogeneous (A-A, B-B) bonds are stronger than heterogeneous bonds (A-B) the cells will attempt to sort into pure populations of the different types (fig. 1(a)). If, however, the A-A bonds

are stronger than the A-B bonds which are in turn stronger than the B-B bonds, then the A cells will migrate centrally while the B cells form a shell around them (fig. 1(b)).

Although, Steinberg proposes that the above behaviours are driven by surface energies he does not clarify the mechanism that allows them to drive the cells. There is now growing evidence that small random movements of the cells caused by membrane ruffling could be an essential ingredient [8]. These movements allow the cells to explore their neighbourhood and thus select stronger bonds.

3 Adhesion Model

Our model provides a control method for a collection of robotic modules. As these modules represent an extremely simplified model of a biological cell we refer to them throughout this paper as A-Cells (Artificial Cells).

The A-Cells are based on the design of the HYDRON module which is a small, near spherical, sub-aquatic robot currently being designed by the HYDRA consortium[1]. The module is capable of independent three dimensional movement. Additionally, neighbouring modules will be able to communicate using eight optical transmitter, receiver pairs located on the HYDRON's hull at the corners of a cube.

3.1 Communication

The A-Cells rely on a relatively simple communication system. Each A-Cell continually broadcasts a single 'transmission signal', with a limited range, from each of its transmitters. This signal contains both information relating directly to the broadcasting A-Cell and a list carrying information aimed at individual cells in the system.

In addition to data transfer the transmission signal also allows A-Cells to determine an approximate position of their neighbours. The intensity of the signal gives a crude indication of the distance it has travelled while the set of detectors that are receiving the signal indicate the general direction of the source.

3.2 Membrane Model

One of the most important structures in living cells is the deformable membrane which surrounds the cell nucleus. Our model provides a very simple approximation of such a membrane by using the A-Cells' communication system to create a virtual force field which extends some distance from the actual A-Cell hull. Like a fluid membrane this force field behaves like a viscous spring, resisting compression from

[1]The HYDRA consortium home page is available at 'http://www.hydra-robot.com/'.

similar force fields generated by neighbouring A-Cells with a force proportional to the speed of the compression.

$$R = (a * m) + (d * v) \quad (1)$$

where R is the repulsion force acting along the line joining the two A-Cell centres (Kgms^{-2}), a the area of membrane (m^2) that has been compressed, m the repulsion factor (Kgm^{-1}s^{-2}, a system parameter), d the damping coefficient (Kgs^{-1}, also a system parameter) and v is the estimated relative velocity (ms^{-1}) concerned along the line joining their centres.

Each A-cell transmits to its neighbours the radius of its physical hull and the distance its membrane (virtual force field) extends from this hull. Therefore, by comparing the approximate distance of neighbouring A-Cells with their relative sizes, an A-Cell can detect membrane collisions and estimate the area of membrane that has been compressed. Also, by monitoring how the distance of each neighbouring A-Cell changes over time it is possible to calculate their relative speeds, and thus the viscous component of the repulsive force.

To ensure that two colliding A-Cells experience an equal and opposite repulsive force some further communication is required. Each of the colliding A-Cells add the repulsive force they calculated to their transmission signal along with the ID of the A-Cell involved in the collision. Each A-Cell can then simply average its own calculated value with that of the other giving equivalent results.

One flaw with this simple model is that each membrane collision is considered separately. Therefore, it is possible for the same area of membrane to collide with more than a single other membrane simultaneously. However, the principal repulsive behaviours of the membrane remain consistent.

3.3 A-CAMs

The membranes of real cells use Cellular Adhesion Molecules (CAMs) to form bonds which are both reversible and selective. We attempt to model these key properties through the use of A-CAMs. Like CAMs, the A-CAMs are capable of both homophilic or heterophilic binding. However, unlike some real CAMs each A-CAM may only bond with a single type. This greatly simplifies the implementation without restricting the power of the model.

When the A-CAMs form bonds they produce a small attractive force pulling the two areas expressing the A-CAMs closer together. Thus bonded A-Cell membranes are pulled together until the attractive force is balanced by the repulsive force generated by the compression.

Clearly when two areas of real cell membrane come into contact only some subset of the CAMs present on each membrane will be apposed and thus able to bond. The size of this subset is dependent on both the number of CAMs on each area

of membrane and the strength of the bond between the CAMs. Greater densities of CAMs will increase the chance that any two are apposed and a greater bond strength will allow the CAMs to seek each other out more assiduously.

To express these relationships with the A-CAMs we use the following equation which exhibits the necessary properties when relating numbers of available A-CAMs to numbers that bond.

$$N = \min(A_{n1}, A_{n2}) \operatorname{sigmoid}\left(m(n1, n2)k + \frac{(A_{n1} - A_{n2})^2}{A_{n1} * A_{n2}}\right) \quad (2)$$

where A_{n1} is the number of A-CAMs of type $n1$ on membrane 1, A_{n2} is the corresponding number of type $n2$ A-CAMs on membrane 2, $m(x, y)$ is the attractive force between A-CAMs of type x and y and k is a system parameter that controls the influence of the bonding strength on bonding numbers.

3.4 Adhesion Sites

The A-Cell membrane model allows the same area of membrane to be in contact with more than a single other membrane simultaneously (see §3.2). However, to achieve more realistic selective bonding, the bonding of each area should be restricted to only one of the membranes in contact.

This is achieved by dividing the membrane surface into a number of equally sized patches or adhesion sites, such that each adhesion site can only bond with a single membrane. If an adhesion site comes into contact with more than a single membrane it makes a probabilistic decision as to with which it should bond. This decision is based on the estimated bond strength of a bond between a unit area of each membrane. To calculate this value, each A-Cell includes in its transmission signal the surface density of every A-CAM it is expressing. These values can then be compared with those of any A-Cell in contact and the estimated bond strength can be calculated.

Having established which adhesion sites wish to bond on each membrane it is then necessary to calculate the number of A-CAMs this represents. As we assume that any A-CAMs being expressed are evenly distributed across the membrane surface this is relatively straight forward. Once calculated these values are communicated between the colliding A-Cells by including them in their respective transmission signals. Each A-Cell then compares each pair of bonding A-CAMs in turn and calculates the total attractive force that will be produced.

This algorithm makes the following two simplifications:

1. It is assumed that adhesion sites that decide to bond on two colliding membranes appose each other. However, it is clear that this might not be the case.

2. Adhesion sites which decide to bond to a specific membrane are not reallocated if unused.

Avoiding these simplifications would require far more complexity and would greatly increase the communication required between A-Cells.

3.5 Random Motion

In §2 we mentioned that random cell movements may play an important role in cell rearrangement. Although turbulence present in the environment may be sufficient, this has also been accommodated in our model by allowing the A-Cells to add some small random force, below some predefined maximum, to each force they produce.

3.6 Model Summary

Each A-Cell repeatedly cycles through the procure of detecting membrane collisions and calculating the associated forces they generate. This has been summarised in figure 2.

4 Simulation Results - 2D Rounding

In this section we present some of our initial results which show the rounding behaviour of the model for simulations carried out in a water like fluid with A-Cell movements constrained to two dimensions.

In each simulation the actual A-Cell hull has a radius of 0.05m and the virtual membrane extends a further 0.5m. The A-Cell mass has been set to 1Kg. Additionally, every A-Cell has been allocated the same membrane properties such that the equilibrium distance (the point where the attractive forces are balanced by the repulsive forces) is 0.55m and the attraction between two cells at this distance is 0.15125N. Finally the membrane's damping coefficient has been set to 0.4Kgs^{-1} and 64 adhesion sites have been used (in two dimensions the adhesion sites divide the equator of the A-Cell membrane).

To investigate the rounding behaviour of an A-Cell aggregate we have used the following compactness measure:

$$C = \frac{4 * \pi * a}{p^2} \tag{3}$$

where a is the area of the aggregate (m^2) and p its perimeter (m). This intuitive measure gives an index between 0 and 1, with 1 the most compact (spherical) and 0 the least compact (linear). Figure 3 shows how this measure changes during a typical run of the simulator with 36 A-Cells starting in a perfect line such that they are the equilibrium distance apart. This demonstrates the common pattern where the A-Cell aggregate quickly increases in compactness before converging to some maximum value.

```
1. Check which signals can be detected and calculate the
   approximate distance and direction of their source.

2. Determine if there are any membrane collisions.

3. For each A-Cell in contact:

   (a) Compare the values in the A-Cell's transmission
       signal with the values being broadcast from this
       A-Cell and calculate any repulsive or attractive
       forces.

4. Sum any repulsive or attractive forces to calculate a
   single force vector.

5. Add any random force to the force vector.

6. For each A-Cell in contact:

   (a) Calculate the repulsive force generated by com-
       pressing the membranes.

   (b) Calculate the estimated bond strength.

7. For each adhesion site:

   (a) Probabilistically decide which membrane to bond
       to.

8. Construct the new transmission signal.

9. Generate the final force.
```

Figure 2: A summary of the procedure employed by the adhesion model.

4.1 Rotational Stability

An interesting observation of this model is the emergence of a global rotation of the aggregate. In the presence of small random forces such as those discussed in following section, the virtual aggregates can slowly begin to rotate around their centre of mass and in some cases a slow build up of forces can cause the rotation to accelerate until the aggregate pulls itself apart.

To prevent this build up of forces we investigated both increasing the time delay between which the A-Cells produce each force and placing a maximum limit on the size of the forces that could be produced. As expected, the results showed that both reducing the maximum force limit and increasing the time delay reduces the average

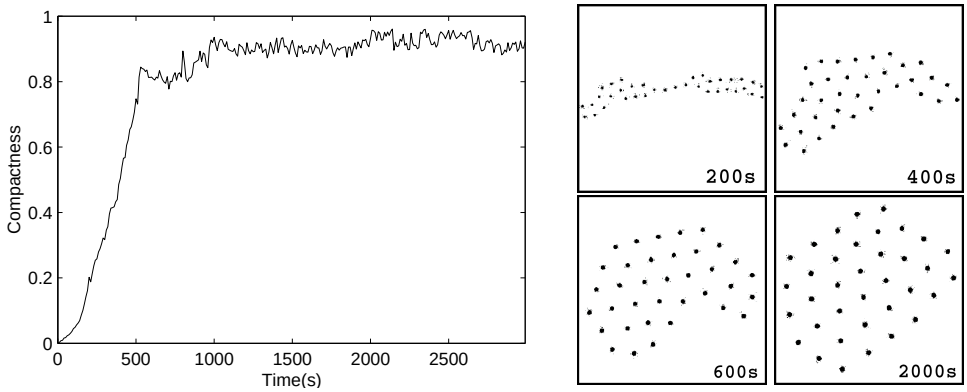


Figure 3: Compactness evolution over time for an example run with parameters discussed in the text. The images on the right show stages in the convergence of the aggregate.

angular speed of the A-Cells. However, each of these options also reduce the rate at which the A-Cells round. A sensible choice of values which greatly reduce the angular speed of the A-cells with minimal impact on their rounding ability are a maximum force of 0.01N and a delay between forces of 0.1s.

4.2 Random

As we have discussed, random movements are thought to aid cell rearrangements (see §2. Here we show the result of adding either small random forces or small amounts of random torque to the A-Cells. Each of these could be caused by the environment or, in the case of the random forces, by the A-Cells themselves. Again the starting formation for each simulation was a perfect line, such that the A-Cells are the equilibrium distance apart. Additionally, to reduce the rotational behaviour of the aggregate, a maximum force of 0.01N and a delay between forces of 0.1s were used. The rounding ability of the simulations is measure by the mean rate at which the A-Cell aggregates round (Compactness/Time) until the compactness value converges (see fig. 4(left)) and the convergence value itself (see fig. 4(right)).

Adding larger random forces should allow the A-Cells to explore more and thus improve their rounding ability. Figure 4 shows there is a slight increase in the compactness of the resulting aggregates as the random forces increase. However, random forces greater than 0.035N caused some A-Cells to break away from the main aggregate.

Random torque, on the other hand, does not directly cause the A-Cells to make random movements. However, by causing the A-Cells' sensors to become misaligned it allows them to form less rigid structures. Figure 5 shows that the level of torque has little effect on the performance of the A-Cells. However, the presence of some torque allows the aggregates to become far more compact.

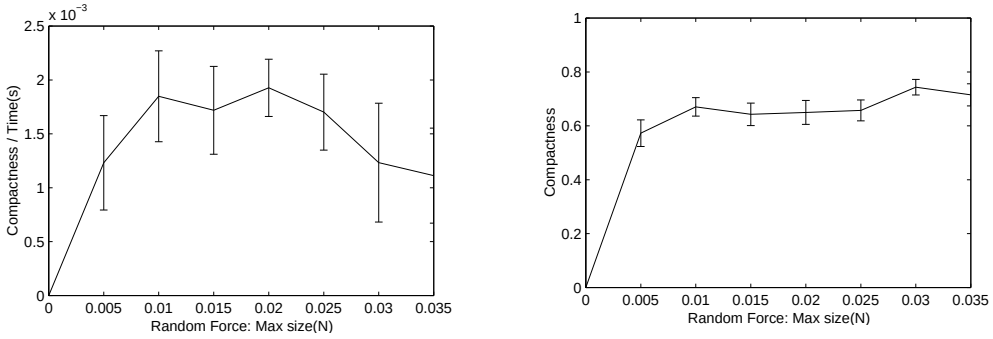


Figure 4: The effect of random forces on the model's rounding ability.

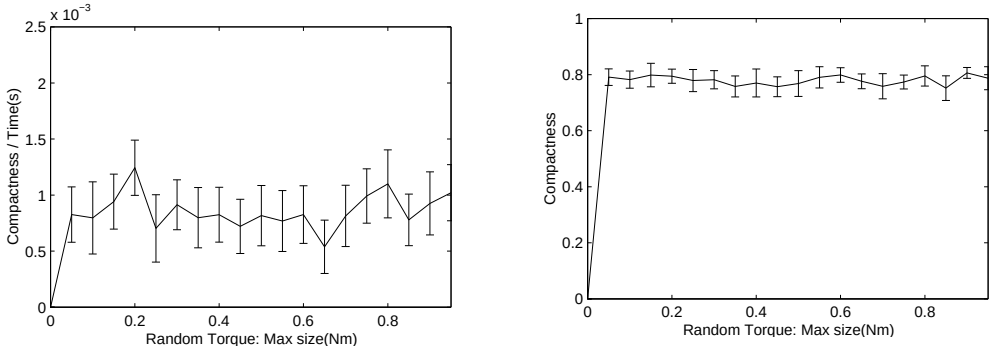


Figure 5: The effect of random torque on the model's rounding ability

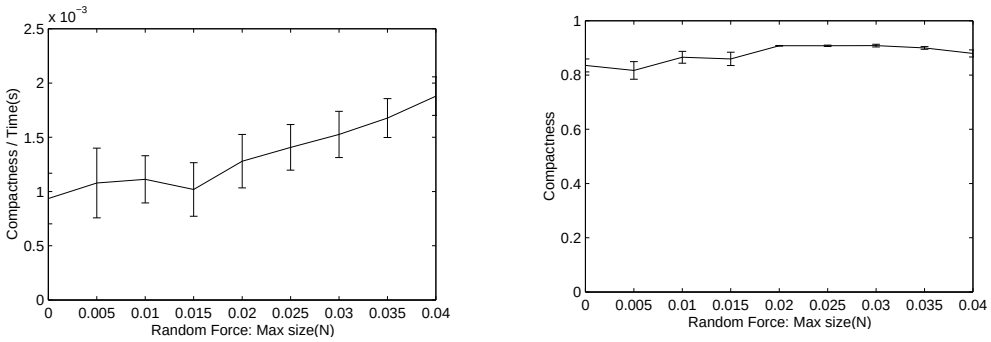


Figure 6: The effect of random forces in addition to 0.01Nm random torque on the model's rounding ability.

The best performance is achieved by a combination of both random torque and random forces. In figure 6 the maximum size of the random torque has been set to 0.01Nm. In this case the resulting aggregates are the most compact, with the rate at which they compact increasing with the level of the random forces being applied. Again, once the random forces become too large ($>$0.04N), A-Cells begin to break away from the main aggregate.

5 Summary

This paper proposed a distributed control method based on the DAH which allows a collection of homogeneous robotic modules to self-assemble into simple hierarchical structures. By allowing the units to continue to search for the most stable configuration, even after these structures are formed, this method also has a degree of self-repair built in. Additionally, we have presented initial results which suggest that our model does not require perfect conditions to operate. In fact small random forces and small amounts of random torque appear to aid the model allowing it to form more compact structures in a shorter time period.

References

[1] Akiya Kamimura, Satoshi Murata, Eiichi Yoshida, Haruhisa Kurokawa, Kohji Tomita, and Shigeru Kokaji. Self-reconfigurable modular robot - experiments on reconfiguration and locomotion. In *IEEE/RSJ International International Conference on Intelligent Robots and Systems*, pages 606–612, 2001.

[2] Satoshi Murata, Haruhisa Kurokawa, Eiichi Yoshida, Kohji Tomita, and Shigeru Kokaji. A 3-d self-reconfigurable structure. In *IEEE International Conference on Robotics and Automation*, pages 432–439, May 1998.

[3] Daniela Rus and Marsette Vona. A basis for self-reconfiguring robots using crystal modules. In *IEEE/RSJ Intelligent Robots and Systems*, volume 3, pages 2184 –2193, 2000.

[4] Kohji Tomita, Satoshi Murata, Haruhisa Kurokawa, Eiichi Yoshida, and Shigeru Kokaji. Self-assembly and self-repair method for a distributed mechanical system. *IEEE Transactions on Robotics and Automation*, 15(6):1035–1045, December 1999.

[5] William M. Spears and Diana F. Gordon. Using artificial physics to control agents. In *International Conference on Information Intelligence and Systems*, 1999.

[6] E. Şahin, T.H. Labella, V. Trianni, J.-L. Deneubourg, P. Rasse, D. Floreano, L. Gambardella, F. Mondada, S. Nolfi, and M. Dorigo. SWARM-BOTS: Pattern formation in a swarm of self-assembling mobile robots. In A. El Kamel, K. Mellouli, and P. Borne, editors, *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Hammamet, Tunisia, October 6-9, 2002. Piscataway, NJ: IEEE Press.

[7] Malcolm S. Steinberg. Reconstruction of tissues by dissociated cells. *Science*, 141(3579):401–408, August 1963.

[8] Jose C.M. Mombach and James A. Glazier. Single cell motion in aggregates of embryonic cells. *Physical Review Letters*, 76(16):3032–3035, April 1996.

Glossary

A-CAM A-Cell Adhesion Molecule: A-CAMs are expressed on the virtual membrane of the A-Cells. They approximate the Cellular Adhesion Molecules (CAMs) expressed on real cell membranes and are capable of both homophilic or heterophilic binding. When the membranes of two A-Cells collide, some subset of the A-CAMs on the areas in contact will bond generating an attractive force which pulls the two A-Cells toward each other.

A-Cell A simple simulated robotic module used throughout this thesis as a testbed for the proposed cellular adhesion model. Each A-Cell consists of a spherically shaped hull and is capable of movement in three dimensions. A number of transmitter/receiver pairs located on the A-Cell's hull allow local communication between neighbouring units.

Adhesion Site An adhesion site controls a region of an A-Cell's virtual membrane and all the A-CAMs present on that region. When the adhesion site comes into contact with other membranes it probabilistically selects which one the A-CAMs will be able to bond with based on the estimated attraction that could be generated. In this way, adhesion sites provide a simple mechanism which prevents the same area of an A-Cell's membrane from forming multiple bonds simultaneously.

Bootstrapping A method of empirically assessing the uncertainty of a small sample without making assumptions about its distribution.

CAM Cell Adhesion Molecule: CAMs are proteins located on the membrane of biological cells. These proteins generate adhesions between the cells, allowing them to form more complex tissues.

CBI Converged Bonding Indices: A scaled measure of the level of bonding which is observed between two of the A-Cell types present in an aggregate once the aggregate reaches an equilibrium configuration.

CC Convergence Compactness: A measure of the compactness value calculated for the equilibrium configuration that is achieved by an aggregate of A-Cells.

Compactness A measure of the 'roundness' of an aggregate.

CR Compacting Rate: The average change in the compactness value of an A-Cell aggregate as it converges to a stable configuration.

CT Convergence Time: A measure of the total time required for an aggregate of A-Cells to reach an equilibrium configuration.

DAH Differential Adhesion Hypothesis: The hypothesis proposed by Steinberg [54] to explain how differences in the adhesion between cells from different tissues can cause simple hierarchical patterns to form when the tissues are mixed.

FOV Field of View: The cone describing the portion of the world that a sensor observes.

GRN Genetic Regulatory Network: A GRN describes the complex interaction between the proteins found in a cell and level of expression of the cell's genes.

Heterotypic A term which describes adhesions formed between cells of different types.

Homotypic A term which describes adhesions formed between cells of the same type.

Bibliography

[1] Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994). *Molecular Biology of the Cell*. Garland Publishing, 3rd edition.

[2] Antonelli, P. L., Rogers, T. D., and Willard, M. A. (1973). Geometry and the exchange principle in cell aggregation kinetics. *Journal of Theoretical Biology*, 41:1–21.

[3] Balch, T. and Hybinette, M. (2000a). Behavior-based coordination of large-scale robot formations. In *Fourth International Conference on MultiAgent Systems*, pages 363–364.

[4] Balch, T. and Hybinette, M. (2000b). Social potentials for scalable multi-robot formations. In *ICRA '00. IEEE International Conference on Robotics and Automation*, volume 1, pages 73–80.

[5] Bongard, J. (2002). Evolving modular genetic regulatory networks. In *IEEE 2002 Congress on Evolutionary Computation (CEC2002)*, pages 1872–1877.

[6] Brodland, G. W. (2004). Computational modeling of cell sorting, tissue engulfment, and related phenomena: A review. *Applied Mechanics Reviews*, 57(1):47–76.

[7] Brodland, G. W. and Chen, H. H. (2000). The mechanics of heterotypic cell aggregates: insights from computer simulations. *Journal of Biomechanical Engineering*, 122(4):402–407.

[8] Chen, H. H. and Brodland, G. W. (2000). Cell-level finite element studies of viscous cells in planar aggregates. *Journal of Biomechanical Engineering*, 122(4):394–401.

[9] CMLabs (2002). *VORTEX Developer Guide*. CMLabs Simulations Inc., Montreal, Canada, 2.0.1 edition.

[10] Şahin, E., Labella, T., Trianni, V., Deneubourg, J.-L., Rasse, P., Floreano, D., Gambardella, L., Mondada, F., Nolfi, S., and Dorigo, M. (2002). SWARM-BOTS: Pattern formation in a swarm of self-assembling mobile robots. In El Kamel, A., Mellouli, K., and Borne, P., editors, *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Hammamet, Tunisia. Piscataway, NJ: IEEE Press.

[11] Cumberbatch, E. and Fitt, A. (2001). *Mathematical Modeling: Case Studies from Industry*. Cambridge University Press.

[12] Davis, G. S., Phillips, H. M., and Steinberg, M. S. (1997). Germ-layer surface tensions and "tissue affinities" inrana pipiensgastrulae: Quantitative measurements. *Developmental Biology*, 192:630–644.

[13] Dey, T. K. and Goswami, S. (2003). Tight cocone: A water-tight surface reconstructor. In *SM '03: Proceedings of the eighth ACM symposium on Solid modeling and applications*, pages 127–134. ACM Press.

[14] Efron, B. and Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, 37(1):36–48.

[15] Eggenberger, P. (1997). Evolving morphologies of simulated 3d organisms based on differential gene expression. In *4th European Conference on Artificial Life*.

[16] Foty, R. and Steinberg, M. S. (2005). The differential adhesion hypothesis: a direct evaluation. *Developmental Biology*, 278:255–263.

[17] Foty, R. A., Forgacs, G., Pfleger, C. M., and Steinberg, M. S. (1994). Liquid properties of embryonic tissues: Measurement of interfacial tensions. *Physical Review Letters*, 72(14):2298–2301.

[18] Fredslund, J. and Matarić, M. J. (2002). A general algorithm for robot formations using local sensing and minimal communication. *IEEE Transactions on Robotics and Automation*, 18(5):837–846.

[19] Fujibayashi, K., Murata, S., Sugawara, K., and Yamamura, M. (2002). Self-organizing formation algorithm for active elements. In *21st IEEE Symposium on Reliable Distributed Systems*, pages 416–421.

[20] Gilbert, S. F. (2003). *Developmental Biology*. Sinauer Associates, Inc., seventh edition.

[21] Glazier, J. A. and Graner, F. (1993). Simulation of differential adhesion driven rearrangement of biological cells. *Physical Review E*, 47(3).

[22] Glazier, J. A., Raphael, R., Graner, F., and Sawada, Y. (1995). The energetics of cell sorting in three dimensions. In *Interplay of Genetic and Physical Processes in the Development of Biological Form*, pages 54–61.

[23] Goel, N. S. and Rogers, G. (1978). Computer simulation of engulfment and other movements of embryonic tissues. *Journal of Theoretical Biology*, 71(1):103–140.

[24] Graham, R. L. (1972). An efficient algorithm for determining the convex hull of a finite planar set. *Inf. Process. Lett.*, 1(4):132–133.

[25] Graner, F. and Sawada, Y. (1993). Can surface adhesion drive cell rearrangement? Part II: a geometrical model. *Journal of Theoretical Biology*, 164:477–506.

[26] Green, P. J. and Silverman, B. W. (1979). Constructing the convex hull of a set of points in the plane. *Comput. J.*, 22(3):262–266.

[27] Groß, R. and Dorigo, M. (2004). Group transport of an object to a target that only some group members may sense. In *Parallel Problem Solving from Nature – 8th International Conference (PPSN VIII)*, volume 3242 of *Lecture Notes in Computer Science*, pages 852–861. Springer Verlag, Berlin, Germany.

[28] Guigue, P. and Devillers, O. (2003). Fast and robust triangle-triangle overlap test using orientation predicates. *Journal of Graphics Tools*, 8(1):25–42.

[29] Honda, H. (1978). Description of cellular patterns by dirichlet domains: the two-dimensional case. *Journal of Theoretical Biology*, 72:523–543.

[30] Honda, H. (1983). Geometrical models for cells in tissues. *International Review of Cytology*, 81:191–248.

[31] Honda, H., Yamanaka, H., and Egushi, G. (1986). Transformation of a polygonal cellular pattern during sexual maturation of the avian oviduct epithelium: computer simulation. *Journal of Embryology and Experimental Morphology*, 98:1–19.

[32] Hosokawa, K., Tsujimori, T., Fujii, T., Kaetsu, H., Asama, H., Kuroda, Y., and Endo, I. (1998). Self-organizing collective robots with morphogenesis in a vertical plane. In *IEEE International Conference on Robotics and Automation*, pages 2858–2863.

[33] Jarvis, R. A. (1973). On the identification of the convex hull of a finite set of points in the plane. *Inf. Process. Lett.*, 2(1):18–21.

[34] Jørgensen, M. W., Østergaard, E. H., and Lund, H. H. (2004). Modular atron: Modules for a self-reconfigurable robot. In *IEEE/RSJ International Conference on Robots and Systems (IROS)*, pages 2068–2073.

[35] Kamimura, A., Murata, S., Yoshida, E., Kurokawa, H., Tomita, K., and Kokaji, S. (2001). Self-reconfigurable modular robot - experiments on reconfiguration and locomotion. In *IEEE/RSJ International International Conference on Intelligent Robots and Systems*, pages 606–612.

[36] Kondacs, A. (2003). Biologically-inspired self-assembly of two-dimensional shapes using global-to-local compilation. In *International Joint Conference on Artificial Intelligence*, pages 633–638.

[37] Kotay, K., Rus, D., Vona, M., and McGray, C. (1998). The self-reconfiguring robotic molecule. In *IEEE International Conference on Robotics and Automation*.

[38] Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, S. L., and Darnell, J. (2003). *Molecular Cell Biology*. W. H. Freeman and Company, 5th edition.

[39] Mamei, M., Vasirani, M., and Zambonelli, F. (2004). Experiments of morphogenesis in swarms of simple mobile robots. *Journal of Applied Artificial Intelligence*, 18(9–10).

[40] Mombach, J. C. and Glazier, J. A. (1996). Single cell motion in aggregates of embryonic cells. *Physical Review Letters*, 76(16):3032–3035.

[41] Mombach, J. C. M., Glazier, J. A., Raphael, R. C., and Zajac, M. (1995). Quantitative comparison between differential adhesion models and cell sorting in the presence and absence of fluctuations. *Physical Review Letters*, 75(11):2244–2247.

[42] Murata, S., Kurokawa, H., Yoshida, E., Tomita, K., and Kokaji, S. (1998). A 3-d self-reconfigurable structure. In *IEEE International Conference on Robotics and Automation*, pages 432–439.

[43] Nagpal, R., Shrobe, H., and Bachrach, J. (2003). Organizing a global coordinate system from local information on an ad hoc sensor network. In *2nd International Workshop on Information Processing in Sensor Networks (IPSN '03)*.

[44] Ottery, P. and Hallam, J. (2004). Steps toward self-reconfigurable robot systems by modelling cellular adhesion mechanisms. In *The 8th Conference on Intelligent Autonomous Systems (IAS-8)*, pages 720–728.

[45] Palsson, E. (2001). A three-dimensional model of cell movement in multicellular systems. *Future Generation Computer Systems*, 17(7):835–852.

[46] Purves, W. K. and Orians, G. H. (1983). *Life: The Science of Biology*. Sinauer Associates, Inc. Publishers and Willard Grant Press.

[47] Quick, T., Nehaniv, C. L., Dautenhahn, K., and Roberts, G. (2003). Evolving embodied genetic regulatory network-driven control systems. In *Seventh European Conference on Artificial Life (ECAL2003)*.

[48] Rieu, J. P., Kataoka, N., and Sawada, Y. (1998). Quantitative analysis of cell motion during sorting in two-dimensional aggregates of dissociated hydra cell. *Physical Review E*, 57(1):924–931.

[49] Rogers, G. and Goel, N. S. (1978). Computer simulation of cellular movements: Cell-sorting, cellular migration through a mass of cells and contact inhibition. *Journal of Theoretical Biology*, 71(1):141–166.

[50] Rus, D. and Vona, M. (2000). A basis for self-reconfiguring robots using crystal modules. In *IEEE/RSJ Intelligent Robots and Systems*, volume 3, pages 2184 –2193.

[51] Saff, E. B. and Kuijlaars, A. B. J. (1997). Distributing many points on a sphere. *Mathematical Intelligencer*, 19:5–11.

[52] Spears, W. and Gordon, D. (1999). Using artificial physics to control agents. In *IEEE International Conference on Information, Intelligence, and Systems*.

[53] Spears, W., Spears, D., Hamann, J., and Heil, R. (2004). Distributed, physics-based control of swarms of vehicles. *Autonomous Robots*, pages 137–162.

[54] Steinberg, M. S. (1963). Reconstruction of tissues by dissociated cells. *Science*, 141(3579):401–408.

[55] Steinberg, M. S. (1970). Does differential adhesion govern self-assembly processes in histogenesis? Equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *Journal Experimental Zoology*, 173(4):395–433.

[56] Stoy, K. and Nagpal, R. (2004). Self-reconfiguration using directed growth. In *7th International Symposium on Distributed Autonomous Robotic Systems (DARS 2004)*, pages 1–10.

[57] Subramanian, M. (2003). Fluid Mechanics - Lecture Notes, Department of Chemical Engineering, Sri Venkateswara College of Engineering, Sriperumbudur-602105 , Tamil Nadu, India. http://www.msubbu.com/ln/fm/Unit-IV/Drag.htm.

[58] Sugihara, K. and Suzuki, I. (1990). Distributed motion coordination of multiple mobile robots. In *5th IEEE International Symposium on Intelligent Control*, volume 1, pages 138–143.

[59] Sulsky, D., Childress, S., and Percus, J. K. (1984). A model of cell sorting. *Journal of Theoretical Biology*, 106:275–301.

[60] Sutton, J. J. (1998). *Biology*. Macmillan Press Ltd.

[61] Taylor, T. (2004). A genetic regulatory network-inspired real-time controller for a group of underwater robots. In *The 8th Conference on Intelligent Autonomous Systems (IAS-8)*, pages 403–412.

[62] Tomita, K., Murata, S., Kurokawa, H., Yoshida, E., and Kokaji, S. (1999). Self-assembly and self-repair method for a distributed mechanical system. *IEEE Transactions on Robotics and Automation*, 15(6):1035–1045.

[63] Ünsal, C. and Bay, J. S. (1994). Spatial self-organization in large populations of mobile robots. In *IEEE International Symposium on Intelligent Control*, pages 249–254.