Research Overview

Jenny Poulton

July 2017

1 Biological Background

DNA, RNA and proteins are biological molecules which carry information essential to how a living organism functions. All of these molecules are polymers; long chains of smaller chemical groups called monomers. The monomers come in multiple different types, and information is encoded into the exact ordering of these monomers (also called the sequence of the polymer). For DNA and RNA the monomers are called nucleotides and for proteins the monomers are called amino acids.

Given that the sequences of these polymers contain information on how the cells function, it is vital that, when these molecules are copied or created, the sequence is accurate. There are three common mechanisms for the copying or creation of these biological polymers as shown in figure 1: DNA is copied from itself by a process of replication, RNA is created, based on the sequence of DNA, by a process called transcription, and proteins are created, based on the sequence of RNA, by a process called translation. DNA replication is vital for cell growth: when a cell reproduces it must have accurate copies of the DNA from the parent cell in both daughter cells so that the new cells can produce the same sequences of RNA and DNA as other cells in the organism. Equally, it is vital that RNA is transcribed accurately from DNA and proteins are translated accurately from RNA so that the proteins will function correctly as they are used by the organism.[1]
It is important to note that, in all these cases, the copy must separate from the template. In DNA replication, the template and copy end up in different cells; when RNA is transcribed from DNA it is then moved from the nucleus to the cytoplasm. When protein is translated from RNA it is then used in many parts of the organism, inside and outside the cell it was produced in. Thus it is important that the accuracy of the sequence of any of these biological polymers is not dependent on there being lasting stabilising interactions between the copied polymer and the template. Any polymers created have to be persistent. This idea provides the basis for this project. [3]

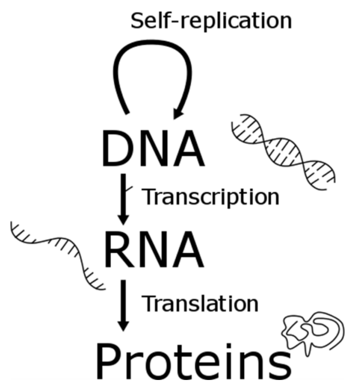

Figure 1: The flow of information in the cell through the processes of polymer copying and creation.

2 Project Aims

Given that creating accurate copies of biological polymers is central to the functioning of organisms, it is vital that we understand the processes of copying, and the generic reaction mechanism that allows copying.
The aim of this project is not to create a working model of any of these examples. These systems are immensely complex and it would be unfeasible to attempt to exactly solve the full processes they undergo. The aim of this project is to design the simplest possible model of an autonomous copying procedure which creates persistent copies. This model incorporates the key stages of copying: creation and separation. Once this model has been solved, we will then analyse models of progressively greater complexity.
By analysing these simplified models and understanding all the possible behaviours of the systems they describe, it is possible to make statements which provide constraints on the behaviour of any system in which a persistent copy is created. This could be in a biological context, or in a purely synthetic system such as in a computational context. In other words, these simple models highlight fundamental principles which will be reflected in more complex systems. Given this is the first model of its kind (no previous model has ever insisted on the persistence of its copies) this model may highlight totally new phenomena. This project aims to highlight the properties of a copying system which lead to increased accuracy, and how the requirement for accuracy places demands on cell resources. It also seeks to consider how these resources can be used most efficiently, and how the requirement for persistent copies changes these principles. These ideas, which are most easily understood from the simplest models, serve to frame our understanding of more complex systems.
It may be unclear to what extent ideas from these simple models can become central to the understanding of real biological systems. As an example, consider kinetic proofreading. Kinetic proofreading is the idea that the specificity of cer-

tain enzymatic processes can be increased by using chemical fuel to push the process out of equilibrium. This idea, which was initially demonstrated for a highly simplified model of enzymatic action, is now crucial to the understanding of much more complex systems and is widely referenced in experimental literature in the field.[2]
Additionally, accurate copying is vital to the concept of a living organism and indeed the simplest idea of life is a replicator. Even though life as we know it now is complex, with copying driven by complex enzymes, original copying mechanisms must have been simpler. Thus our simplified systems directly probe the properties of early life. As yet, no-one has constructed an enzyme-free synthetic copier driven purely by chemical processes. Being able to build a synthetic copier would be a major step towards engineering synthetic life-like systems. The principles highlighted by our simple models will guide the design of these synthetic copiers.

References

[1] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. Molecular biology of the cell (garland science, new york, 2002). *There is no corresponding record for this reference*, 1997.

[2] J. J. Hopfield. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proceedings of the National Academy of Sciences of the United States of America*, 71(10):4135–4139, Oct 1974. LR: 20130926; JID: 7505876; 0 (Amino Acids); 9014-25-9 (RNA, Transfer); OID: NLM: PMC434344; ppublish.

[3] Thomas E. Ouldridge and Pieter Rein ten Wolde. Fundamental costs in the production and destruction of persistent polymer copies. *Physical Review Letters*, 118(15):158103, 2017.