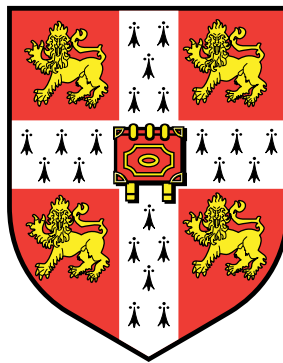


Mathematical and Experimental Approaches to the Dimer Catastrophe Theory

Christopher Martyn Field

Gonville & Caius College
Department of Genetics
University of Cambridge



A dissertation submitted for the degree of
Doctor of Philosophy
November 12, 2010

Title

Mathematical and Experimental Approaches to the Dimer Catastrophe Theory

Abstract

Multicopy plasmids rely on random distribution for stable inheritance by daughter cells at division. Threats to plasmid copy number increase the probability of plasmid loss, which can be detrimental to both plasmid and host. Plasmid dimers emerge through homologous recombination. Dimers have two independent origins of replication and thus have a replicative advantage and reduced copy number. Models of plasmid behaviour suggest that dimers would overtake a cell population, but that this can be prevented if they impose a small metabolic load, which has been observed *in vivo*. Plasmid ColE1 also contains a *cer* site, which allows for dimer resolution by XerCD site-specific recombination. A small RNA, Rcd, is expressed from the *cer* site in dimers and interacts with tryptophanase to increase the concentration of indole in the cell. It is proposed that, as indole inhibits cell division, Rcd imposes a checkpoint on the cell until plasmid dimers are resolved.

In this work, plasmid behaviour in a growing cell population was modelled stochastically in more detail than previous work. A plasmid replication model suggested that dimers replicate to more than half the average copy number of monomers, perhaps accounting for their increased metabolic load. A cell population model suggested that the presence of dimer-only cells decreased the average plasmid stability by less than in previous models, which used a fixed plasmid copy number. The rate of dimer resolution required to affect plasmid stability was unreasonably high, indicating the necessity of the Rcd checkpoint. The model thus suggested that the checkpoint may be an escape route for dimer-only cells rather than an immediate response to the emergence of an initial dimer.

The Rcd checkpoint itself was also subject to critical analysis. It was realised that neither inhibition of cell division nor cell growth were sufficient to assist dimer resolution; inhibition of plasmid replication was required. Experiments *in vivo* found that indole inhibited plasmid replication at a concentration that may be achievable endogenously. DNA gyrase was investigated as a component of the mechanism of this inhibition, and indole was found to inhibit its supercoiling activity *in vitro*.

Acknowledgements

My sincerest thanks go to David Summers, and although I'm sure that I've been difficult at times, I hope to have given something back in exchange for his encouragement and wisdom. I owe countless favours to Philip Oliver for all his assistance and the many occasions he has put up with my rambling on about mathematics or computing. I will even go so far as to thank John Archer for the teaching opportunities he inadvertently gave me.

I have thoroughly enjoyed my time working with my colleagues in Lab 102, and thank them for putting up with my particular ways. I am grateful for the provision of chocolate and conversation over the uncountably many tea breaks. The cast includes, in order of appearance: Duncan, Ian, Steve, Nic, Matt, Chih, Elaine, Claudio, Azlin, Sunkjukta, Antonio, James and Sylvia. Guest starring my long-suffering underlings: Cath, Alice, Anne-Laure and Jen.

I would like to thank my friends and family for their support; the last four years have probably been a bit of a mystery to them, but at last I have something to show for it.

A final thanks goes to a miscellany of helpful and supportive people. Marisa, for advice and taking an interest in my work. Jenny Barna, for her technical advice and patience. Irene, for drinks, dinners and delightful company. Ryanne, for her unwavering support and for keeping me sane.

Engineering taught me how to build things, science taught me how to take them apart.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. The length of this dissertation is no more than 60,000 words.

Table of Contents

List of Figures	17
List of Tables	19
Abbreviations	22
1 Introduction	23
1.1 Plasmid Biology	23
1.1.1 Classification	23
1.1.2 Structure	25
1.1.3 Components	26
1.2 Horizontal Gene Transfer	26
1.2.1 Transduction	27
1.2.2 Transformation	27
1.2.3 Conjugation	28
1.2.4 Bacteriocins	28
1.2.5 Bacterial Adaptability	29
1.3 Plasmid Replication	29
1.3.1 Principles of Replication Control	30

1.3.2	ColE1 Replication Control	31
1.3.3	ColE1 Derivatives	34
1.4	Plasmid Stability	34
1.4.1	Low-Copy Plasmid Stability	35
1.4.2	High-Copy Plasmid Stability	36
1.5	Stability Mechanisms of ColE1	38
1.5.1	Xer- <i>cer</i> Multimer Resolution	38
1.5.2	The Rcd Checkpoint Hypothesis	41
1.6	Type II Topoisomerases	44
1.6.1	DNA Gyrase	45
1.6.2	Topoisomerase IV	47
1.6.3	Topoisomerase Roles in Plasmid Replication and Control	48
1.7	Mathematical Models of ColE1 Replication and Control	48
1.7.1	Deterministic vs. Stochastic Systems	48
1.7.2	Deterministic Models	49
1.7.3	Stochastic Models	50
1.8	Mathematical Models of the Dimer Catastrophe	50
1.9	Aims and Objectives	51
2	Materials and Methods	53
2.1	Strains	53
2.2	Plasmids	54
2.3	Media	54
2.3.1	Lysogeny Broth (LB), with Agar (LA)	54
2.3.2	Super Optimal Broth with Catabolite Repression (SOC)	54
2.4	Antibiotics	55
2.5	Buffers and Solutions	55
2.5.1	ATP-Free Reaction Buffer for DNA Gyrase	55
2.5.2	Electrophoresis Buffer (TAE)	55
2.5.3	Ethidium Bromide	56
2.5.4	Indole and Indole Analogs	56

2.6	Microbiological Techniques	56
2.6.1	Cell Culture	56
2.6.2	Spectrophotometry	57
2.6.3	Centrifugation	57
2.6.4	Transformation	57
2.7	DNA Manipulation	57
2.7.1	Plasmid DNA Extraction	57
2.7.2	Gel Electrophoresis	58
2.7.3	Gel Extraction	58
2.7.4	Restriction Digests	58
2.8	Specific Assays	58
2.8.1	Indole-Plasmid Assay	58
2.8.2	Gel Densitometry	59
2.8.3	Indole-Gyrase Assay	59
2.9	Computer Modelling	59
2.9.1	Programming	59
2.9.2	Processing	60
2.9.3	Plasmid Stability Index	60
3	Modelling Plasmid Behaviour in a Cell Population	63
3.1	Introduction	63
3.2	Simulation Design	64
3.2.1	Chemostatic Environment	64
3.2.2	Data Structure	65
3.3	Simulation Operation	66
3.4	Testing	68
3.4.1	Variable Division Time	68
3.4.2	Runtime Optimisation	68
3.5	Results	70
3.5.1	General Behaviour	70
3.5.2	Dimer Formation Rate	70

3.5.3	Dimer Growth Penalty	73
3.5.4	Dimer Resolution Rate	73
3.5.5	Initial Conditions	76
3.6	Conclusions	76
4	Modelling ColE1 Copy Number Distribution	79
4.1	Introduction	79
4.2	Model Design	80
4.2.1	System State	80
4.2.2	Transitions	80
4.2.3	Algorithm	82
4.3	Parameters	83
4.3.1	Cell Parameters	83
4.3.2	Plasmid Parameters	84
4.3.3	Replication Parameters	84
4.4	Results	85
4.4.1	Average Copy Number Estimates	86
4.4.2	Plasmid Copy Number Distributions	86
4.4.3	Dimer Formation and Resolution	88
4.5	Conclusions	90
5	An Improved Model of Plasmid Behaviour in a Cell Population	93
5.1	Introduction	93
5.2	Simulation Design and Operation	94
5.2.1	Copy Number	94
5.2.2	Metabolic Load	94
5.2.3	Dimer Formation Rate	95
5.2.4	Initial Conditions	96
5.2.5	Simulation and Sampling Time	96
5.2.6	Parallelisation	96
5.3	Basic operation	96

5.3.1	Average Steady-State Behaviour	97
5.3.2	Plasmid Distribution	97
5.4	Dimer Formation Rate	98
5.5	Division Time Penalty	105
5.6	Dimer Resolution	105
5.7	Conclusions	108
6	The Effect of Indole on Plasmid Replication	109
6.1	Introduction	109
6.2	Examination of the Rcd Checkpoint Hypothesis	110
6.2.1	Inhibition of Cell Division is Insufficient to Ensure Plasmid Stability	110
6.2.2	Inhibition of Cell Growth is Insufficient to Prevent Plasmid Replication	111
6.3	The Effect of Indole on Plasmid Replication	112
6.3.1	Plasmid Replication in Chloramphenicol-Treated Cells	112
6.3.2	Indole Prevents Plasmid Amplification	113
6.3.3	Indole Alone Inhibits Plasmid Replication	113
6.4	Concentration Dependence of Plasmid Replication Inhibition by Indole	116
6.5	Replication Inhibition for Monomers and Dimers	117
6.6	The Effects of Indole Analogs on Plasmid Replication	117
6.7	Conclusions	120
6.7.1	Indole Has Multiple Targets	123
6.7.2	A Modified Rcd Checkpoint Hypothesis	124
7	The Mechanism of Plasmid Replication Inhibition by Indole	125
7.1	Introduction	125
7.2	Indole Inhibits DNA Gyrase <i>in vitro</i>	126
7.3	Indole Does Not React with ATP or the DNA Substrate	127
7.4	Concentration Dependence of DNA Gyrase Inhibition	129
7.5	The Inhibition of DNA Gyrase by Indole is Reversible	129
7.5.1	Pre-Incubation of Indole with DNA Gyrase	129
7.5.2	Recovery of Supercoiling Activity After Indole Treatment	131

7.6	Indole Does Not Cause Double-Strand Breaks	131
7.7	The Kinetics of DNA Gyrase Inhibition by Indole	133
7.8	Conclusions	134
7.8.1	Indole Inhibits DNA Gyrase	134
7.8.2	Indole Probably Targets GyrB	135
8	Discussion	137
8.1	Computer Modelling	137
8.1.1	The Role of the Rcd Checkpoint	137
8.1.2	Population-Wide Effects	138
8.1.3	Future Work	139
8.2	Indole and the Rcd Checkpoint	140
8.2.1	The Modified Rcd Checkpoint Hypothesis	140
8.2.2	Indole Production and Concentration	141
8.2.3	Future Work	141
	References	161

List of Figures

1.1	Incompatibility between two plasmids	24
1.2	The kinetics of plasmid replication	30
1.3	Some theoretical plasmid copy number control systems	32
1.4	A representation of ColE1 replication control	33
1.5	The dimer catastrophe	37
1.6	Sequence of the <i>cer</i> site	39
1.7	The Hodgman Model of synaptic complex formation at <i>cer</i>	41
1.8	The activity of the enzyme tryptophanase	43
1.9	The activity of DNA gyrase	46
3.1	Comparison of simulation behaviour with different sampling frequencies . .	69
3.2	Simulation behaviour with distributed interdivision times	69
3.3	Comparison of simulation behaviour with different population sizes	71
3.4	Distribution of plasmid dimers in the population at different time points . .	72
3.5	Simulation behaviour with different dimer formation rates	74
3.6	Plasmid stability with different dimer formation rates	74
3.7	Simulation behaviour with different division time penalties	75
3.8	Plasmid stability with different division time penalties	75

3.9	Simulation behaviour with different dimer resolution probabilities	77
3.10	Plasmid stability with different dimer resolution probabilities	77
4.1	State transitions and their rates in the plasmid replication model	81
4.2	Probability distribution of plasmid copy number	89
4.3	Probability distribution of plasmid copy number (logarithmic scale)	89
4.4	Plasmid copy number over time with dimer formation and resolution	91
5.1	Simulation behaviour with copy number variation	99
5.2	Plasmid stability with copy number variation	99
5.3	Plasmid distribution with copy number variation	100
5.4	Monomer distribution with copy number variation	101
5.5	Dimer distribution with copy number variation	102
5.6	Monomer-dimer distribution with copy number variation	103
5.7	Simulation behaviour with different dimer formation rates	104
5.8	Plasmid stability with different dimer formation rates	104
5.9	Simulation behaviour with different division time penalties	106
5.10	Plasmid stability with different division time penalties	106
5.11	Simulation behaviour with different dimer resolution rates	107
5.12	Plasmid stability with different dimer resolution rates	107
6.1	Comparison of delayed and normal cell division	110
6.2	OD ₆₀₀ of cultures treated with combinations of Cm, indole and NA	114
6.3	Gel densitometry data for combinations of Cm, indole and NA	114
6.4	Agarose gel images for combinations of Cm, indole and NA	115
6.5	OD ₆₀₀ of cultures treated with Cm and different indole concentrations	118
6.6	Gel densitometry data for different indole concentrations	118
6.7	Agarose gel images for different indole concentrations	119
6.8	Agarose gel image for different indole analogs	122
6.9	Gel densitometry data for different indole analogs	122
7.1	Chemical structures of indole, indolinone and compound 1	126

7.2	Supercoiling assay to demonstrate DNA gyrase inhibition by indole	128
7.3	Supercoiling assay with pre-incubation step	130
7.4	Concentration dependence of DNA gyrase inhibition by indole	130
7.5	Supercoiling assay to demonstrate reversible inhibition	132
7.6	Recovery of DNA gyrase supercoiling activity	132
7.7	Supercoiling assay with different ATP concentrations	134

List of Tables

2.1	Strains of <i>Escherichia coli</i> used during this work.	53
2.2	Plasmids used during this work.	54
2.3	Antibiotics used in this work.	55
3.1	Performance of different priority queue implementations	66
3.2	Population model parameters	67
4.1	Replication model parameters	83
4.2	Steady-state average plasmid copy numbers	87
4.3	Steady-state average inhibitor levels	87
4.4	Plasmid distribution statistics	87
5.1	Combined model parameters	97
6.1	Chemical analogs of indole	121

Abbreviations

ATP	Adenosine triphosphate
BSA	Bovine serum albumin
Cm	Chloramphenicol
CTP	Cytidine triphosphate
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
GFP	Green fluorescent protein
NA	Nalidixic acid
OD₆₀₀	Optical density measured at 600 nm
PCR	Polymerase chain reaction
RAM	Random access memory
Rcd	Regulator of cell division
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
SDW	Sterilised distilled water
TAE	Tris-acetate-EDTA
tRNA	Transfer RNA
UV	Ultraviolet light
YFP	Yellow fluorescent protein

Chapter

1

Introduction

1.1 Plasmid Biology

The bacterial plasmid was first identified as an ‘infective hereditary factor’ ([Lederberg et al., 1952](#)) that, unlike bacteriophage, was non-lytic ([Hayes, 1953](#)). The name was proposed as ‘a generic term for any extrachromosomal hereditary determinant’ ([Lederberg, 1952](#)). At first, plasmids were only of interest to those investigating intercellular gene transfer, but by the 1960s it was determined that they were responsible for the spread of antibiotic resistance in pathogens ([Watanabe, 1963](#)). Today, they are ubiquitous tools for genetic engineering and still of major concern with regards to multiple antibiotic resistance.

1.1.1 Classification

The most useful definition of a plasmid is as a non-essential extrachromosomal element. This implies that it ought to be possible to ‘cure’ the host of the plasmid, but there remains ambiguity for some megaplasms. For instance, one of the two megaplasms of *Rhizobium meliloti* carries essential genes such as the arginine tRNA ([Galibert et al., 2001](#)) and both are capable of cointegration with the chromosome ([Guo et al., 2003](#)). It was recently proposed that such elements be distinguished from both plasmids and chromosomes as ‘chromids’ ([Harrison et al., 2010](#)).

Plasmids are often determined to be either ‘low-copy’ or ‘high-copy’, referring to their average copy number in a particular host. There is no defined cut-off between the two, though low-copy plasmids are typically larger and actively partitioned at cell division (see [Section 1.4.1](#)), whereas high-copy plasmids are smaller and randomly distributed.

Plasmids have been classified historically according to a range of criteria, such as conjugative ability, drug-resistance and colicin production (Meynell et al., 1968). The predominant modern scheme is by incompatibility group; when two plasmids are unable to coexist in the same cell, they are said to be incompatible. This is typically tested by introducing a second plasmid to cells that already contain a first, and looking for loss of one or the other during subsequent culturing. Couturier et al. (1988) offer a good review of the 30 or so incompatibility groups currently in use. Wang et al. (2009) recently proposed a more rigorous classification scheme for plasmid vectors used in biotechnology based on the plasmid replicon, its resistance markers and the promoter used for gene expression.

The molecular basis for incompatibility lies in the replication control systems employed by different plasmids. Most produce a *trans*-acting repressor of replication, the concentration of which is proportional to the number of plasmids in the cell. Two plasmids that share the same control mechanism will co-repress one another's replication such that the total copy number achievable in the cell is shared by the two of them. Over time, cells with only one of the two plasmids will emerge; the replication control system does not distinguish between the two and, should one replicate more often than the other in a particular cell, positive feedback reinforces the more dominant plasmid (see Figure 1.1).

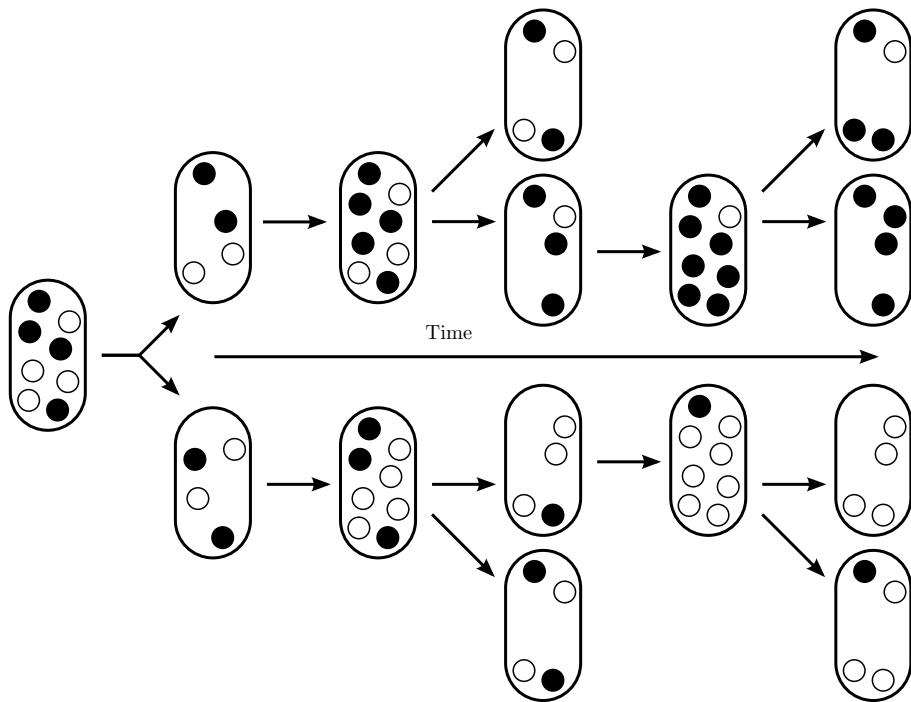


Figure 1.1: Incompatibility between two plasmids, shown as black and white circles. Starting at equal copy number in the same cell, distribution at cell division results in a bias towards one or other plasmid. This bias is amplified by subsequent replication and division cycles until cells emerge with only one of the two plasmids.

Incompatibility can also present itself as a result of shared or similar partitioning systems in low-copy plasmids (Nordström et al., 1980). If two such plasmids are present in the same cell, they are still paired successfully at cell division, but not necessarily with one of their own kind. Thus a daughter cell can inherit more of one plasmid than another, and whether the two share the same replication control system or not, will ultimately segregate by mismatched partitioning.

The results of an incompatibility test can sometimes be difficult to interpret. Certain plasmids have more than one system responsible for their replication control. If such a plasmid were to share a replicon with another plasmid in the same cell, it could be maintained by its second replicon, which may be less or not at all repressed by the single-replicon plasmid. The single-replicon plasmid would, however, still have its copy number suppressed by the dual-replicon plasmid. Certain stability mechanisms, such as host-killing, can also obfuscate results by eliminating cells which have lost a particular plasmid (see Section 1.4.1). The rate at which one plasmid displaces another will depend on their initial frequencies, their sensitivity to the repressor and the metabolic burden they impose on the host. Cullum and Broda (1979) made an early attempt to model the process and calculate segregation rates.

With the advent of high-throughput DNA sequencing techniques, plasmid replicons can be more rapidly identified by PCR (Carattoli et al., 2005). Further, entire sequences can be aligned to construct phylogenies of plasmid families (Jensen et al., 2010). This has shifted the focus of classification away from incompatibility typing to a more considered view of the relationships between different plasmids.

1.1.2 Structure

Plasmids were only identified as being comprised of DNA ten years after their initial discovery (Marmur et al., 1961). Most, and by far the majority in *E. coli*, are negatively supercoiled circles of double-stranded DNA. Supercoiling influences a variety of processes in the cell, such as transcription and replication, and is maintained by the action of DNA gyrase (see Section 1.6.1). The circularity of most plasmids makes them easy to extract in preference to the chromosome, by heat (Holmes and Quigley, 1981) or alkaline lysis (Birnboim and Doly, 1979). Modern preparative techniques use a combination of alkaline lysis and the absorption of DNA by silica (Vogelstein and Gillespie, 1979), optimised for plasmid DNA.

Linear plasmids can also be found in a variety of bacterial species (reviewed in Hinnebusch and Tilly, 1993; Meinhardt et al., 1997). Structurally, linear plasmids have either covalently closed ends in the form of hairpin loops, such as those of *Borrelia burgdorferi* (Hinnebusch and Barbour, 1991), or proteins bound to their 5' termini, such as pSLA2 of *Streptomyces rochei* (Hirochika et al., 1984).

Plasmid size varies considerably, from around 2 kbp for a small plasmid such as pHD2 of *Bacillus thuringiensis* (McDowell and Mann, 1991), up to 1.35 Mbp for the aforementioned megaplasmid of *Rhizobium meliloti* (Galibert et al., 2001). The metabolic load imposed by an individual plasmid is increased if it is made larger (Cheah et al., 1987). Increasing the size of a plasmid, for instance by inserting a cloned gene, will typically reduce its copy number to limit the total metabolic load in the host.

1.1.3 Components

The basic replicon is the backbone of the plasmid, defined as the smallest length of DNA able to replicate with wild-type copy number (Nordström, 1985). It consists of an origin of replication, at which the process is initiated, as well as elements to control this initiation. Control is exerted either by a series of genes that encode for proteins to initiate and repress replication, as for plasmid R1, or by a region from which different RNAs are transcribed to achieve the same, as for plasmid ColE1. The details of replication control are discussed in Section 1.3.

Plasmid-selfish functions, including replication, enhance stability and horizontal transfer. Active partitioning systems in low-copy plasmids, host-killing systems, multimer resolution sites and their associated recombinases improve segregational stability (see Section 1.4). Conjugative ability and mobilisation allow for horizontal transfer and the opportunity to colonise new hosts (see Section 1.2).

As well functions to ensure replication and stable maintenance, a plasmid can encode a wide range of characteristics. One can consider all mobile genetic elements to be composed of different modules, each providing a specific function (Toussaint and Merlin, 2002). From the point of view of the plasmid, individual components can appear to be selfish, ensuring its continued existence, or mutualistic, providing advantage to the host. Some components, such as colicin production, can be considered both selfish and mutualistic, dependent on circumstance (see Section 1.2.4).

Mutualistic functions confer novel phenotypes on a host and enhance its fitness. The most infamous example is antibiotic resistance, but resistance to other toxins, such as metal ions, is also common (reviewed in Foster, 1983). Genes can also encode metabolic functions, virulence factors, toxin production and more (see Stanisich, 1988, for a comprehensive list).

1.2 Horizontal Gene Transfer

In the absence of sexual reproduction, bacterial adaptation is driven by horizontal gene transfer. It is estimated that 16% of the *E. coli* chromosome arose through horizontal

transfer (Lawrence and Ochman, 1997; Médigue et al., 1991), and along with bacteriophage and transposons, plasmids play a significant role in the development of the bacterial genome. It has even been demonstrated that bacterial plasmids can cross the domain boundary and transfer to plants and fungi (Buchanan-Wollaston et al., 1987; Heinemann and Sprague Jr, 1989), implying that they may influence evolution in eukaryotic organisms. The ability to colonise new hosts by horizontal transmission allows for greater propagation of plasmids than through inheritance alone.

Horizontal transfer is achieved by three mechanisms: transduction, transformation and conjugation.

1.2.1 Transduction

Transduction is the transfer of genetic material from one cell to another by bacteriophage (Zinder and Lederberg, 1952). During phage development, transducing particles are formed, in which chromosomal or plasmid DNA has been accidentally packaged. The size of the DNA packaged is critical; it cannot be too large or too small for the phage head (Saye et al., 1987). The original host is lysed and a new host infected, whereupon the packaged DNA is potentially integrated into the new host chromosome. This is not always successful, and transferred DNA will only persist for a short while before degradation. Plasmids have an advantage at this stage; provided that they can recircularise and are compatible with the host's replication system, they will autonomously replicate and therefore persist in the new host.

1.2.2 Transformation

Transformation is the take-up of naked DNA from the environment and cells that are capable of doing so are 'competent'. Some bacteria, such as certain species of *Neisseria*, are perpetually competent (Catlin and Cunningham, 1961) and some, such as *Bacillus subtilis*, regulate their competence according to physiological conditions (Dubnau, 1991). The effective uptake of DNA can require the presence of specific sequences, for example in *Neisseria gonorrhoeae* (Goodman and Scocca, 1988). In the laboratory, most bacterial strains can be transformed either by chemical treatment or electroporation (Sambrook and Russell, 2001).

The mechanisms underlying natural competence and transformation will not be discussed in detail here; Johnsborg et al. (2007) and Chen and Dubnau (2004) offer reviews of the subject. Suffice to say that plasmids, as they are generally closed circles, are likely to survive longer in the environment than other DNA, and are therefore primary candidates for natural transformation.

1.2.3 Conjugation

Horizontal transfer of plasmid DNA is particularly associated with conjugation, a parasexual process that involves direct transfer of DNA from a host to a recipient during cell-cell contact (Lederberg and Tatum, 1946). In Gram-negative bacteria, pili are essential (Achtman et al., 1978), as they are thought to initiate contact and pull the surface of the cells together to allow a transport pore to form. Two types of pili have been identified: long flexible pili approximately 1 μm long and short rigid pili approximately 0.1 μm long (Bradley, 1980). Some plasmids make use of both types (Bradley, 1984). Gram-positive bacteria do not have pili, and the mechanism by which they come together is unclear (Grohmann et al., 2003).

Once the cells are brought together, conjugation proceeds with nicking at *oriT* and unwinding of the plasmid before a single strand is passed to the recipient cell. Synthesis of replacement and complementary strands then occurs in the host and recipient cells respectively, before recircularisation of the recipient plasmid. The mechanism is reviewed in detail in Willetts and Wilkins (1984) and Lanka and Wilkins (1995).

Some plasmids are mobilisable; they can be transferred by conjugation, but lack the machinery required to bring cells together for transport and control DNA processing during the process. These functions must be provided by a co-resident conjugative plasmid. Taking ColE1 as an example, mobilisation requires an *oriT* on the plasmid (Bastia, 1978), the *mob* region, containing genes *mbeABCD*, (Boyd et al., 1989) and the *tra* gene products of a conjugative plasmid (Willetts and Wilkins, 1984). ColE1 changes conformation to an open circular form in association with proteins at *oriT* (Clewell and Helinski, 1969). Varsaki et al. (2003, 2009) recently characterised MbeC as a DNA-binding protein that bends the *nic* site to assist the relaxase, MbeA.

1.2.4 Bacteriocins

The production of, and immunity to, bacteriocins allows a host to antagonise other strains it may be competing with. The phenotype is widespread among plasmids from bacteria isolated in the pre-antibiotic era (Hughes and Datta, 1983). These systems can be encoded by the host or by a plasmid and the bacteriocin toxin can be a simple protein or complex of proteins. A well-studied example is the system of plasmid ColE1 (reviewed with other examples in Konisky, 1982).

ColE1 encodes a colicin, *cea*, a gene responsible for cell lysis, *kil* and an immunity gene *imm* (Sabik et al., 1983). Transcription of *cea* and *kil* is repressed by LexA, but this protein is disabled by RecA activity when the SOS response is triggered by DNA damage. Thus, a cell in such a situation synthesises a burst of colicin and is lysed by *kil*, releasing the colicin into the environment. Leakage from the operon also produces a low

background level of colicin. Those cells with *imm* are immune to the exogenous colicin, which otherwise depolarises the cell membrane and severely disrupts metabolism (Gould and Cramer, 1977).

The ecological consequences of colicin production are complex. Colicin will kill any bacterial invaders that haven't yet obtained ColE1, any cells native to an environment that ColE1 has just invaded and any cells that may have lost the plasmid through segregational instability.

1.2.5 Bacterial Adaptability

The evolutionary advantage of plasmids is derived from their modularity, their provision of a wide range of functions for selective advantage and their ability to horizontally transfer to new hosts (Reanney, 1976). They act almost as books in a library of genetic information, with their hosts borrowing adaptations as selection demands. Eberhard (1990) describes the advantage of a plasmid-borne gene over a chromosomal gene in his local adaptation hypothesis. Many of the plasmid-encoded phenotypes are only transiently useful to a host, in a particular time or place. Provided that a plasmid's rate of horizontal transfer is greater than the reduction in its rate of vertical transmission due to the metabolic load it imposes on the host, it will propagate more rapidly than any chromosomal gene. Further, the plasmid's ability to find new hosts allows it to associate with potentially superior genes of other plasmids and chromosomes. Different phenotypes will associate on the same plasmid, and if they are useful in the same conditions, positive feedback will keep them together and enhance the advantage the plasmid provides. For the host, the ability to lose genes that are only locally useful, but otherwise penalising, gives considerable adaptability.

1.3 Plasmid Replication

By definition, the most essential plasmid function is persistence. Once in a host, a plasmid must replicate sufficiently so that at cell division, at least one copy is passed to each daughter cell. Failure results in plasmid loss, which is detrimental for the plasmid, denied the chance to replicate in that cell line, and sometimes to the host, denied a selective advantage or deliberately harmed by the plasmid it lost.

Plasmids replicate autonomously, but must borrow components of the host's replication machinery to do so. For both low-copy and high-copy plasmids there is a dilemma: replicate too often, and the host will be overly burdened and at a selective disadvantage, but fail to replicate enough and the plasmid will be lost from the population. A plasmid must synchronise its replication with the growth rate of the host, and so each encodes elements to control its own copy number.

1.3.1 Principles of Replication Control

As a basic principle of copy number control, a steady concentration of plasmids should be maintained in a growing cell and a consistent copy number achieved in time for distribution at cell division. Any deviations from this behaviour ought to be corrected as efficiently as possible. A single plasmid, introduced to a cell by transformation or conjugation, is indeed capable of rapidly establishing itself in the subsequent population at average copy number (Highlander and Novick, 1987).

A replication control system must ensure that, for a plasmid above the mean copy number, replication is reduced to less than once per plasmid per generation, and for a plasmid below the mean copy number, replication is increased to more than once per plasmid per generation. This presents a number of plausible kinetics, summarised in Figure 1.2. Curve (a) is a step function, wherein the plasmid will replicate as fast as possible when under-represented and not at all when at or over the mean, similarly to the plasmid form of bacteriophage P1 (described in Das et al., 2005). Curve (b) is a hyperbolic function, wherein the plasmid replicates at a rate inversely proportional to its copy number, similarly to plasmid R1 (described in Nordström and Wagner, 1994). Finally, curve (c) is an exponential function, a compromise between the other two kinetics, illustrated by the focus of this work, plasmid ColE1 (described in Section 1.3.2).

Whatever the precise form of the kinetics, stable control requires negative feedback; any other system would amplify small perturbations from the stable state and run away. The feedback signal in this case is the concentration of a plasmid-encoded replication repressor, and two theoretical models have been proposed to explain how it can exert control.

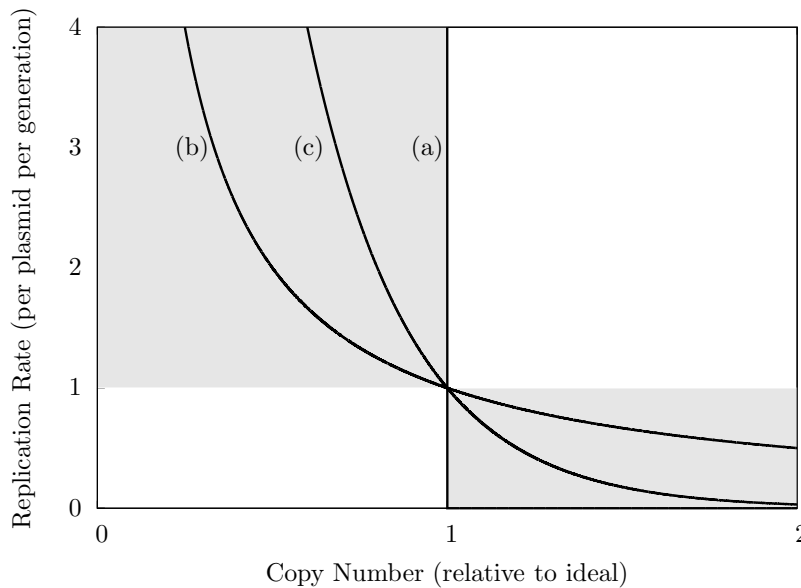


Figure 1.2: The kinetics of plasmid replication (from Nordström et al., 1984). Only behaviour in the shaded area will result in stable maintenance of mean copy number. Curve (a) is a step function, similar to P1, curve (b) is hyperbolic, similar to R1 and curve (c) is exponential, similar to ColE1.

In the autorepressor model (Sompayrac and Maaloe, 1973), replication is initiated by a *rep* protein on the same operon as the repressor. Expression of the operon is inhibited by the repressor such that *rep* concentration is maintained and replication will occur at a constant rate. A possible genetic organisation for such a system is shown in Figure 1.3(a). The model also allows for a single protein that is both the autorepressor and initiator, which is similar to the structure of the control system of plasmid P1, and is shown in Figure 1.3(b).

In the inhibitor dilution model (Pritchard et al., 1969), production of the inhibitor and initiator are controlled separately. Ideally, the inhibitor maintains a concentration proportional to the plasmid copy number and either represses replication directly or inhibits expression of the initiator, shown in Figures 1.3(c) and (d). As the cell grows, the inhibitor is diluted, allowing replication of the plasmid, which in turn increases the rate of inhibitor production. This allows the plasmid to replicate at a rate proportional to the growth rate of the cell. If the inhibitor concentration does not achieve proportionality with the plasmid copy number fast enough, the plasmid will continue to replicate and overshoot the intended copy number. In control terms, the system has become unstable due to too great a phase shift in the feedback signal. This is prevented by rapid production and degradation of the inhibitor.

1.3.2 ColE1 Replication Control

Here, the replication and copy number control of ColE1 is discussed in detail, but for a review of other mechanisms see Solar et al. (1998). Replication of ColE1 does not require the expression of plasmid-encoded proteins (Donoghue and Sharp, 1978), unlike many other plasmids. Instead, an RNA pre-primer, RNAII, is transcribed from a constitutive promoter 555 bp upstream of the origin of replication (Itoh and Tomizawa, 1980). RNAII naturally folds into a complex secondary structure with multiple stem-loops (Masukata and Tomizawa, 1986). This allows a DNA-RNAII complex to form at the origin of replication, the RNA strand of which is cleaved by RNaseH to produce a primer for leading-strand synthesis by DNA Polymerase I (Itoh and Tomizawa, 1980). After initiation, DNA Polymerase III takes over replication (Staudenbauer, 1976). Meanwhile, lagging strand synthesis is initiated at a primosome binding site downstream of the origin, revealed by unwinding of the DNA (Zavitz and Marians, 1991). Due to the relative stability of the proteins required, replication of ColE1 is able to continue in the absence of *de novo* protein synthesis, for instance in the presence of chloramphenicol (Clewett, 1972).

Control of ColE1 replication is exerted by the activity of a second RNA, RNAI, transcribed from a constitutive promoter 445 bp upstream of the origin of replication, in the opposite direction to RNAII (Tomizawa et al., 1981). It is thus complementary to RNAII. RNAI also forms a structure with multiple stem loops and binds RNAII to inhibit primer formation (Tomizawa, 1984). The binding process begins with interaction between the

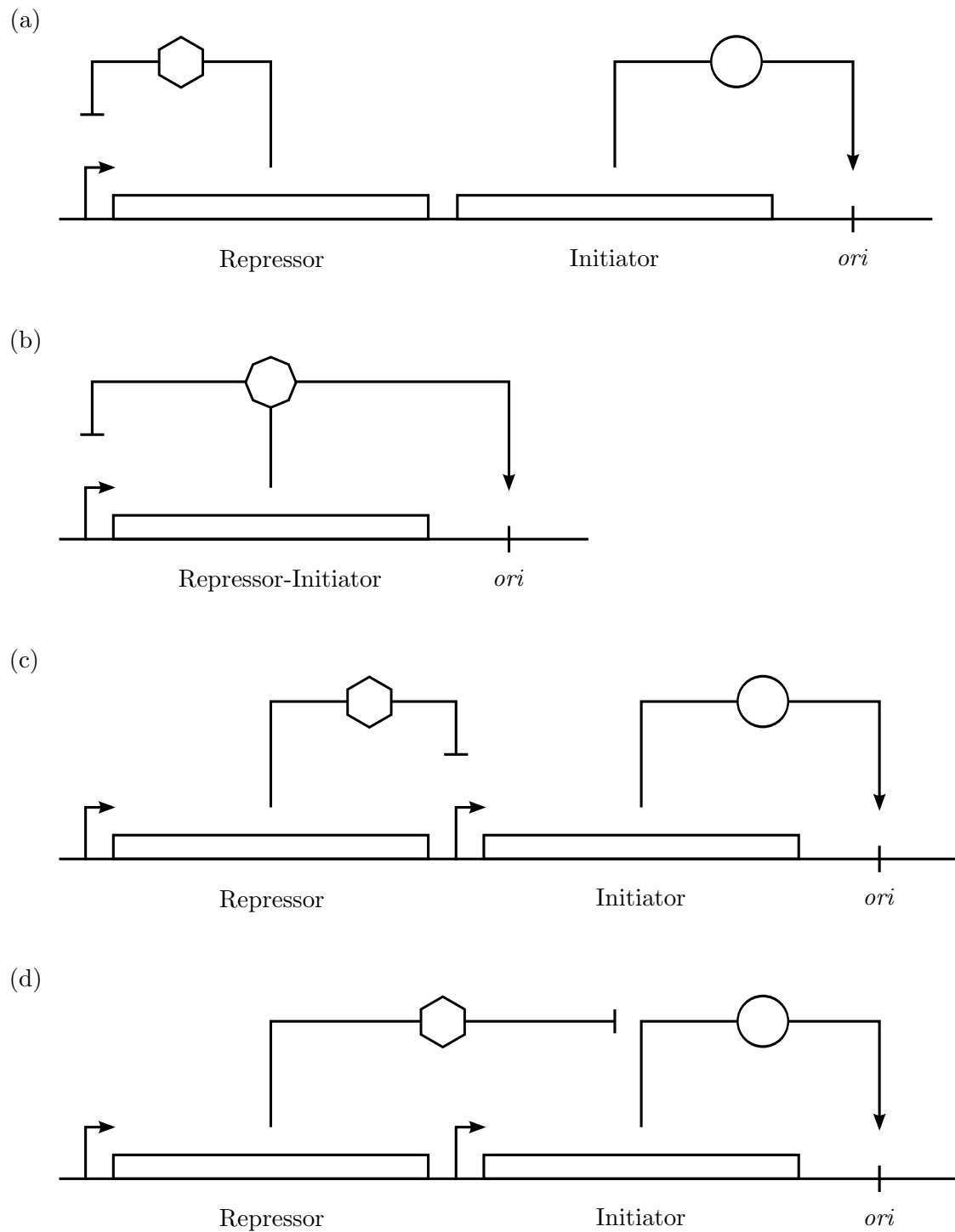


Figure 1.3: Some theoretical plasmid copy number control systems; (a) an autorepressor and initiator on the same operon, (b) a single gene product acts as both autorepressor and initiator, (c) a repressor inhibits transcription of a separately expressed initiator, (d) a repressor directly inhibits a separately expressed initiator.

complementary loops of the two RNAs in a so-called ‘kissing complex’, which encourages base pairing from the 5’ end of RNAI. This in turn unravels the stem loops as pairing propagates (Tomizawa, 1990). A simple representation of ColE1 replication control is shown in Figure 1.4.

RNAI binding has the effect of altering the secondary structure of RNAII so as to inhibit the formation of the DNA-RNAII complex (Masukata and Tomizawa, 1986). However, RNAI inhibition is only effective in a certain window of opportunity during the transcription of RNAII. It cannot bind at all before RNAII has reached a length of 100 nucleotides, due to allosteric inhibition by RNA polymerase. After a length of 360 nucleotides is reached, although RNAI can still bind, there is no effect on primer formation (Tomizawa, 1986). Transcription proceeds at approximately 50 nucleotides per second (Hippel et al., 1984) so the window for inhibition is open for only 5 seconds. It is estimated that only 1 in 20 RNAII transcripts is able to form a primer (Lin-Chao and Bremer, 1987) and of those, only half ever successfully complete replication (Brendel and Perelson, 1993).

ColE1 also features a secondary control system in the form of the plasmid-encoded protein Rom (Twigg and Sherratt, 1980). Rom binds to the stem-loop structures formed by RNAI and RNAII, protecting the complex they form from cleavage by RNAaseV₁ and from alkylation of their phosphate groups (Eguchi and Tomizawa, 1990). This reduces the dissociation rate of the complex about 300-fold. ColE1 derivative plasmids lacking Rom have a copy number several times higher than usual (Som and Tomizawa, 1983; Twigg and Sherratt, 1980).

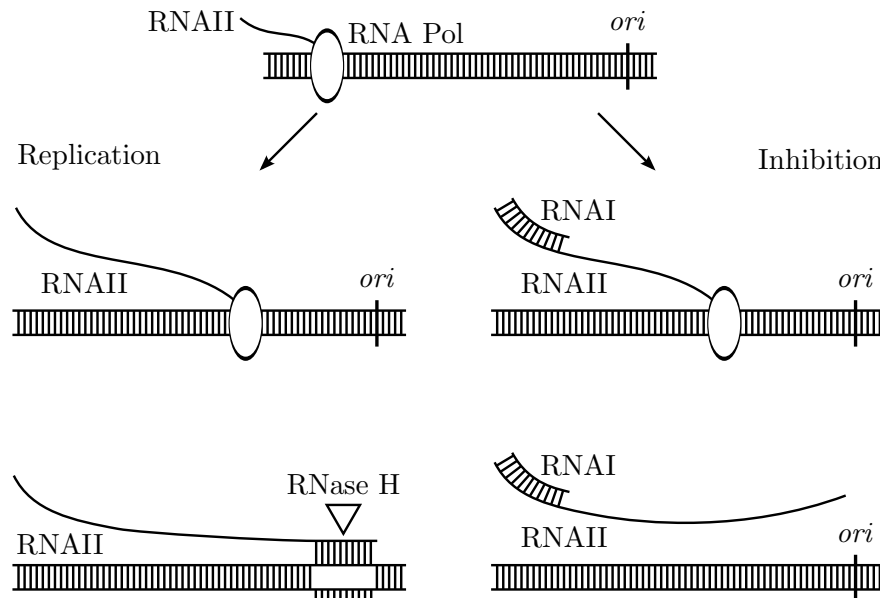


Figure 1.4: A representation of ColE1 replication control. On the left, replication is successfully primed by the cleavage of unimpeded RNAII transcription. On the right, RNAI forms a complex with RNAII to disrupt its secondary structure, preventing replication priming.

The role of Rom in replication control has not been completely resolved. [Ehrenberg \(1996\)](#) suggests that as it determines the effectiveness of inhibition in conjunction with RNAI, and its own concentration is proportional to plasmid copy number, it allows for sharper response to changes in plasmid concentration. [Paulsson et al. \(1998\)](#) suggest Rom may exist to bring the probability of replication at high copy number closer to zero and thus limit the dynamic range of the control system. [Summers \(1996\)](#) suggests that the absence of Rom in a post-conjugal recipient allows for rapid establishment of the plasmid, before the control system returns to normal as the Rom level increases. Finally, [Atlung et al. \(1999\)](#) suggest that Rom exists to prevent over-replication in slow growing cells and thus reduce the metabolic burden imposed by the plasmid.

1.3.3 ColE1 Derivatives

A number of plasmids have been derived from ColE1 for use as cloning vectors in molecular biology, some with modifications to the replication control elements to increase copy number. The majority use the origin of a ColE1-like plasmid, pMB1 ([Betlach et al., 1976](#)). Of these, the cloning vector pBR322 has control elements and copy number closest to ColE1 ([Bolivar et al., 1977](#)). The pUC series of cloning vectors have a point mutation in RNAII that significantly increases copy number ([Lin-Chao et al., 1992](#)) and the lineage is Rom⁻ ([Vieira and Messing, 1982](#)).

1.4 Plasmid Stability

For a plasmid to persist in a host cell line, it must be both structurally and segregationally stable. For structural stability, a plasmid must maintain a consistent nucleotide sequence. Short, direct or inverted repeat sequences can generate deletions ([Schaaper et al., 1986](#)) and topoisomerase activity can cause rearrangement ([Ikeda et al., 1981](#)). Further, a plasmid entering a new host may be subject to restriction and modification by host-encoded endonucleases (reviewed in [Boyer, 1971](#)).

This work focuses on segregational stability, which requires that a plasmid to be passed on to both daughters at cell division. It is usually determined experimentally by allowing a bacterial culture to grow without selection, after an initial period of selection for the plasmid. Cells are tested for the presence of the plasmid after a given time in a variety of ways (see [Livermore and Brown, 2001](#), for β -lactamase resistance). Two factors determine the apparent rate of loss of the plasmid: the rate at which plasmid-free cells arise and the rate with which such cells accumulate due to their metabolic advantage over plasmid-bearing cells ([Boe et al., 1987](#)). Metabolic load is dependent upon plasmid size, copy number and whether or not it expresses costly gene products ([Cheah et al., 1987](#); [Seo and Bailey, 1985](#)). High-copy and low-copy number plasmids have adopted different strategies

to minimise the rate at which plasmid-free cells arise; those of low-copy number plasmids will be discussed only briefly here.

1.4.1 Low-Copy Plasmid Stability

Active Partitioning

Low-copy plasmids contain partitioning, or *par*, loci that function to actively segregate them at cell division (reviewed in [Salje, 2010](#)). These systems show some similarity to those responsible for chromosome segregation in eukaryotes ([Gerdes et al., 2000](#)). All known examples of these systems consist of two *trans*-acting proteins and at least one *cis*-acting centromere-like site, at which the proteins act ([Abeles et al., 1985](#); [Gerdes and Molin, 1986](#); [Ogura and Hiraga, 1983](#)). The first protein binds DNA at the centromere-like site, pairing plasmids at the mid-cell position ([Davis and Austin, 1988](#); [Edgar et al., 2001](#); [Jensen et al., 1998](#); [Mori et al., 1989](#)). The second is an ATPase that forms the filaments necessary to separate the plasmid pairs ([Barillà et al., 2005](#); [Ebersbach and Gerdes, 2004](#); [Møller Jensen et al., 2002](#)). The expression of these proteins is tightly auto-regulated ([Friedman and Austin, 1988](#); [Hirano et al., 1998](#); [Jensen et al., 1994](#)). [Ebersbach and Gerdes \(2005\)](#) and [Gerdes et al. \(2010\)](#) offer reviews of these mechanisms in detail.

Host-Killing Mechanisms

Low-copy plasmids sometimes encode a stability mechanism that acts to eliminate plasmid-free cells. These host-killing systems typically involve a stable toxic protein, such as CcdA in F or Hok in R1, and an unstable antitoxin, which may be a protein, such as CcdB in F, or an antisense RNA that exerts transcriptional control, such as *sok* in R1 ([Gerdes et al., 1988](#); [Tam and Kline, 1989](#)). A cell that inherits no plasmids after cell division will rapidly lose immunity and suffer as a result. [Hayes \(2003\)](#) offers a review of these mechanisms in detail, whilst [Van Melder and Saavedra De Bast \(2009\)](#) postulate on the roles of both plasmid-borne and host-encoded systems.

Multimer Resolution

As multiple copies of a plasmid exist at once inside the cell, plasmid dimers and higher-order multimers can emerge through homologous recombination (reviewed in [Smith, 1988](#)). For low-copy plasmids, this can result in their partitioning system ineffectively trying to push apart two sites on the same molecule ([Austin et al., 1981](#)). To combat this, some encode site-specific recombination systems capable of resolving dimers back into monomers, such as the Cre-*lox* system of P1 (reviewed in [Van Duyne, 2001](#)).

1.4.2 High-Copy Plasmid Stability

If multicopy plasmids are randomly distributed throughout the cell, the probability of one daughter cell inheriting all n copies of the plasmid at cell division, and the other none, is:

$$P_{loss} = 2^{(1-n)} \quad (1.1)$$

So for a plasmid with a copy number of 20, the probability of plasmid loss is 1.9×10^{-6} . In this case, one might expect a plasmid-free cell to arise every 500,000 cell divisions. The probability of plasmid loss will increase two-fold if plasmid copy number is reduced by just one. This highlights the critical importance of good copy number control; not only must a sufficiently high average copy be achieved, but the distribution around that average must be as narrow as possible.

Plasmid Clustering

Equation 1.1 assumes that plasmids are randomly distributed throughout the cell. However, experiments in which they were fluorescently tagged have suggested that plasmids might be localised to the $\frac{1}{4}$ and $\frac{3}{4}$ cell positions (Pogliano et al., 2001). If the plasmids diffuse as two clusters, one might expect a 50% rate of plasmid loss, or if they remained in their positions at cell division, perfect stability. The GFP tags were fusions with LacI, a protein that binds its operator sites in dimeric form and is capable of binding two such sites simultaneously in tetrameric form (O’Gorman et al., 1980). This makes it possible that a single LacI could bind two different plasmids, and with multiple operator sites per plasmid, many could be artificially bound together in a cluster. However, a study using a similar technique, but with a YFP-TetR protein fusion, suggested that ColE1 itself was localised at the cell poles (Yao et al., 2007).

Nordström and Gerdes (2003) propose that it does not matter if plasmids cluster, so long as they are recruited to the centre of the cell for replication and both products can then diffuse around the cell freely. This would make n in Equation 1.1 somewhere between the copy number at cell division (assuming all plasmids decluster) and the number of replication events before cell division plus one (assuming that only one plasmid manages to decluster). The matter certainly needs more investigation, ideally using a different technique to observe the plasmid localisation inside the cell.

A second form of plasmid clustering arises from catenation as a result of circular plasmid replication (Adams et al., 1992). Catenanes are likely resolved by Topoisomerase IV (Peng and Mariani, 1993), discussed in Section 1.6, but the process could be slow enough that catenanes exist at cell division. This form of clustering would probably only slightly reduce the effective copy number of the plasmid.

The Dimer Catastrophe

As mentioned above, plasmid dimers can emerge through homologous recombination. Dimers have two active origins of replication in only one molecule, so whilst the copy number control system perceives two entities, there is only one for the purposes of distribution at cell division. Thus, a single dimer of a high-copy number plasmid effectively reduces the copy number by 1, which doubles the probability of plasmid loss.

For ColE1, both origins of a dimer are independently active; they both produce RNAII and have the potential to initiate replication (Summers et al., 1993). This means that a single dimer is twice as likely to replicate as a single monomer, which is a significant competitive advantage. Further, termination of ColE1 replication is thought to involve meeting with the second replication fork, either stalled or moving in the other direction. This means that initiation of replication from a dimer origin produces another dimer, such that they are maintained in a resolution-deficient strain (Summers and Sherratt, 1988).

This leads to what is known as ‘The Dimer Catastrophe’ (Summers et al., 1993), depicted in Figure 1.5. After the emergence of an initial dimer, it and its descendants will rapidly establish itself in a cell line due to its replicative advantage. Eventually, cells containing only dimers will arise, containing only half the number of plasmids of a normal, monomer-

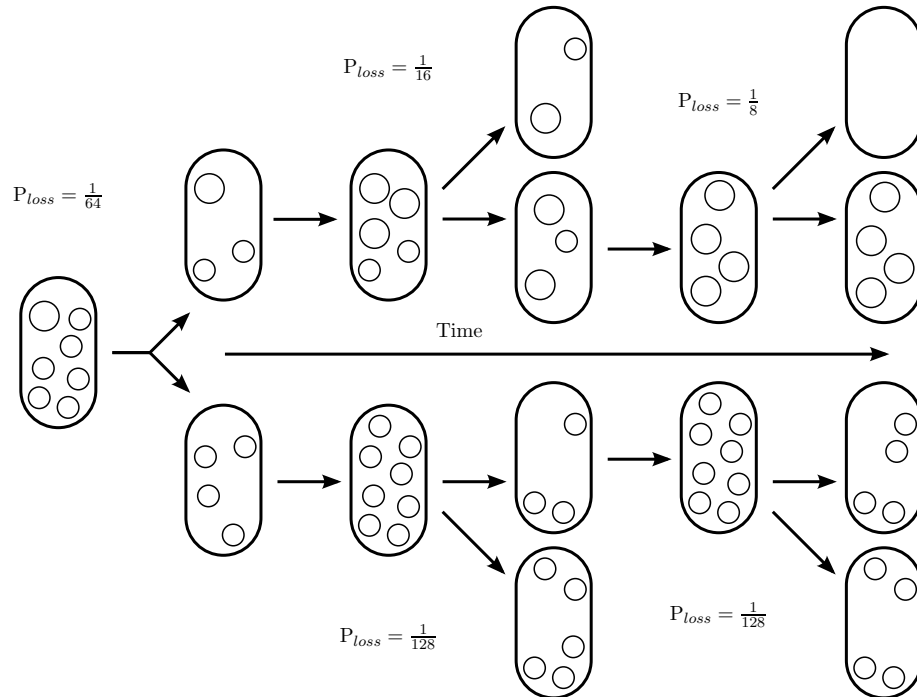


Figure 1.5: The dimer catastrophe. Plasmid monomers are depicted as small circles, plasmid dimers as large circles. After the emergence of an initial dimer, they rapidly establish themselves in the upper cell line until a dimer-only cell emerges with a high probability of plasmid loss.

only cell. From the previous example, the probability of plasmid loss for a copy number of 20 is 1.9×10^{-6} ; the dimer-only equivalent of this cell would have only 10 plasmids, and a P_{loss} of 2×10^{-3} , over 1,000 times greater. Worse still, further recombination events could lead to even larger plasmids, which have a yet greater replicative advantage, and reduce the effective copy number even more.

This runaway multimerisation is prevented *in vivo*, because dimers impose a higher metabolic load on the cell than monomers. For plasmid pBR322, dimer-only cells grow approximately 3% slower than monomer-only cells (Summers et al., 1993). This may be due to over-replication of dimers such that their copy number is more than half that of monomers (Chiang and Bremer, 1988). Summers et al. (1993) observed the fate of 600 *rec*⁺ cells containing pBR322 and found that only 2.3% were dimer-only, 1.8% were mixed monomer-dimer and the remaining 95.9% were monomer-only. Consistent with this observation, a simple model of the system predicted that assigning a metabolic penalty to dimers would restrict them to making up no more than 5% of the total plasmid content of the population, the majority of which would be in dimer-only cells (see Section 1.8).

1.5 Stability Mechanisms of ColE1

1.5.1 Xer-*cer* Multimer Resolution

Whilst the additional metabolic penalty dimers impose reduces their impact on the stability of the plasmid, ColE1 also contains the 240 bp *cer* site, which significantly increases stability by allowing multimer resolution (Summers and Sherratt, 1984). Site-specific recombination at *cer* requires the host-encoded proteins ArgR (Arginine Repressor; Stirling et al., 1988), PepA (Aminopeptidase A; Stirling et al., 1989), and recombinases XerC (Collops et al., 1990) and XerD (Blakely et al., 1993). Recent work has shown that a fifth host-encoded protein, Fis (*F*actor for *i*nversion *s*timulation), is present in the nucleoprotein complex (Blaby and Summers, 2009). It is not involved in site-specific recombination, but instead implicated in the control of the P_{cer} promoter (see Section 1.5.2).

cer and *cer*-Like Sites

The organisation of the *cer* site is depicted in Figure 1.6. The XerCD binding region of the *cer* site has a similar arrangement to those of other integrase recombinases (reviewed in Stark et al., 1992). It has a central region 8 bp long, flanked by 11 bp recombinase binding sites for XerC and XerD (Blakely et al., 1993). Adjacent to the recombinase binding site is a 210 bp accessory sequence, including an 18 bp ArgR binding site 100 bp upstream of the XerC binding site (Stirling et al., 1988). Recombination at *cer* is restricted to sites on the same molecule in direct repeat (Alén et al., 1997), and is topologically constrained by the accessory proteins ArgR and PepA (Guhathakurta et al., 1996).

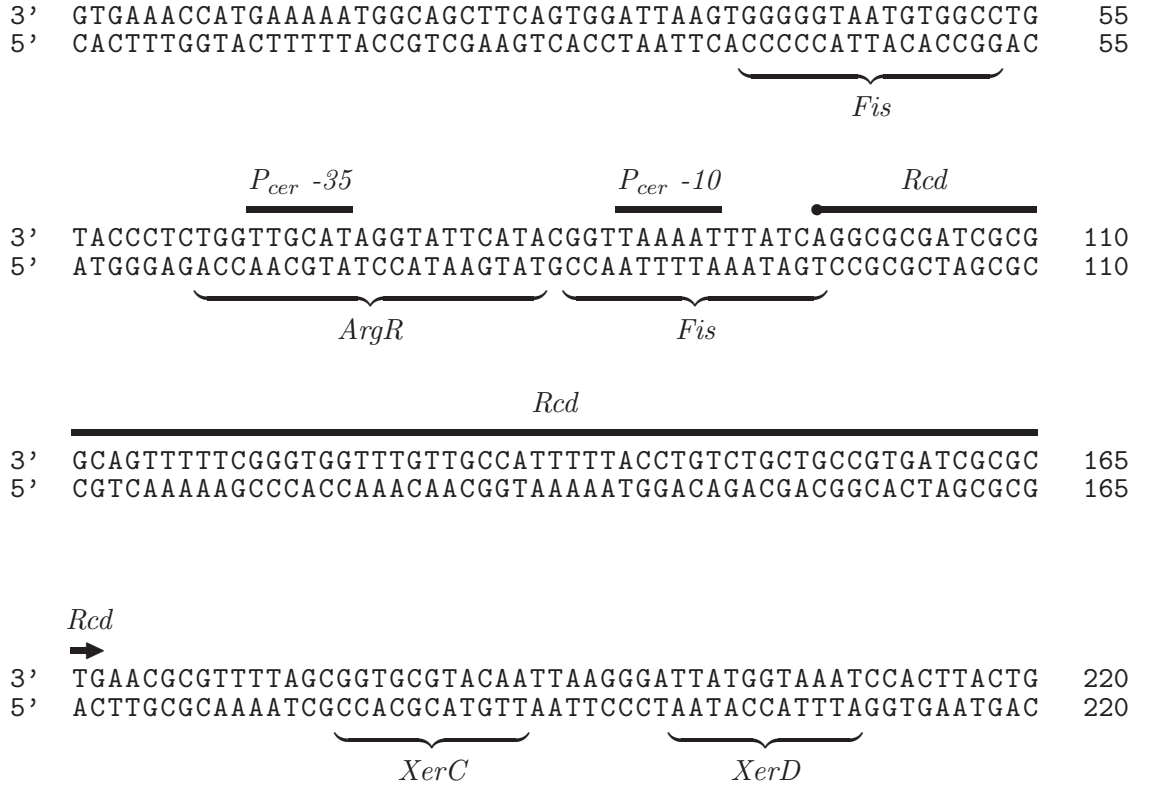


Figure 1.6: Sequence of the *cer* site with P_{cer} , *Rcd* coding region and protein binding sites.

The recombinases XerC and XerD are also implicated in the resolution of chromosomal dimers at the *cer*-homologous *dif* site (Blakely et al., 1991; Kuempel et al., 1991). In contrast to *cer*, its central region is only 6 bp long (Sherratt et al., 1993). Recombination at *dif* is unrestrained and can occur both inter- and intra-molecularly; there is also no requirement for ArgR or PepA (Blakely et al., 1991). Resolution of chromosomal dimers does, however, require FtsK (Steiner et al., 1999).

A second *cer*-like site, *psi*, is the target of XerCD recombination in plasmid pSC101 (Cornet et al., 1994); this requires PepA (Colloms et al., 1996) and ArcA (Colloms et al., 1998). *psi* has a 6 bp central region and a 160 bp accessory sequence adjacent to the core (Colloms et al., 1996; Cornet et al., 1994). Recombination at *psi*, as at *cer*, is restricted to sites on the same molecule in direct repeat, topologically constrained by the accessory proteins PepA and ArcA (Colloms et al., 1997).

Protein Roles

XerC and XerD are members of the λ integrase family of recombinases with 37% sequence similarity to one another (Blakely et al., 1993; Colloms et al., 1990). Both proteins are 33 kDa and form a heterodimer in complex with DNA that has been modelled by Subra-

manya et al. (1997). At *cer*, XerC must bind its site before XerD (Blakely et al., 1993) and is responsible for the first strand exchange (Arciszewska and Sherratt, 1995), to form a Holliday junction (McCulloch et al., 1994). It is not yet clear how the second strand exchange is completed.

ArgR is a highly conserved transcription factor found in all bacterial species (Makarova et al., 2001). The protein is 16.5 kDa and forms a hexamer that is stabilised by binding L-arginine (Van Duyne et al., 1996). It represses genes involved in the biosynthesis and transport of arginine, as well as regulating its own expression (Charlier et al., 1992).

PepA controls the transcription of the carbamoyl phosphate synthesis pathway, which itself is an intermediate involved in the pyrimidine nucleotide pathway (Charlier et al., 1995). The protein is 54.9 kDa and forms a hexamer (Sträter et al., 1999). As the name suggests, PepA is also involved in protein degradation, but this activity is not required for its role in site-specific recombination.

Both ArgR and PepA are necessary for XerCD recombination at *cer* (Stirling et al., 1989, 1988). Together, they impose topological constraint on the system, allowing recombination between sites in direct repeat on the same molecule only (Guhathakurta et al., 1996). Hence, XerCD recombination can only resolve multimers back to monomers.

Fis is an 11.2 kDa DNA-binding protein (Johnson et al., 1988). It is thought to influence the transcriptional behaviour of about 21% of the *E. coli* genome (Cho et al., 2008), either directly or indirectly, through modulating the production of DNA gyrase (Schneider et al., 1999) and topoisomerase I (Weinstein-Fischer and Altuvia, 2007). Levels of Fis in the cell are very high during exponential phase growth, but almost undetectable in stationary phase (Ali Azam et al., 1999). This is due to its transcriptional dependency on the nucleotide triphosphate CTP (Walker et al., 2004), though it is also modulated by the level of chromosomal supercoiling (Schneider et al., 2000).

Synaptic Complex Formation

There are two different models which attempt to explain how the proteins might come together and join two *cer* sites in a synaptic complex.

The Hodgman model suggests that a sub-complex forms on the single site of a plasmid monomer (Figure 1.7(a)), such that two sites come together to form a combined complex when a dimer is formed (Hodgman et al., 1998) (Figure 1.7(b)). The structure then isomerises, moving the DNA bound to the PepA hexamer of one site to the PepA hexamer of the other (Figure 1.7(c)), allowing strand exchange. Selectivity is imposed because the key interaction between PepA and ArgR is supposedly weak; two sites *in trans* will not stay together long enough for isomerisation, whereas two sites *in cis* will allow this.

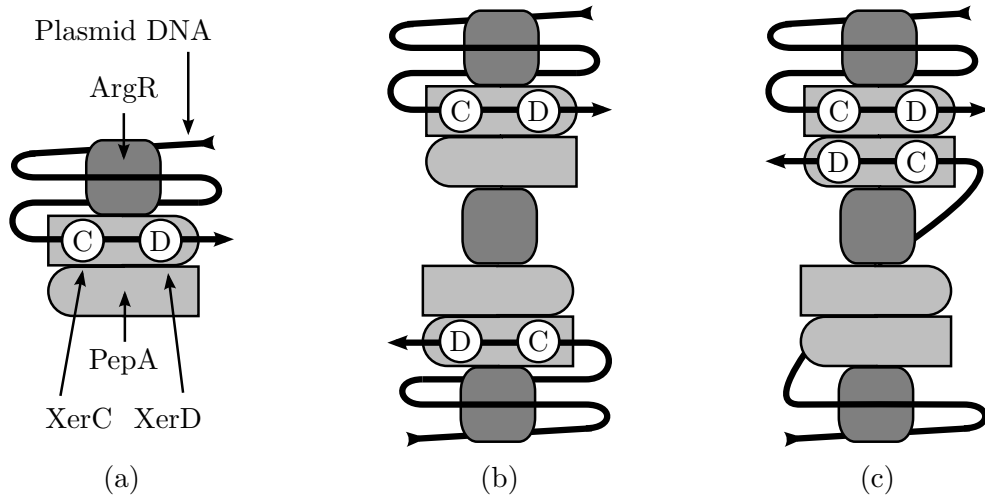


Figure 1.7: The Hodgman Model of synaptic complex formation for XerCD recombination; (a) the site forms at a single *cer* site, (b) two sites come together in antiparallel formation, (c) migration of one *cer* site and XerCD forms a synaptic complex for recombination.

The Colloms model suggests instead that the synaptic complex forms *de novo* between two sites on a plasmid dimer (Colloms et al., 1997). In this complex, the DNA of both sites is immediately bound to the same PepA hexamer, allowing strand exchange (Reijns et al., 2005). Selectivity is imposed because, as a result of the interwrapping of DNA about PepA and ArgR, three supercoils are trapped in the site, which would be energetically unfavourable *in trans* (Alén et al., 1997).

Recent work by Minh et al. (2009) looked at the structure of PepA-*cer* complexes *in vitro* using atomic force microscopy. Their results strongly suggested that two PepA hexamers are required to assemble two *cer* sites in direct repeat into a synaptic complex, supporting the Hodgman model.

1.5.2 The Rcd Checkpoint Hypothesis

In addition to its role in dimer resolution, the ColE1 *cer* site is also responsible for the expression of a short RNA, Rcd, whose transcription is driven by the P_{cer} promoter that lies within the site (Figure 1.6) (Patient and Summers, 1993). Disruption of this promoter by mutation was found to reduce plasmid stability without disrupting dimer resolution by XerCD. This suggests that dimer resolution is required, but not sufficient, to ensure plasmid stability.

It was later confirmed that mutational inactivation of Rcd, rather than of the promoter itself, caused plasmid instability (Balding et al., 2006). In this work, Rcd inactivation

was caused by point mutations at 14 different locations, and at least 2 mutant transcripts were no less stable than wild-type Rcd. These mutations were shown to cause no loss of efficiency in dimer resolution, yet still reduce the stability of plasmids into which they were introduced.

Cells in which Rcd was overexpressed were found to arrest at the point of cell division when grown on solid media, hence *Regulator of cell division* (Patient and Summers, 1993). It was proposed that Rcd imposed a checkpoint to ensure that all dimers are resolved before cell division.

Monomer-Dimer Control of Rcd Expression

Transcription of Rcd was found to occur only when plasmid multimers were present (Patient and Summers, 1993). The spacing between the -10 and -35 boxes of the promoter is relatively short, only 15 bp, and its activity can be increased with a longer spacer. It was initially proposed that differential expression between monomers and dimers was achieved by twisting the promoter within the synaptic complex, making it more favourable for RNA polymerase (Chatwin and Summers, 2001).

Recent work, however, has led to a re-evaluation of the model for regulation of the P_{cer} promoter. Blaby and Summers (2009) identified two potential Fis binding sites in *cer*, one overlapping the -10 region of P_{cer} and the other upstream of its -35 region. Fis was found to bind only the former *in vitro*. However, mutations in both sites were found to cause plasmid instability *in vivo*, without loss of dimer resolution efficiency. Rcd production from a wild-type *cer* site in plasmid monomers was shown to increase significantly in DS941 Δfis compared to DS941. Further, Rcd production was elevated in plasmid monomers with Fis binding site mutations. Both these results indicate that Fis represses Rcd expression, and that the promoter is potentially active in plasmid monomers. Rcd production was also strongly elevated in $\Delta xerC$ and $\Delta xerD$ mutants, but not in $\Delta argR$ or $\Delta pepA$ mutants, implying that XerCD also repress P_{cer} in monomers, irrespective of Fis.

These results suggested that P_{cer} is in a potentially active configuration in the plasmid monomer complex and hence are inconsistent with the model of Chatwin and Summers (2001). A new model of P_{cer} control was constructed, in which it is proposed that a single-site complex assembles at *cer*, wherein the promoter is potentially activated by twisting, with Fis, XerC and XerD repressing P_{cer} transcription together. When two sites come together, isomerisation will occur according to the Hodgman model (Figure 1.7), bringing one of the two XerCD recombinases into the synaptic complex, but away from one of the promoters. This then allows expression of Rcd from P_{cer} , as Fis alone cannot repress it. The promoter's short spacer may still play a role in preventing a transcriptional burst when the complex is disassembled by the replication fork during monomer replication.

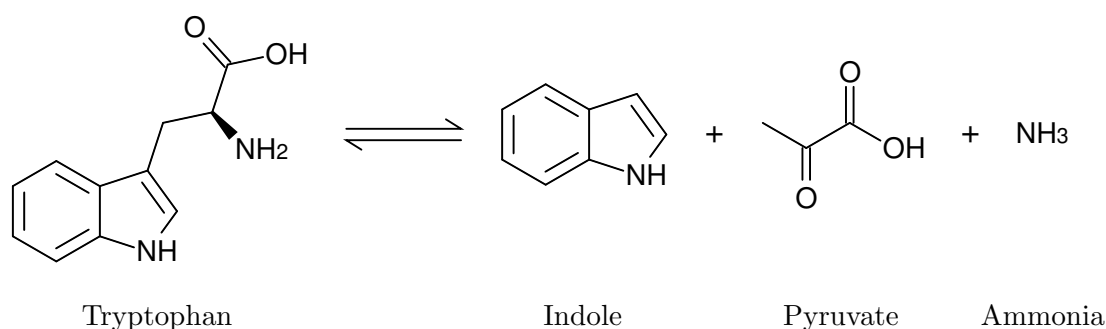


Figure 1.8: The activity of the enzyme tryptophanase; tryptophan is converted into indole, pyruvate and ammonia.

The Mechanism of Action of Rcd

The target of Rcd remained undiscovered for more than a decade after the first description of its role in plasmid stability. At first, it appeared as though Rcd might act as an antisense RNA. Analysis of other *cer*-like sites by [Sharpe et al. \(1999\)](#) revealed several Rcd analogues and a conserved region of 15nt. Work *in silico* suggested that Rcd forms a stem loop structure with a conserved region at the head of the RNA, which was a strong candidate for the active site. However, whole genome analysis and hybridisation assays failed to identify a functional target for antisense activity, suggesting that Rcd might interact directly with a protein target.

[Chant and Summers \(2007\)](#) sought the target of Rcd by RNA affinity chromatography, wherein *in vitro* transcribed Rcd was affixed to a column, through which cell lysate was applied. This study identified the enzyme tryptophanase as an Rcd-binding protein. Tryptophanase catalyses a reaction that converts tryptophan to pyruvate, ammonia and indole in equimolar ratio ([Wood et al., 1947](#)) (Figure 1.8). In the presence of Rcd, the enzyme's affinity for tryptophan increases five-fold *in vitro*, and the concentration of indole in a cell culture at low density increases three- to five-fold ([Chant and Summers, 2007](#)).

In liquid culture, indole had little effect on *E. coli* growth (measured by increase in OD₆₀₀) at concentrations less than 3 mM, but strongly inhibited it as the concentration increased beyond this, up to 6 mM ([Chant and Summers, 2007](#)). [Garbe et al. \(2000\)](#) observed loss of viability in a culture treated with 5 mM indole, supported by work in this laboratory, which observed cell lysis at concentrations of 5 mM and higher. Critically, the cell division phenotype seen when Rcd is overexpressed has also been observed in the presence of 5 mM exogenous indole (Pinero and Summers, *in prep.*). When observed under a microscope, indole-treated cells do not divide, instead growing slowly into filaments of 2-4 cells in length. It is therefore indole, not Rcd, that is directly responsible for the observed inhibition of cell division and growth.

Indole is an aromatic heterocyclic organic compound. It is composed of a five-membered pyrrole ring that contains nitrogen, fused to a six-membered benzene ring (Roychowdhury and Basak, 1975). Indole production is a diagnostic marker for the identification of *E. coli* (Sonnenwirth, 1980).

Wang et al. (2001) showed that indole can act as an extracellular signalling molecule, activating expression of *astD*, an aldehyde dehydrogenase, *tnaB*, a tryptophan transporter, and *gabT*, the first enzyme in the 4-aminobutyrate pathway. Di Martino et al. (2002) demonstrated that tryptophanase was involved in biofilm formation and later showed that indole could restore biofilm formation activity in a tryptophanase deficient mutant (Di Martino et al., 2003). Hirakawa et al. (2005) implicated indole in the regulation of xenobiotic exporter expression, showing that indole treatment of *E. coli* cells could confer resistance to rhodamine 6G and SDS through activation of the *mdtEF* and *acrD* genes respectively. Indole clearly has wide-ranging effects on the cell.

The Revised Rcd Checkpoint

The description of the role of Rcd in plasmid stability was adjusted to incorporate the discovery of its interaction with tryptophanase. The Rcd checkpoint hypothesis asserts that when a dimer emerges through homologous recombination, Rcd is transcribed from the *cer* site synaptic complex. It then interacts with tryptophanase to increase the production of indole inside the cell, leading to a transient increase in intracellular indole concentration, which arrests cell division. Dimer resolution by Xer-*cer* converts multimers back to monomers and the cell is then allowed to pass the checkpoint and divide. This is necessary, because it is supposed that the rate of dimer resolution is naturally slow (supported by Oram et al., 1997). Rcd synthesis ceases as dimers become monomers, indole concentration falls as it is exported and Rcd concentration falls as it is diluted by cell growth, restoring the cell to normal growth and a monomer-only existence.

1.6 Type II Topoisomerases

Topoisomerases are a class of enzymes found throughout living organisms, responsible for the necessary topological rearrangement of DNA during replication, transcription, recombination and partitioning processes (reviewed in Champoux, 2001; Wang, 2002). Within a closed circular molecule, the number of times the two DNA strands are wound about each other is described by the linkage number. This can vary at a local level, though any alteration must be compensated for elsewhere in the molecule, and overall linkage number can only be altered through strand-breakage; these issues are resolved by topoisomerases (reviewed in Espeli and Marians, 2004).

Topoisomerases are classed as either Type I, in which the DNA strands are broken one at a time or Type II, in which they are broken together. Type IA topoisomerases create a single-strand break in negatively supercoiled DNA and pass another single strand or double helix through the gap, enabling them to, amongst other things, relax negatively supercoiled DNA. Type IB topoisomerases relax both positively and negatively supercoiled DNA by creating a single-strand break and rotating the protein-DNA complex before religation. Type IIA and type IIB topoisomerases are structurally distinct families that create a double-strand break and pass another intact double helix through the gap. This requires ATP hydrolysis, but enables them to perform a variety of topological rearrangements such as introducing or removing supercoils and catenating or decatenating closed circles.

In *Escherichia coli*, topoisomerases I and III are type IA, whereas DNA gyrase and topoisomerase IV are type IIA (Wang, 2002). In general, DNA gyrase is responsible for maintaining negative supercoils, antagonised by the relaxation activity of topoisomerases I and IV (Zechiedrich et al., 2000). Topoisomerases IV and III ensure that precatenanes produced by the replication process are resolved, preventing the creation of concatenated daughter molecules (Adams et al., 1992; DiGate and Marians, 1988; Wang et al., 2008). Of particular interest to plasmids, and to this work, are DNA gyrase and topoisomerase IV (reviewed in Drlica and Zhao, 1997; Nöllmann et al., 2007).

1.6.1 DNA Gyrase

DNA gyrase was first identified by Gellert et al. (1976) and shown to introduce negative supercoils to closed circular DNA in a reaction dependent on ATP, Mg^{2+} and stimulated by spermidine. In the absence of ATP, it relaxes supercoiled DNA instead (Gellert et al., 1977), though it is not clear if it performs this function *in vivo*. In the cell, DNA gyrase maintains negative supercoiling (Drlica and Snyder, 1978), which is known to affect promoter strength (Botchan et al., 1973). It also resolves positive supercoils generated ahead of RNA polymerase in transcription and ensures fork movement ahead of DNA polymerase in replication (Kreuzer and Cozzarelli, 1979). It has been implicated in decatenation of the chromosome (Steck and Drlica, 1984), but that role is now believed to be performed by the more efficient topoisomerase IV (see Section 1.6.2).

Structurally, the enzyme consists of two subunits in an A_2B_2 complex, which were originally identified as NalA and Cou, named for the antibiotic resistances they determined (Mizuuchi et al., 1978), but later renamed GyrA and GyrB, respectively. GyrA is involved in the breakage and reunion of DNA (Sugino et al., 1977), whereas GyrB is responsible for ATP hydrolysis (Sugino et al., 1978). Recently, Costenaro et al. (2005, 2007) determined the modular structure of GyrA and GyrB by small-angle x-ray scattering. GyrA consists of an amino-terminal domain containing the active-site tyrosines and a carboxyl-terminal domain which wraps the DNA around itself. GyrB is organised into three domains; the ATPase, a Toprim domain that coordinates Mg^{2+} ions and a Tail for DNA binding.

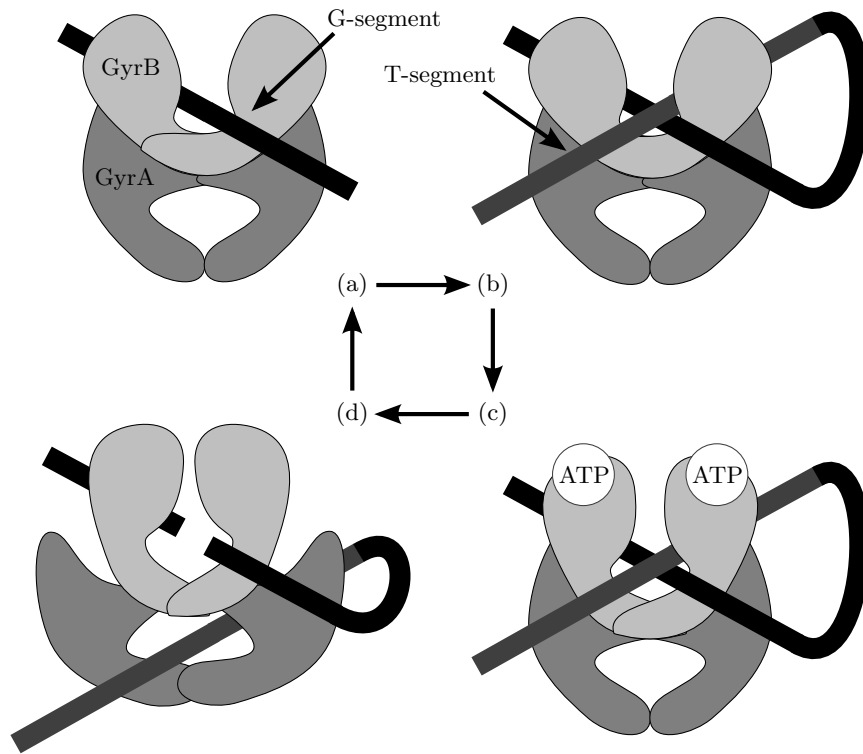


Figure 1.9: The activity of DNA gyrase; (a) subunits GyrA and GyrB form a DNA-protein complex with the G-segment, (b) the structure of GyrA encourages looping of the T-segment to make a cross formation, (c) ATP binding at GyrB causes conformational change in the complex, (d) hydrolysis results in cleavage of the G-segment, through which the T-segment is passed, increasing the linkage number of the molecule by 2, (a) the G-segment is repaired, the T-segment released and the enzyme is reset.

The supercoiling mechanism of gyrase (Figure 1.9) is explained by the ‘sign-inversion’ model (Brown and Cozzarelli, 1979). The enzyme first binds two DNA duplexes in a cross formation. One segment, the gate or G-segment, is cleaved, and the other, the transport or T-segment, is passed through the gap before it is resealed. Thus, the previously positive crossing becomes a negative crossing, changing the linking number of the molecule by 2, which is characteristic of type II topoisomerases.

Mechanistically, GyrA binds the G-segment at its N-terminal domain (Cabral et al., 1997), then its C-terminal domain loops the nearby DNA to form the T segment (Reece and Maxwell, 1991). ATP binding at GyrB causes conformational changes that lead to transesterification cleavage of the G-segment by tyrosines at the GyrA active site (Horowitz and Wang, 1987). The T-segment is then driven through the break by the sequential hydrolysis of two ATP (Baird et al., 1999). Finally, the break is repaired and all components of the reaction are released.

Gyrase is the target of a number of antibiotics, including coumarins and quinolones, which target either GyrA or GyrB, with different consequences (reviewed in Maxwell, 1997; Oblak et al., 2007). Those that target GyrA, such as nalidixic acid, are thought to disrupt the

breakage and reunion behaviour of the enzyme, trapping the DNA-enzyme complex and resulting in double-strand breaks (Gellert et al., 1977; Sugino et al., 1977). Those that target GyrB, such as novobiocin, are believed to competitively inhibit the ATPase activity of the subunit (Sugino et al., 1978). Work by Lewis et al. (1996) has suggested that these inhibitors do not explicitly compete for the ATP binding site, but instead bind in close proximity.

1.6.2 Topoisomerase IV

Topoisomerase IV was first identified by Kato et al. (1990), who found that the *parC* gene, essential for chromosome partition in *E. coli*, showed strong sequence similarity with *gyrA*. The second subunit, encoded by *parE*, was discovered in the same work and showed strong sequence similarity with *gyrB*. The complete enzyme is of the form C₂E₂. The full length crystal structure of ParC was determined by Corbett et al. (2005); it was found to be highly similar to that of GyrA, though with a more open conformation. Only a partial structure has been found for ParE (Bellon et al., 2004), which was highly similar to that of GyrB.

The enzyme was shown to relax both positively and negatively supercoiled DNA, as well as decatenate knotted DNA in the presence of ATP and Mg²⁺ (Kato et al., 1992). In the cell, topoisomerase IV acts in concert with topoisomerase I against DNA gyrase to maintain the correct level of negative supercoiling (Zechiedrich et al., 2000). More critically, it is responsible for the decatenation of chromosomes prior to cell division (Kato et al., 1990), though DNA gyrase can perform the same function with 1% efficiency, if necessary (Zechiedrich and Cozzarelli, 1995). Adams et al. (1992) have shown that topoisomerase IV is also responsible for the decatenation of plasmids during their replication.

The mechanism of action of topoisomerase IV is broadly similar to that of DNA gyrase, except that it does not wrap a nearby section of DNA around itself (Peng and Mariani, 1995). This means that topoisomerase IV favours intermolecular strand passage; it would rather transport a DNA duplex from another molecule through the double-strand break it creates than pass a section of a molecule through itself, as gyrase prefers.

Topoisomerase IV is a target for the same antibiotics as DNA gyrase, due to their sequence and structural similarity (Khodursky et al., 1995). However, it typically requires a significantly higher concentration to inhibit topoisomerase IV compared to gyrase (Hoshino et al., 1994). In the work of Bellon et al. (2004), equivalent single amino acid residue substitution in the side chains of ParE and GyrB produced completely opposite changes in the novobiocin resistance and kinetics of the enzymes, suggesting an explanation for their different responses to antibiotics.

1.6.3 Topoisomerase Roles in Plasmid Replication and Control

DNA gyrase has a wide range of activities in the cell and could therefore affect the behaviour of plasmids in a number of ways. The supercoiled state of the plasmid could potentially affect the transcriptional activity of any plasmid-encoded genes, including RNAI and RNAII for ColE1, which control replication initiation. The enzyme is also an essential component of ColE1 replication (Kaguni and Kornberg, 1984), which is prevented either by a defective GyrB (Orr and Staudenbauer, 1981) or by the addition of a known gyrase inhibitor (Gellert et al., 1977). Inhibition of gyrase activity preferentially suppresses plasmid replication over that of the chromosome (Uhlen and Nordström, 1985). A plasmid can be cured from a strain by the addition of a sub-lethal concentration of such an inhibitor (Wolfson et al., 1982), and inhibitors that target GyrB do so more efficiently (Hooper et al., 1984).

Topoisomerase IV will also affect the supercoiled state of plasmids due to its relaxation activity. It has also been implicated in decatenation of the products of replication (Adams et al., 1992; Peng and Marians, 1993), which is encouraged by the supercoiling activity of DNA gyrase (Martínez-Robles et al., 2009).

1.7 Mathematical Models of ColE1 Replication and Control

1.7.1 Deterministic vs. Stochastic Systems

Modelling biological systems typically involves simplification; it is often not known how cellular components outside of a specific subsystem affect that subsystem, or it is at least too complicated to include them all. Values or rates determined *in vivo* are key to accurately representing a system, but they are almost always averages across a population, and deviation from this average is often not captured; the technology to observe behaviour on the level of a single cell is still in development (see Elowitz et al., 2002; Ozbudak et al., 2002, for examples of this kind of work). Data from controlled experiments *in vitro* is often more useful, as they allow characterisation of specific, isolated reactions.

Mathematically, a model can be deterministic or stochastic. In a deterministic model, there is no randomness in the behaviour of the model; its state at any given time can be predicted from its initial state. Such a model would consist of a series of differential equations that describe the relationships between the different variables in the system, which are often continuous, rather than discrete. These models only explain the behaviour of biological systems accurately when they involve thousands of molecules and are based on *in vivo* population averages.

Systems involving molecules that are few in number, such as plasmids in a cell, are more accurately described by stochastic models. A stochastic model takes into account the random behaviour of a system; its state at any given time is not easily predictable, though some of its behaviour may be. Such a model requires that the behaviour of the system be described in terms of probabilities; how likely it is to react one way or another. The Monte Carlo method (described in [Metropolis, 1987](#)) is often used to then observe the behaviour of the system, wherein, with the help of a random number generator, the system is tested over and over again, predicting both population averages and variation from the mean. When the reactions and reaction rates of a system are defined, the Gillespie algorithm ([Gillespie, 1977](#)) can be used to determine statistically accurate behaviour for that system.

1.7.2 Deterministic Models

The first mathematical model of ColE1 replication was published by [Ataai and Shuler \(1986\)](#). It considered the likely average copy numbers that might be produced for different time windows, within which RNAII is susceptible to inhibition by RNAI. The interaction between the two was correctly modelled exponentially, but the statistical variation from the mean copy number was not considered. The work supported the supposition of [Tomizawa and Som \(1984\)](#) that the Rom protein increases the binding constant between RNAI and RNAII and rejected one hypothesis of [Cesareni et al. \(1984\)](#) that RNAI might exert control by making RNAII more susceptible to degradation by RNase.

The model of [Bremer and Lin-Chao \(1986\)](#) was considerably more involved, deriving plasmid behaviour from a pair of ordinary differential equations for plasmid and RNAI concentration. They assumed hyperbolic inhibition and also made an attempt to consider replication of plasmid multimers, mistakenly concluding that they would have the same copy number as monomers. [Keasling and Palsson \(1989\)](#) took the equations and assumptions of this model and did a comprehensive parametric analysis of the dynamics of copy number control. The work used hyperbolic inhibition but did suggest that the effects of Rom could be included in the interaction constant between RNAI and RNAII due to its saturation in the cell.

[Brenner and Tomizawa \(1991\)](#) published a minimal model along with quantitative data on the number of RNAI and RNAII transcripts found *in vivo*. They also measured the half-life of the two RNAs and, using steady-state assumptions, estimated transcription rates. In contrast, the model of [Brendel and Perelson \(1993\)](#) was intricately detailed, considering all of the possible states of the components of control and the transitions between those states. However, this model assumed hyperbolic inhibition and was more concerned with the plasmid achieving a sufficiently high average copy number than reducing variation through efficient control. This work solved its differential equations analytically, but the model itself was later formalised as a stochastic Petri net and simulated by [Goss and Peccoud \(1998\)](#).

1.7.3 Stochastic Models

The approach to modelling the system changed with [Ehrenberg \(1996\)](#), who used a master equation approach to analyse the difference between hyperbolic and exponential control. Copy number control is dependent on a discrete and relatively low number of molecules in the cell; it is therefore sensitive to stochastic events that change the cell's state, in terms of these molecules. The master equation approach takes this into account and gives information on the statistical variation in the system. This work specifically demonstrated that exponential control reduced the likelihood of plasmid loss for ColE1 when compared to hyperbolic control. Further, it suggested that Rom could narrow the distribution of plasmid copy number before cell division, though, due to limited resources, did not consider the case where the plasmid is randomly distributed at cell division and used very low copy numbers.

This effort was followed up in [Paulsson et al. \(1998\)](#), in which the requirements for rapid copy number adjustment were considered with a deterministic model. The work touched briefly on the replication of plasmid dimers, suggesting that they would be at least equal in copy number to monomers with hyperbolic control, and likely more than half the copy number of monomers with exponential control. In a second follow-up, [Paulsson and Ehrenberg \(1998\)](#) expanded on the original stochastic model to consider random distribution of the plasmid at cell division and the parameter ranges that would reasonably stabilise the plasmid. Finally, [Paulsson and Ehrenberg \(2001\)](#) reviewed in detail the construction and findings of all these models and the equivalents for plasmid R1, analysing the impact of noise on copy number control.

1.8 Mathematical Models of the Dimer Catastrophe

The dimer catastrophe hypothesis was developed using a computer model of plasmid behaviour in a growing cell population ([Summers et al., 1993](#)). The model was relatively simple; each 'cell' was summarised as the number of plasmid monomers and dimers it contained, and the time at which it was due to divide. The simulation would determine which cell was next due to divide and then determine the distribution of plasmids to daughter cells at random. Each daughter cell then had its entire lifetime simulated immediately; the plasmid would replicate back to full copy number, each time choosing an origin to replicate from at random, such that a dimer was twice as likely to replicate as a monomer. After testing for the slim chance that a recombination event might create a dimer from two monomers, the division times of the new cells were calculated and then put back into the population. The population was limited to 125,000 cells and when reached, it would be subcultured randomly back to 8,000 cells before proceeding. Information on the population was written to file periodically until the desired simulation time elapsed.

This model demonstrated several interesting behaviours. Firstly, it showed that dimers would indeed out-replicate monomers and eventually comprise the whole population. It was then found that just a small metabolic penalty for each dimer (such that a dimer-only cell grew 10% more slowly than a monomer-only cell) stabilised the population and reduced the overall proportion of dimers to approximately 4.5% of the total plasmids, over half of which were in dimer-only cells. The difference in growth rates was then found to occur *in vivo*. A number of initial conditions were tested, each arriving at the same steady-state distribution of plasmids in the population over time.

Prior to the start of the present work, this model was reconstructed with the sparse information available in the publication (MEng thesis, Field, 2006). Modern hardware allowed for a more flexible simulation and the testing of many parameter sets. As well as reproducing the results of the original, relationships between parameters and the stability of the plasmid in the model were discovered. Plasmid stability was shown to be logarithmically related to the rate of homologous recombination, that is, the rate of dimer formation. The metabolic penalty associated with dimers also affected the steady-state of the population.

Two additional effects were tested within the framework of this model. Firstly, inclusion of dimer resolution only significantly affected the steady-state when the probability of resolution for each dimer was 20% per generation, which is perhaps unreasonably high for the real system. Secondly, plasmid clustering was incorporated into the system in the form of plasmid pairing at a given rate. It was found to have very little effect on plasmid stability at low rates and, counter-intuitively, to improve stability at high rates, perhaps because paired dimers reduced the chance of both daughter cells inheriting dimers.

1.9 Aims and Objectives

This work set out to investigate the dimer catastrophe theory and to understand how plasmid stability is achieved by the mechanisms encoded by ColE1. Parallel approaches were taken *in silico* and through experimentation, aiming to:

- Construct a new framework for the computer model of plasmid behaviour in a cell population, as existing models were technically flawed.
- Implement a model for ColE1 plasmid replication, to take into account copy number variation as well as the dynamics of recombination.
- Introduce this model to the new framework, to allow for a more accurate assessment of how the different model parameters affect the behaviour of the system and the effectiveness of dimer resolution in the absence of the Rcd checkpoint.
- Critically analyse the Rcd checkpoint and specify the role of indole therein.
- Identify a target for indole, to understand the mechanism by which the Rcd checkpoint maximises the stable maintenance of ColE1.

Chapter

2

Materials and Methods

2.1 Strains

Strain	Genotype	Reference
W3110	λ^- , <i>IN(rrnD – rrnE)</i> 1, <i>rph</i> -1	Bachmann (1972)
JC8679	<i>thr</i> -1, <i>araC</i> 14, <i>leuB</i> 6(<i>Am</i>), Δ (<i>gpt-proA</i>)62, <i>lacY</i> 1, <i>tsx</i> -33, <i>glnV</i> 44(<i>AS</i>), <i>galK</i> 2(<i>Oc</i>), λ^- , <i>sbcA</i> 23(<i>Rac</i>), <i>his</i> -60, <i>relA</i> 1, <i>recB</i> 21, <i>recC</i> 22, <i>rpsL</i> 31(<i>strR</i>), <i>xylA</i> 5, <i>mtl</i> -1, <i>argE</i> 3(<i>Oc</i>), <i>thi</i> -1	Gillen et al. (1981)
DS941	<i>thr</i> -1, <i>araC</i> 14, <i>leuB</i> 6(<i>Am</i>), Δ (<i>gpt-proA</i>)62, <i>lacY</i> 1, <i>lacZ</i> Δ M15, <i>lacI</i> ^q , <i>tsx</i> -33, <i>qsr</i> '-0, <i>glnV</i> 44(<i>AS</i>), <i>galK</i> 2(<i>Oc</i>), λ^- , <i>Rac</i> -0, <i>hisG</i> 4(<i>Oc</i>), <i>rfbC</i> 1, <i>mgl</i> -51, <i>rpoS</i> 396(<i>Am</i>), <i>rpsL</i> 31(<i>strR</i>), <i>kdgK</i> 51, <i>xylA</i> 5, <i>mtl</i> -1, <i>recF</i> , <i>argE</i> 3(<i>Oc</i>), <i>thi</i> -1	Summers and Sherratt (1988)

Table 2.1: Strains of *Escherichia coli* used during this work.

2.2 Plasmids

Plasmid	Size (bp)	Description	Reference
pBR322	4361	pMB1 origin, amp^R , tet^R	Bolivar et al. (1977)
pUC18	2686	pMB1 origin, Δrop , amp^R , $lacI$, $lacZ$	Norrander et al. (1983)
pUC19	2686	pMB1 origin, Δrop , amp^R , $lacI$, $lacZ$	Norrander et al. (1983)

Table 2.2: Plasmids used during this work.

2.3 Media

All media was prepared with deionised water for a total volume of 1 litre. The pH was adjusted with NaOH. Each solution was autoclaved at 121 °C for 20 minutes prior to use.

2.3.1 Lysogeny Broth (LB), with Agar (LA)

NaCl	10 g
Tryptone	10 g
Yeast extract	5 g
pH 7.5	

For LA, 15g agar was added.

2.3.2 Super Optimal Broth with Catabolite Repression (SOC)

Tryptone	20 g	(2%)
Yeast extract	5 g	(0.5%)
Glucose	3.6 g	(20 mM)
MgSO ₄	1.2 g	(10 mM)
MgCl	0.95 g	(10 mM)
NaCl	0.58 g	(10 mM)
KCl	0.19 g	(2.5 mM)
pH 7.0		

2.4 Antibiotics

Antibiotic	Stock Concentration	Working Concentration
Carbenicillin	100 mg ml ⁻¹	100 µg ml ⁻¹
Chloramphenicol	34 mg ml ⁻¹	34 µg ml ⁻¹
Nalidixic Acid	30 mg ml ⁻¹	30 µg ml ⁻¹

Table 2.3: Antibiotics used in this work.

2.5 Buffers and Solutions

2.5.1 ATP-Free Reaction Buffer for DNA Gyrase

The recipe for this buffer is derived from the standard reaction buffer for DNA gyrase (New England Biolabs). An ATP-free version was required to allow for a lower concentration of ATP in the reaction. The standard reaction buffer contains 1.75 mM ATP.

Tris-HCl	35 mM
KCl	24 mM
Spermidine	5 mM
MgCl ₂	4 mM
DTT	2 mM
BSA	0.1 mg ml ⁻¹
Glycerol	6.5%

pH 7.5

2.5.2 Electrophoresis Buffer (TAE)

This 50× stock solution was prepared with deionised water for a total volume of 1 litre. It was autoclaved at 121 °C for 20 minutes and diluted to 1× prior to use.

Tris base	242.2 g
Glacial acetic acid	57.1 ml
0.5 M EDTA (pH 8.0)	100 ml

2.5.3 Ethidium Bromide

A stock solution of 10 mg ml^{-1} ethidium bromide was prepared by dissolving a 10 mg tablet in 1 ml of deionised water. For the staining of agarose gel, a working solution was prepared by adding 100 μl of stock solution to 2 litres of $1\times\text{TAE}$ buffer, for a final concentration of $0.5 \mu\text{g ml}^{-1}$.

2.5.4 Indole and Indole Analogs

Stock solutions of these chemicals were prepared fresh before each experiment. All solutions were prepared in ethanol, with the exception of Indole-3-acetic acid, which was prepared in water.

Chemical	Stock Concentration
Indole	0.5 M
1-Acetyldoline	0.25 M
3- β -Indoleacrylic acid	0.125 M
Indole-3-acetic acid	0.5 M
Indoline	0.5 M
Isoquinoline	0.5 M
Pyrrole	0.5 M
Quinoline	0.5 M
Tryptamine	0.5 M

2.6 Microbiological Techniques

2.6.1 Cell Culture

Escherichia coli strains were routinely grown in LB medium at 37°C with 120 rpm shaking. Typically, a 10 ml culture was inoculated at low density in a 20 ml plastic universal tube, with antibiotics as necessary, and left to grow for around 16 hours overnight.

After transformation and for short-term storage, strains were spread or streaked out to single colony on LA plates, with antibiotics as necessary. Strains were left to grow overnight at 37°C before storage at 4°C , wrapped in parafilm to prevent dehydration.

For long-term storage, 1.35 ml of stationary phase culture was mixed with 0.45 ml sterile 60% glycerol in a 2 ml cryovial. After vortexing, this was frozen in liquid nitrogen and stored at -80°C .

2.6.2 Spectrophotometry

To assess the cell density of a liquid culture, its optical density was normally measured at 600 nm with a DU650 spectrophotometer (Beckman).

2.6.3 Centrifugation

Volumes of liquid less than 1.5 ml were typically centrifuged at $12,000g$ in a Minispin tabletop centrifuge (Eppendorf). Those greater than 1.5 ml were typically centrifuged at $3,220g$ in a 5810R floor centrifuge (Eppendorf).

2.6.4 Transformation

For electrocompetent cells, an overnight culture of *E. coli* was chilled to 4°C and spun down; the supernatant discarded. The cells were then resuspended in 10 ml chilled SDW and spun down again. This was repeated for a total of 3 washes before cells were resuspended in just 200 μl chilled SDW.

For transformation, 50 μl of this suspension was transferred to an electroporation cuvette with a 0.1 cm gap and approximately 1 μg plasmid DNA was gently mixed in. Cells were then subject to a 1.68 kV shock through a 200 Ω -25 μF resistor-capacitor pair with a Gene Pulser electroporator (Biorad). 0.5 ml of either LB or SOC was immediately added and cells were left to recover at 37°C for 1 hour before plating on LA with appropriate antibiotics.

2.7 DNA Manipulation

2.7.1 Plasmid DNA Extraction

Plasmid DNA was routinely extracted using the Qiaprep Spin Miniprep kit, according to the manufacturer's instructions (Qiagen).

2.7.2 Gel Electrophoresis

Gels were typically 1% (w/v) agarose and measured 150 mm×100 mm×4 mm. 5 µl samples were mixed with 1 µl 5×DNA Loading Buffer (Bioline) prior to loading. Electrophoresis was carried out in 1×TAE buffer, normally at 80 V for 90 minutes.

For visualisation, gels were stained post-electrophoresis in 1×TAE with 0.5 µg ml⁻¹ ethidium bromide for 20 minutes, prepared every day or two. A UV transilluminator hood and camera were used to capture the necessary images (Biorad).

2.7.3 Gel Extraction

Where a particular band of DNA was required for transformation, care was taken not to expose the agarose gel to excessive UV. A scalpel blade was used to cut the band from the agarose gel, with the transilluminator set to low-power mode. The DNA was then purified using the Qiaquick Spin kit, according to the manufacturer's instructions (Qiagen).

2.7.4 Restriction Digests

For restriction of plasmid DNA, 20 µl at approximately 0.5 µg µl⁻¹ was mixed with 20 µl sterilised distilled water, 5 µl of the enzyme stock solution and 5 µl of the manufacturer's suggested buffer (New England Biolabs). The mixture was left at 37 °C for 1 hour before heat inactivation at 65 °C for 20 minutes.

2.8 Specific Assays

2.8.1 Indole-Plasmid Assay

The cultures in this assay were 100 ml in volume, inoculated with 1 ml overnight culture and grown in 200 ml conical flasks in a 37 °C water bath (Grant). Carbenicillin was added to the medium to maintain the plasmid. OD₆₀₀ was regularly measured until it reached around 0.2, at which point the appropriate chemical solutions were added. Sample volumes were normalised according to OD₆₀₀ such that approximately the same cell mass was taken for each. Plasmid DNA was extracted immediately and the samples stored at 4 °C until required for electrophoresis.

2.8.2 Gel Densitometry

Agarose gel images were analysed with Quantity One software (Biorad). Each image was normalised to the same average brightness before a 3-pixel weighted mean filter was applied to remove gaussian noise. The gel lanes were identified manually, the background level of brightness removed and the individual bands selected using both manual intervention and software automation. Each of the bands was then idealised as a gaussian distribution of brightness before their weights were calculated. For each sample, the calculated weights for the bands in the appropriate lane were summed for simplicity, even though the relationship between brightness and DNA weight is not absolutely linear. Then to account for additive noise, the sample value for t_0 was subtracted from each sample in the data series.

2.8.3 Indole-Gyrase Assay

A typical mixture for this assay was as follows:

DNA Gyrase (1/25 dilution from stock)	1 μ l
Reaction Buffer	6 μ l
Relaxed pUC19 DNA (1 μ g μ l ⁻¹)	1 μ l
Indole (various stock concentrations)	0.4 μ l
or	
Nalidixic Acid (30 mg ml ⁻¹)	6 μ l
SDW	to total
Total Volume	30 μ l

5 μ l samples were taken at regular intervals. After heat inactivation at 65 °C for 20 minutes, samples were stored at 4 °C. DNA was separated according to its superhelical state by agarose gel electrophoresis, and visualised by ethidium bromide staining.

2.9 Computer Modelling

2.9.1 Programming

All models were programmed in C++. In addition to the standard libraries, iostream, fstream, iomanip, string, sstream, vector, cmath and algorithm were used. To generate high quality pseudo-random numbers, a Mersenne twister generator was used (Matsumoto & Nishimura, 1998), the code for which was courtesy of Richard J. Wagner.

Models were tested on a home computer with dual 1.6 GHz Atom processors (Intel) and 1 GB RAM. The code was initially compiled with the GCC compiler, version 4.3.3-5ubuntu4 and later with version 4.3.2-1.1.

2.9.2 Processing

For processing the models with various input parameters, the majority were run on the ‘elephant’ server of the School of Biosciences, University of Cambridge, with 8×4-core 2.3 GHz Opteron processors (AMD) and 128 GB RAM.

Parallel processing was achieved with CamGrid, a distributed computing resource that connects over 1150 idle machines around the University of Cambridge for high throughput use. The system is managed by the Condor set of software tools, version 7.2.5. The machines connected to the network have mostly 64-bit processors, and the majority run a Linux operating system.

2.9.3 Plasmid Stability Index

The calculation of the plasmid stability index for C cells, where each cell, c , contained a number of monomers, M_n , and dimers, D_n , was as follows:

$$P_{loss} = \frac{1}{C} \times \sum_c 2^{(1-(M_c+D_c))} \quad (2.1)$$

This is the probability of plasmid loss averaged across the population. The true probability of all daughter cells inheriting at least one plasmid, should the population all divide at once, is:

$$1 - P_{loss} = \prod_c (1 - 2^{(1-(M_c+D_c))})$$

As plasmid copy number is discrete, this can be simplified by considering the number of cells, C_p , containing p plasmids:

$$1 - P_{loss} = \prod_p (1 - 2^{(1-p)})^{C_p}$$

Taking logarithms:

$$\ln(1 - P_{loss}) = \sum_p C_p \ln(1 - 2^{(1-p)})$$

Using the Maclaurin series for the expansion of $\ln(1 - x)$:

$$\sum_{n=1}^{\infty} \frac{(P_{loss})^n}{n} = \sum_{n=1}^{\infty} \sum_p C_p (2^{(1-p)})^n$$

Approximate to the first order and normalise according to the total number of cells, C :

$$P_{loss} = \frac{1}{C} \sum_p (C_p 2^{(1-p)})$$

Which is equivalent to equation 2.1. This holds well where the minimum number of plasmids in a cell is half of the maximum, but is less accurate when there is a larger factor between these limits.

Modelling Plasmid Behaviour in a Cell Population

3.1 Introduction

The original computer model of the dimer catastrophe simulated the plasmid content of a cell population as it underwent exponential growth (Summers et al., 1993). The computer hardware of the time imposed certain limitations on the scope of this simulation:

- The population could not exceed a given maximum, and was thus ‘subcultured’ at regular intervals.
- No matter what the initial contents of the cell were, the model plasmid would always replicate up to its full, fixed copy number.
- The state of the cell population was recorded only as a count of the number of cells containing a given number of dimers, averaged across all simulation runs.
- The total simulation time and the number of runs were limited.

With these restrictions, the model showed that imposing even a small growth penalty on cells containing dimers restrained the catastrophe, and prevented runaway multimerisation. It also showed that the majority of dimers in the cell population were in cells containing only dimers, and that the steady-state concentration of dimers was related to the rate of dimer formation by homologous recombination.

In work prior to the study reported in this thesis, the original simulation was reconstructed with modern software in order to investigate further aspects of plasmid behaviour (MEng thesis, Field, 2006). This second version operated considerably faster than the original such

that more accurate averages could be calculated from a greater number of individual runs. After recreating the results of the original, this model went on to show that the plasmid stability index (the average probability of plasmid loss across the whole cell population, see Chapter 2) was logarithmically related to the dimer formation rate. Dimer resolution, that is dimer-to-monomer recombination, was included in the updated model and it was found that a resolution chance of around 20% per dimer per generation was required to have a significant effect on plasmid stability. Plasmid clustering (see Section 1.4.2) was also investigated and predicted to have a minor beneficial effect on stability. This model was useful but still suffered from many of the drawbacks of the original: the cell population was subcultured, the plasmid replicated back to full, fixed copy number and the simulation results were not being thoroughly statistically analysed.

This chapter describes the construction of a third computer simulation of plasmid behaviour in a cell population, which overcomes some of these limitations. This simulation no longer relies on subculturing and performs a more detailed analysis of the available data. Most importantly, it establishes a framework for the development of a model with realistic copy number distribution.

3.2 Simulation Design

The previous models operated in the manner discussed in Chapter 1, modelling the population as a computationally expensive priority queue, in which cells were sorted by their due division time. Dimer-containing cells were penalised with a longer growth period such that they would become less numerous than dimer-free cells over time. As the cell population grew, the simulation would require more memory for storage and take longer to perform each cell division operation due to the nature of the priority queue. A necessary solution was to ‘subculture’ the population at random to reduce its size when it reached the limit. This was time-consuming itself, and data recorded immediately after subculturing was less representative than data recorded immediately before subculturing. Thus, the state of the population was recorded at its maximum and the sample times were averaged across multiple runs, which was potentially misleading.

3.2.1 Chemostatic Environment

The alternative to subculturing was simulating a chemostatic environment, in which the cell count is maintained at a constant level by removing cells whenever the population exceeds a certain number. A strict chemostat would remove a cell every time a cell was added to the population, that is, at every cell division. A more realistic chemostat would have a calculated probability of removing a cell, which would increase with overpopulation and decrease with underpopulation. An attempt was made to implement this in the

reconstructed model (MEng thesis, Field, 2006). When each cell was given a probability of survival after cell division, adjusted such that the population would tend back towards a mean size, it was found that dimer-only cells very quickly came to dominate the population. This was because every cell was treated equally after cell division; monomer-only cells lost their advantage because faster growth did not contribute towards their survival and the division time penalty imposed on dimer-containing cells did not hinder their survival.

A more natural way to implement a chemostatic environment would have involved a radical overhaul of the model itself. Rather than stabilising the system with a penalty to the growth time of dimer-containing cells, such cells could instead suffer from a fitness penalty that would directly increase their probability of removal from the population. When a cell divides, each of the daughters could then be challenged to survive based on their plasmid content; monomer-only cells would have a survival probability of 1 and others would suffer a small penalty per dimer they contained. The population could be maintained at a constant number by scaling the fitness according to the population count, as in the more realistic chemostat model above. The principle advantage of this solution is that the cell population would not require ordering, with the program passing over each cell in turn for the entire population, and then repeating from the start again with the survivors of that generation. Whilst this would reduce computation time, it is not an accurate portrayal of the biology of the system (it would be equivalent to plasmid dimers producing a toxin) and the growth time penalty cannot be directly translated into a fitness penalty. Further, if future development of the model were to include an algorithm for variable copy number, it would surely be time-dependent and thus incompatible with this time-independent simulation.

It was fortunately possible, however, to produce a functioning chemostatic environment for the model reported here; it removes a cell at random whenever the population exceeds a maximum size, as per a strict chemostat. However, at this point, the way in which the data were being stored and accessed within the program itself were the cause of further difficulty.

3.2.2 Data Structure

The cell population was modelled by the simulation as an abstract data type known as a priority queue. Such a queue sorts its contents according to a priority assigned to each element, in this case the time at which a given cell is due to divide, and allows only the topmost element to be accessed or removed from the queue. A priority queue can be implemented either as a sorted list, in which the data storage structure determines the performance of operations on the queue, or as an unsorted list, in which the data storage structure does not affect performance, but the topmost element must be found each time.

	Sorted List (Heap)	Unsorted List	Sorted List (Linear)
Insertion	$\mathcal{O}(\log n)$	$\mathcal{O}(1)$	$\mathcal{O}(\log n)$
Access	$\mathcal{O}(1)$	$\mathcal{O}(n)$	$\mathcal{O}(1)$
Removal	$\mathcal{O}(\log n)$	$\mathcal{O}(1)$	$\mathcal{O}(1)$

Table 3.1: Performance of different priority queue implementations; the \mathcal{O} notation describes the worst case use of computing resources required for each operation as a function of the population size n .

Of the many possible data storage structures for a sorted list implementation, a binary heap is the most common. In a heap, each element is considered below one other element and above two other elements, where the priority of each element determines the hierarchy. In the programming language used for this simulation (C++, see Chapter 2), it is not intended that a random element might be removed from a binary heap, so the standard implementation performs badly for this operation. The alternative structure was a linear list, in which the topmost element is readily accessible and removal of a random element maintains the ordering of the elements. Inserting an element requires that its correct place in the list be found, achieved fastest by using a modified binary search algorithm.

A summary of performance for the key operations can be seen in Table 3.1. The sorted list, stored as a linear list, has the best achievable performance for the purposes of this work, but the overall complexity of the simulation is still high, at $\mathcal{O}(n \log n)$, where n is the size of the cell population. To obtain results in reasonable time, the population size is thus somewhat restricted.

3.3 Simulation Operation

As described above, the cell population is modelled as a non-idealised priority queue. Each element in the queue represents an individual cell, about which only pertinent information is stored: the number of monomers it contains, the number of dimers it contains and the time at which it is due to divide. The starting cell population and the simulation parameters, such as plasmid copy number, are determined by two files that are loaded at the start of the program. Table 3.2 summarises the simulation parameters with typical values. The default starting population is a single cell at full copy number that has undergone a single dimer formation event, such that it contains 38 monomers and 1 dimer.

The simulation proceeds by performing a series of operations on the cell at the top of the queue, that is, the cell that is next due to divide. First, the time recorded by the simulation is set to equal the time at which this cell divides, such that the simulation jumps through time in discrete, variable-size periods. The plasmids in the cell are divided randomly between two new cells, representing the daughters that are the result of division.

Parameter	Typical Value	Description
n_{pop}	10000	Maximum population size
n_{copy}	40	Plasmid copy number
P_{res}	0%	Dimer resolution probability
R_{rec}	3.8×10^{-3}	Dimer formation rate
t_{div}	30 minutes	Mean cell division time
t_{dim}	0.15 minutes	Division time penalty per dimer
T_{run}	7200 minutes	Total simulation time
t_{sample}	60 minutes	Sample period
n_{run}	1000	Number of repeats

Table 3.2: Population model parameters with typical values drawn from [Summers et al. \(1993\)](#) or chosen for computational convenience.

Then, provided that the daughter has inherited at least one plasmid, its plasmids are replicated at random back to full copy number, acknowledging that dimers are twice as likely to replicate as monomers. Daughter cells that inherit no plasmids are ‘killed’ by the simulation and removed from the queue. The surviving daughter cells are then tested for dimer formation and resolution events. The probability of a dimer formation event occurring, P_{rec} , is based on the plasmid content of the cell, assuming that the likelihood of two such events is negligible, as follows:

$$P_{rec} = R_{rec} \times \frac{M \times (M - 1)}{(M + D) \times (M + D - 1)}$$

In the other direction, each dimer has a probability, equal to the parameter P_{res} , of being resolved into two monomers via XerCD recombination. The time at which the daughters are due to divide is calculated based on their plasmid contents; it equals the base division time for a monomer-only cell plus a small penalty per dimer. This is slightly altered from the calculation used in the original model, in order to give true linearity. Finally, again provided that the specific daughter contains plasmids, the daughter cells are put back into the queue in the appropriate places and the parent cell removed. Should the population exceed its maximum at this point, a cell is randomly removed from the queue. The cell that is then at the top of the queue is subject to the same series of operations, and so on, until the simulation time limit is reached.

At regular intervals, the queue undergoes a census to determine what proportion of the population contains a given number of monomers, dimers and total plasmids. When the simulation time limit is reached, it resets and is repeated until the appropriate number of repeats have been completed. At the very end, the accumulated data is analysed and output to files for later use. As well as saving the results of every census, the means and standard deviations of these results are calculated across all the repeats. The stability indices of each census are also calculated and similar statistics are produced.

3.4 Testing

Initial tests of the simulation raised two concerns: that there might be significant behaviour occurring faster than the sampling frequency could capture, affecting the subsequent analysis of results, and that the available runtime would limit the size of the simulated population, which might affect the observed behaviour of the system.

3.4.1 Variable Division Time

Figure 3.1 shows the average percentage of monomer-only and dimer-only cells in the population recorded by the simulation at different sampling frequencies. In Figure 3.1(a), step-like behaviour can be observed in the first half of the simulation. This might be a natural result of deriving the population from a single cell; synchronised division causes synchronised movement towards steady-state. However, as the sampling frequency for this simulation is half the frequency of division for monomer-only cells, it is possible that the behaviour is due to aliasing, that is, the poor sampling frequency distorting observations. To test this hypothesis, the simulation was repeated with a sampling period of 10 minutes rather than 60 minutes. In Figure 3.1(b), the extremes of the resultant, rapidly-changing curves roughly correspond to the curves in Figure 3.1(a), rather than the more representative midpoints. The overall zig-zag behaviour is undesirable when trying to observe steady-state trends.

Further investigation revealed that it was indeed synchronisation of cell division times in the initial stages of the simulation that was causing the behaviour. To correct this issue, the solution was to introduce variation into the division times of individual cells, whilst maintaining the division time penalty associated with dimers. A number of studies report that the interdivision times of cells that are initially synchronised are normally distributed, with a coefficient of variance in the range of 12% to 22% (Koppes et al., 1980; Schaechter et al., 1962; Shehata and Marr, 1970). For this simulation, a normal distribution with a coefficient of variance of 12.5% was selected. In Figure 3.2(a), the steady-state levels of monomer-only and dimer-only cells with a variable division time are compared to those of cells with a fixed division time. The simulation was retested with sample periods of 60 minutes and 10 minutes and as seen in Figure 3.2(b), there is no difference in behaviour between the two; they both show a smooth tendency towards the steady-state.

3.4.2 Runtime Optimisation

The parameters with the most significant effect on the runtime of the simulation are the simulation time limit, the number of repeats and the population size. The simulation time limit must be in the region of 6,000 to 8,000 minutes for the population to reach steady-

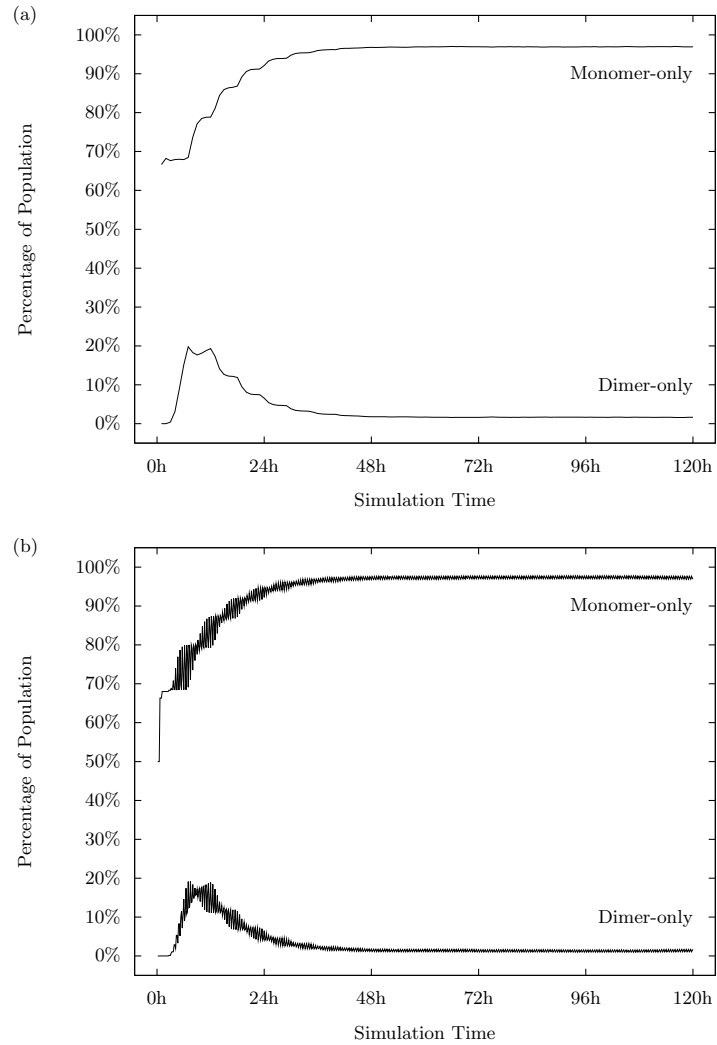


Figure 3.1: Comparison of simulation behaviour with different sampling frequencies; (a) 60 minute sample period, (b) 10 minute sample period.

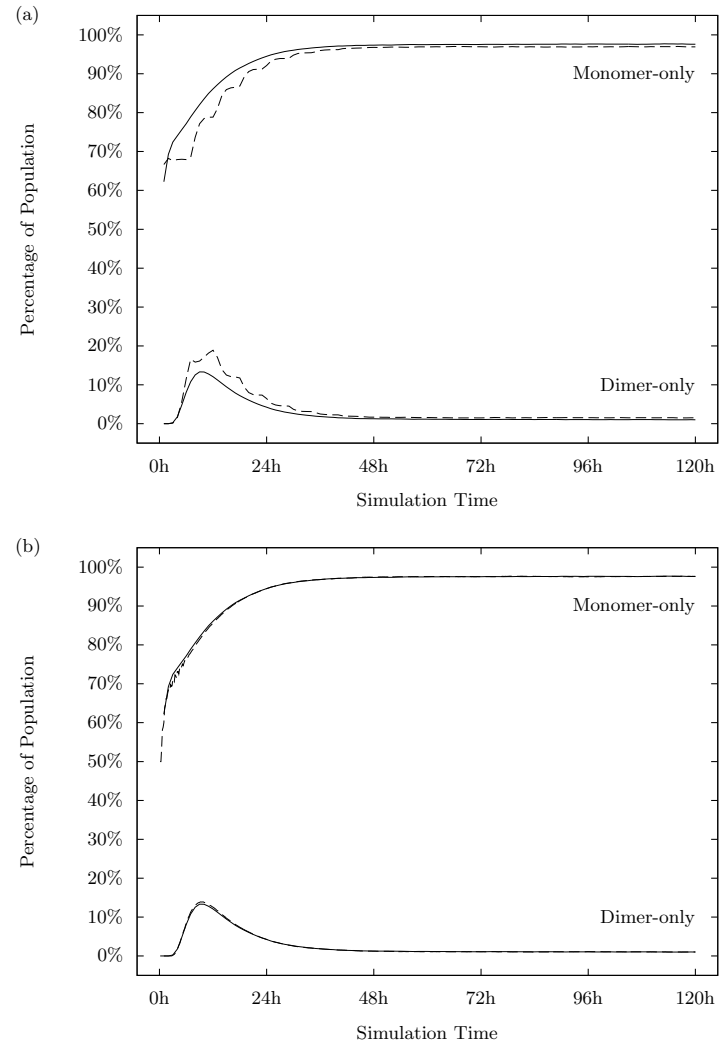


Figure 3.2: Comparative simulation behaviour; (a) 60 minute sample period with (—) and without (---) variable division time, (b) variable division time with 60 minute sample period (—) and 10 minute sample period (---)

state, so cannot be altered to significantly reduce the runtime. The number of repeats will determine how smooth the final data is, and it is important to ensure that enough are performed to include rare events. The population size determines the resolution of each data point, that is, the significance of a single cell within the population. As the runtime for a single repeat is of complexity $\mathcal{O}(n \log n)$, $10r$ repeats of population size n will run faster than r repeats of population size $10n$.

One concern with the simulation was that the restricted population size might affect the behaviour of the system itself. To test whether this was the case, the simulation was run 100 times with population sizes of 1,000, 2,000, 5,000 and 10,000. The results are shown in Figure 3.3. The population size has not affected the steady-state of the system when the results are averaged, but the standard deviation is larger for a smaller population, likely due to the poor resolution available. The variance of a system with a given set of parameters can only be useful in comparison to a system with a different set of parameters, rather than providing insight into behaviour *in vivo*.

3.5 Results

3.5.1 General Behaviour

The behaviour of the simulation with default parameter values as listed in Table 3.2, was shown to be similar to that of Summers et al. (1993). The distribution of dimers in the population at various time points can be seen in Figure 3.4. As reported by Summers et al. (1993), the system moves towards a steady state wherein the majority of dimers inhabit dimer-only cells. The percentage of monomer-only and dimer-only cells in the population are approximately 97.6% and 1.1% respectively at steady-state. This is more and less, respectively, than seen in the model of Summers et al. (1993). The most likely reason for the difference is the inclusion of variation in cell division time; it is likely that Summers et al. (1993) were unable to test their simulation sufficiently to realise that a fixed cell division time would create sample bias, as discussed here previously.

3.5.2 Dimer Formation Rate

The rate at which homologous recombination converts monomer pairs into dimers was varied to observe the effects on the behaviour of the system. Values for R_{rec} between 0.1 times and 10 times the default rate of 3.8×10^{-3} were used. It can be seen in Figure 3.5 that the steady-state level of monomer-only and dimer-only cells is dependent on the dimer formation rate. Further, Figure 3.6 shows that the stability index of the system at steady-state is logarithmically related to the dimer formation rate.

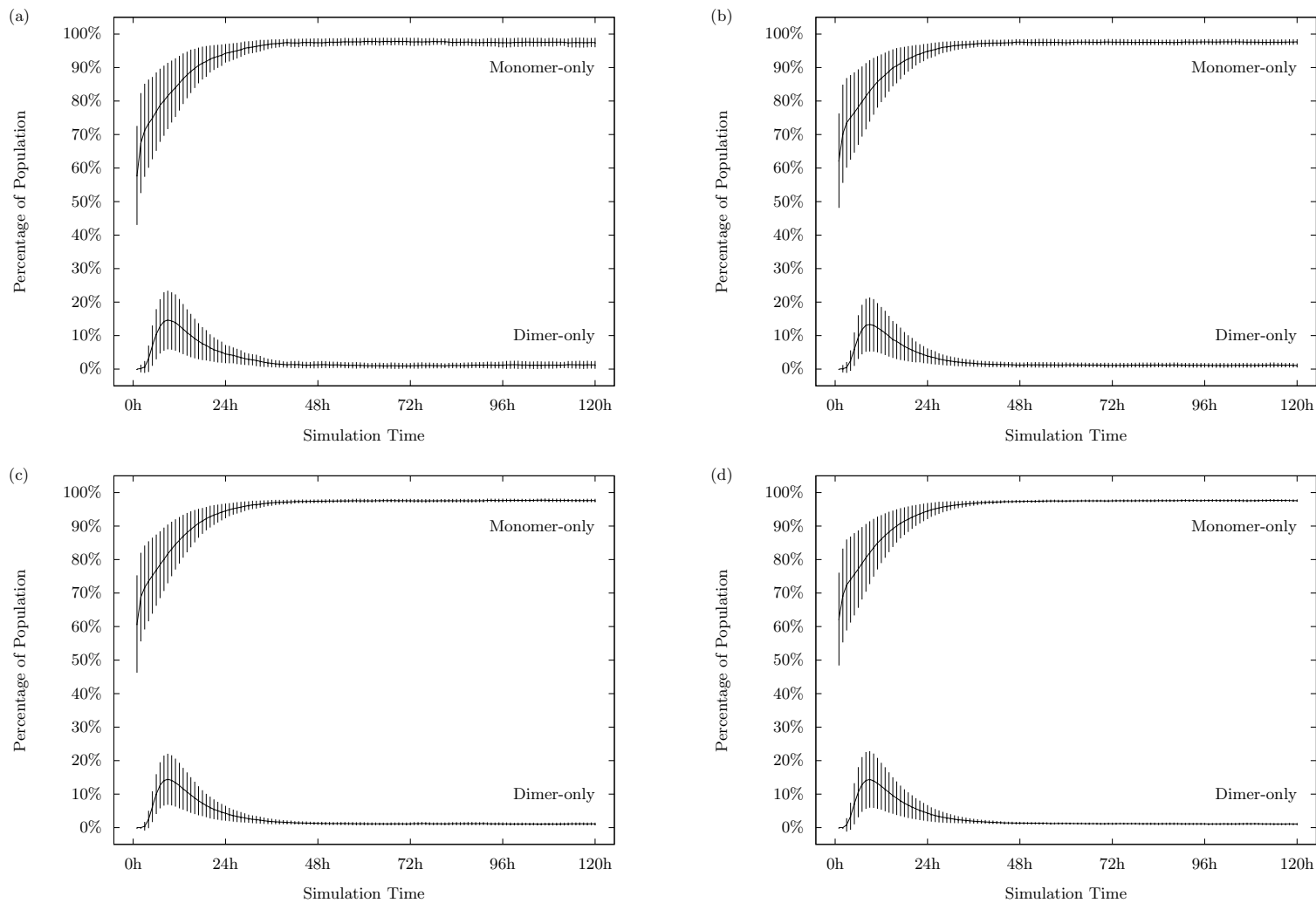


Figure 3.3: Comparison of simulation behaviour with different population sizes; (a) 1,000 cells, (b) 2,000 cells, (c) 5,000 cells, (d) 10,000 cells. Vertical bars indicate the standard deviation of each sample across the 100 runs.

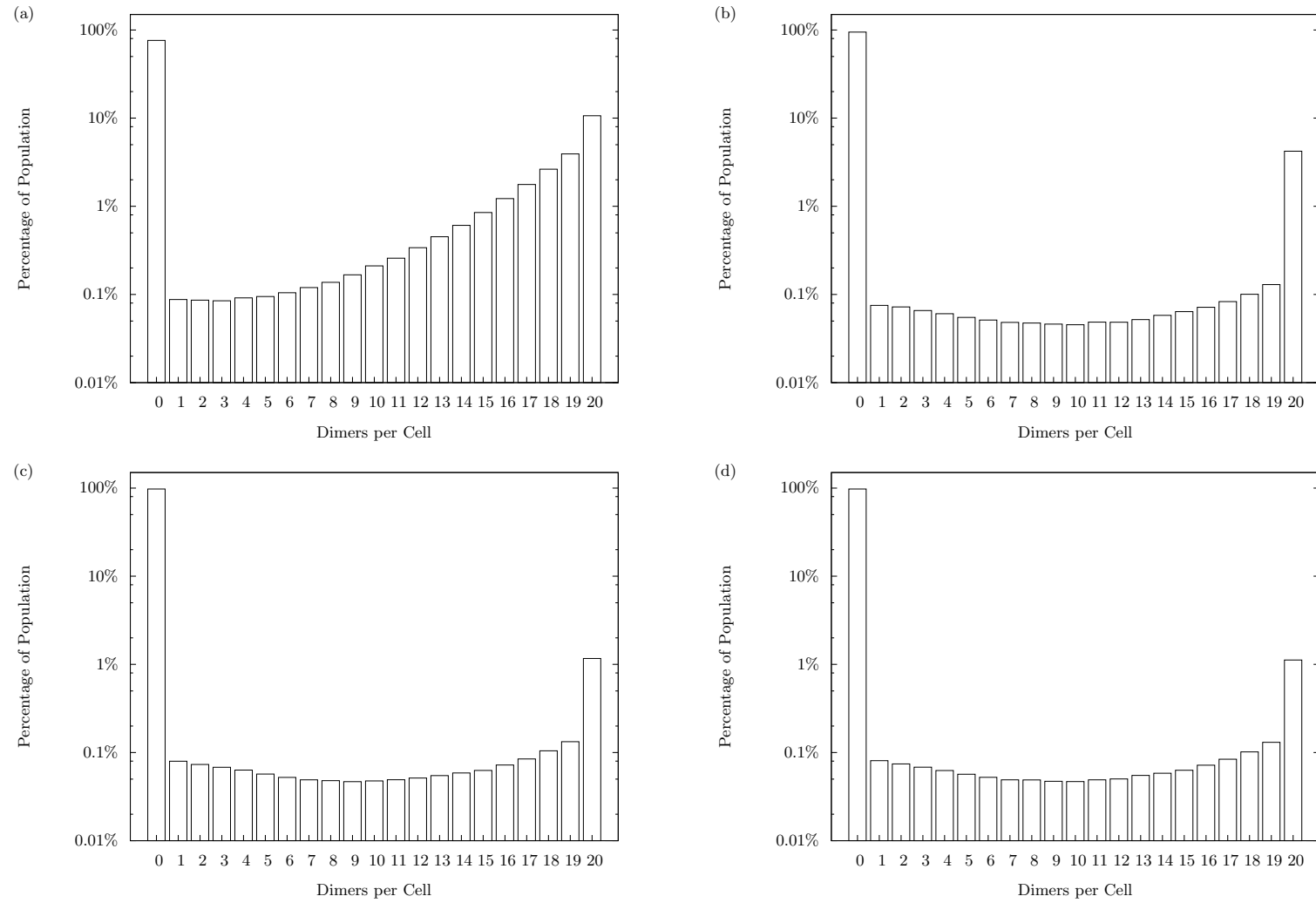


Figure 3.4: Distribution of plasmid dimers in the population at different time points; (a) 6 hours, (b) 24 hours, (c) 60 hours, (d) 120 hours.

It is also worth noting that the system takes longer to reach steady-state with a higher dimer formation rate. This is more evident in Figure 3.6, where the system reaches steady-state faster when the final stability is closer to the peak instability at around 6 hours.

3.5.3 Dimer Growth Penalty

The division time penalty imposed on cells containing dimers was varied to observe the effects on the behaviour of the system. Values for t_{dim} between 0.03 minutes and 0.75 minutes per dimer were used, equal to 0.1% and 2.5% of the 30 minute generation time respectively. This corresponds to a division time penalty of between 2% and 50% for dimer-only cells compared to monomer-only cells. It can be seen in Figure 3.7 that the more that dimer-containing cells are penalised, the faster the system reaches the steady-state.

The steady-state level of monomer-only cells increases with a greater penalty and the level of dimer-only cells correspondingly decreases. This is reflected in the steady-state stability indices for the different penalties, shown in Figure 3.8. As for the dimer formation rate, the relationship between stability index and dimer-associated growth penalty is logarithmic, though it is inversely so in this case.

3.5.4 Dimer Resolution Rate

Plasmid ColE1 contains a site-specific recombination site that enables dimer resolution via the XerCD recombinase (see Section 1.5.1). The rate at which this occurs is unknown, but is thought to be slow, necessitating the intervention of the Rcd checkpoint. Dimer resolution was incorporated into the model by giving each dimer a chance of resolution prior to cell division. To observe how effective the resolution system needs to be to have a significant impact on the stability of the system, resolution probabilities between 1% and 20% per dimer per generation, were tested in a system with a the default dimer formation rate of 3.8×10^{-3} .

Figure 3.9 shows that dimer resolution has very little effect on the steady-state level of monomer-only cells in the population until it reaches at least 20%. However, the peak level of dimer-only cells is immediately reduced with even a 1% chance of resolution per dimer. The steady-state level of dimer-only cells is reduced only slightly as the dimer resolution chance increases.

Looking at the effects on the overall stability of the plasmid in Figure 3.10, a 1% probability of dimer resolution per dimer per generation, increases the stability from 2.2×10^{-8} to 1.7×10^{-8} . This is not as effective as increasing the plasmid copy number by 1, which would increase the stability to 1.1×10^{-8} . For every 3% increase in the dimer resolution rate, plasmid stability increases by approximately the same amount as it would do by

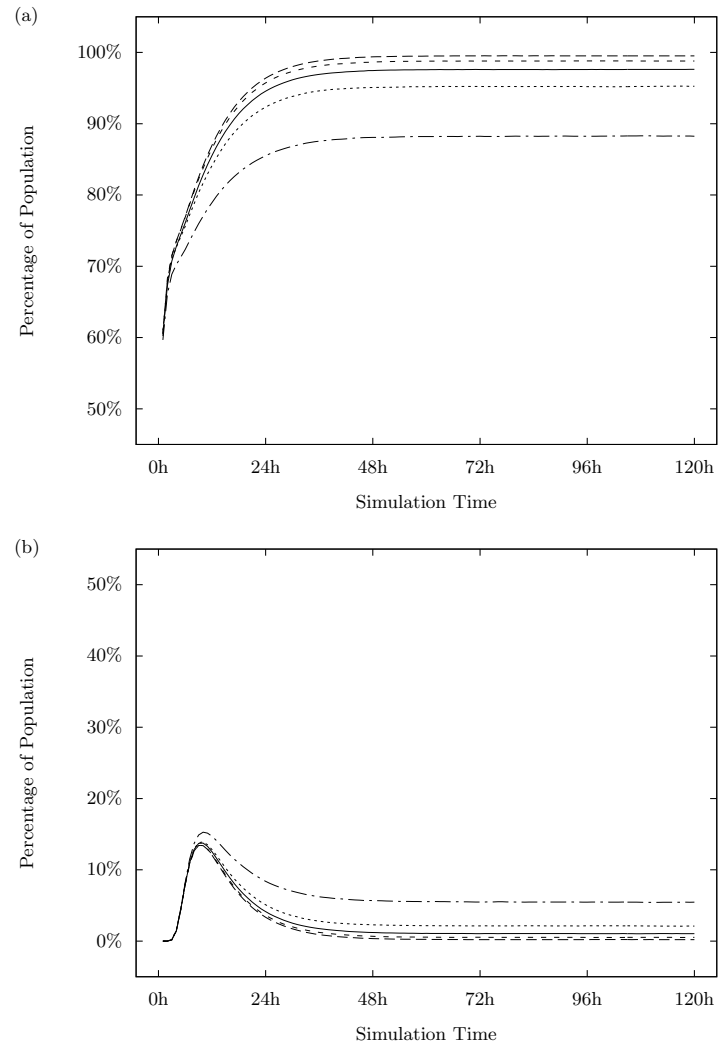


Figure 3.5: Simulation behaviour with different dimer formation rates; (a) Monomer-only cells, (b) Dimer-only cells.

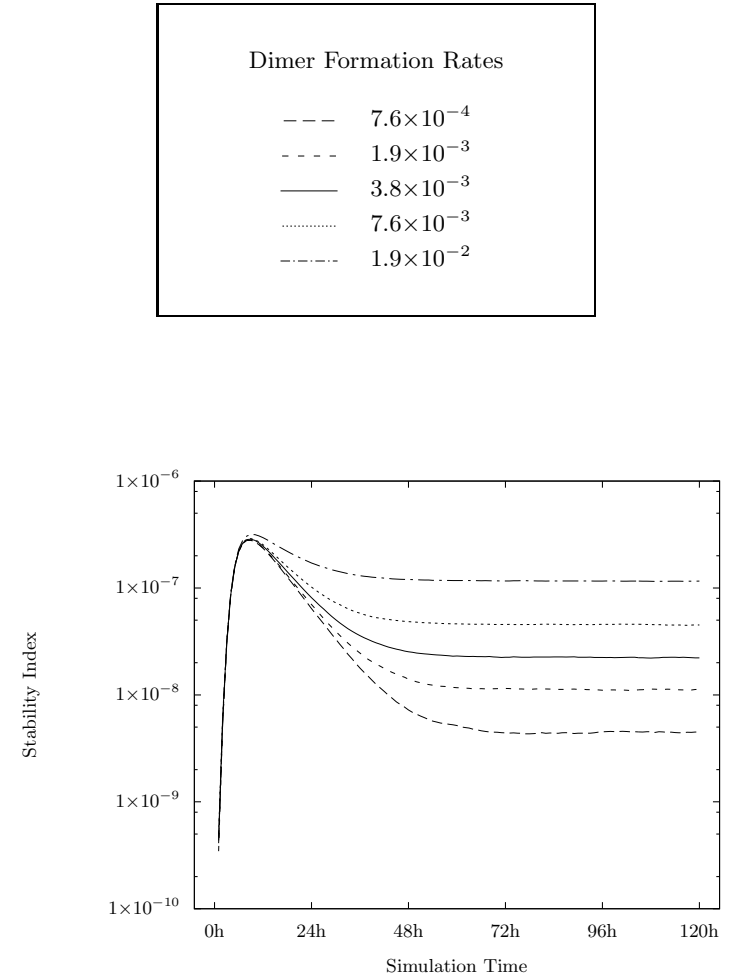


Figure 3.6: Plasmid stability with different dimer formation rates.

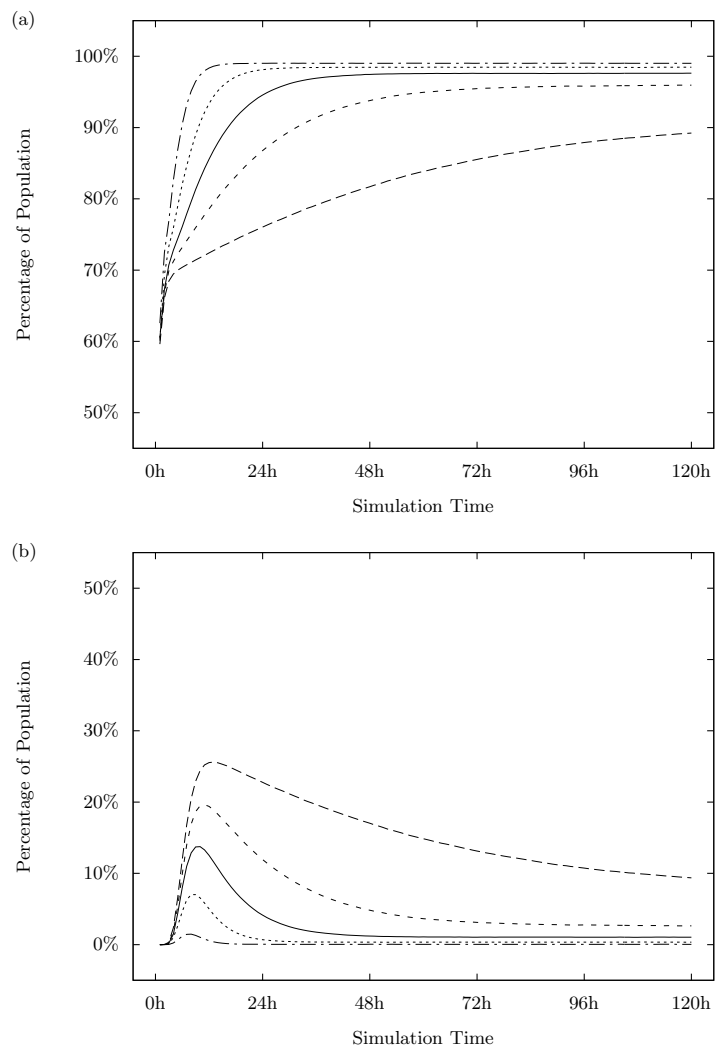


Figure 3.7: Simulation behaviour with different division time penalties for dimers; (a) Monomer-only cells, (b) Dimer-only cells.

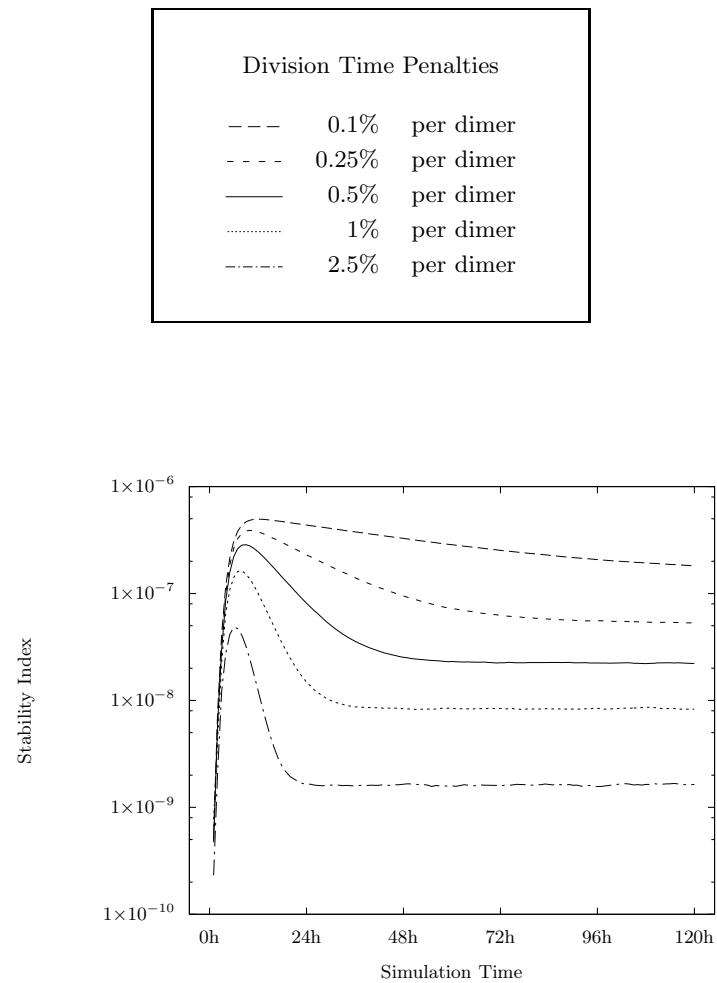


Figure 3.8: Plasmid stability with different division time penalties for dimers.

increasing the copy number by 1, which is 2.5% of the monomer copy number. Whether this is effective or not depends on the disadvantages associated with increasing plasmid copy number. It is also worth noting that even with a 20% probability of dimer resolution per dimer per generation, the system does not reach steady-state any faster, although peak instability is reduced by the same amount as the eventual steady-state stability.

3.5.5 Initial Conditions

It was important to verify that the system reaches the same steady-state from different initial conditions. The simulation was run with default parameter values and a low resolution rate of 1% per dimer per generation. The starting conditions were: a cell containing 20 monomers and 10 dimers, a cell containing 40 monomers and 0 dimers and a cell containing 1 monomer and 20 dimers.

In each case, given enough time, the steady state reached was the same as for a system begun with a cell containing 38 monomers and 1 dimer (data not shown). Although the dimer-only cells dominated the population early in the run when the initial cells were dimer-heavy, the small 1% resolution chance per dimer per generation, and the division time penalty associated with dimers, eventually caused them to be overtaken by the monomer-only cells. For the population from an initial cell containing no dimers, the system tended smoothly towards steady-state as dimer formation eventually occurred in each simulated repeat.

3.6 Conclusions

The simulation of Summers et al. (1993) was successfully recreated with an underlying chemostatic model of population control, rather than subculturing. Additionally, normally distributed cell division time was introduced to represent bacterial growth more accurately, and to prevent the bias caused by synchronised cell division.

The results of the simulation described in this chapter broadly agree with the conclusions of Summers et al. (1993) and the reconstructed simulation work discussed in Section 1.8:

- Imposing even a relatively small division time penalty on dimer-containing cells contains the dimer catastrophe and stabilises the system.
- The majority of dimers in the population are hosted by dimer-only cells.
- The plasmid stability index was logarithmically related to the dimer formation rate and the division time penalty associated with dimer-containing cells.

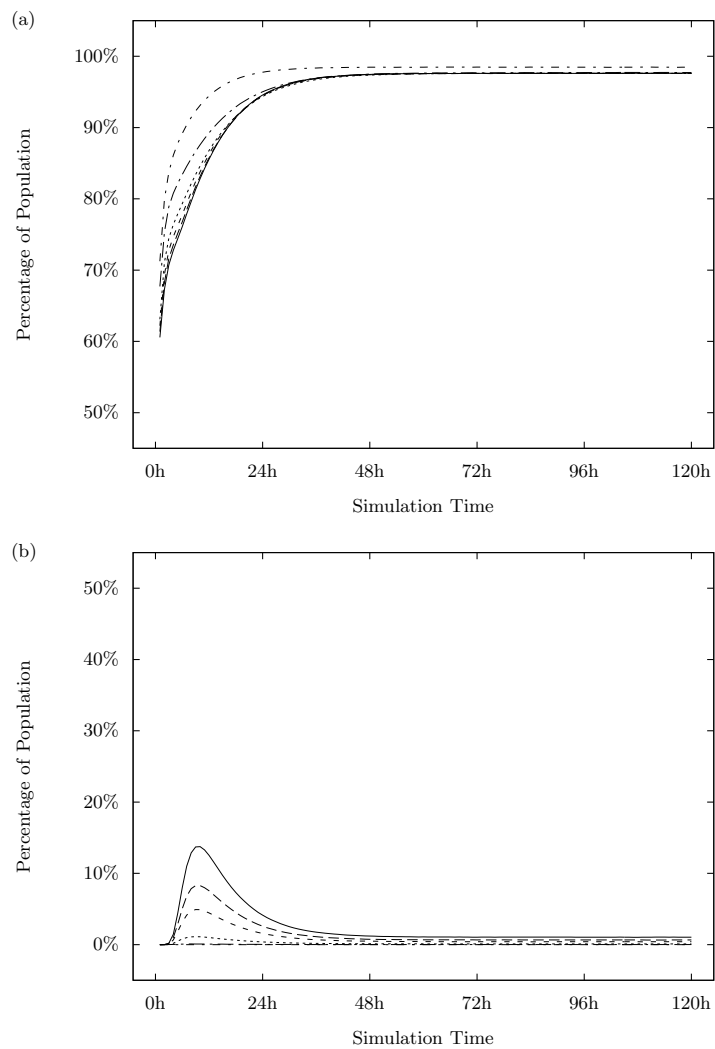


Figure 3.9: Simulation behaviour with different dimer resolution probabilities; (a) Monomer-only cells, (b) Dimer-only cells.

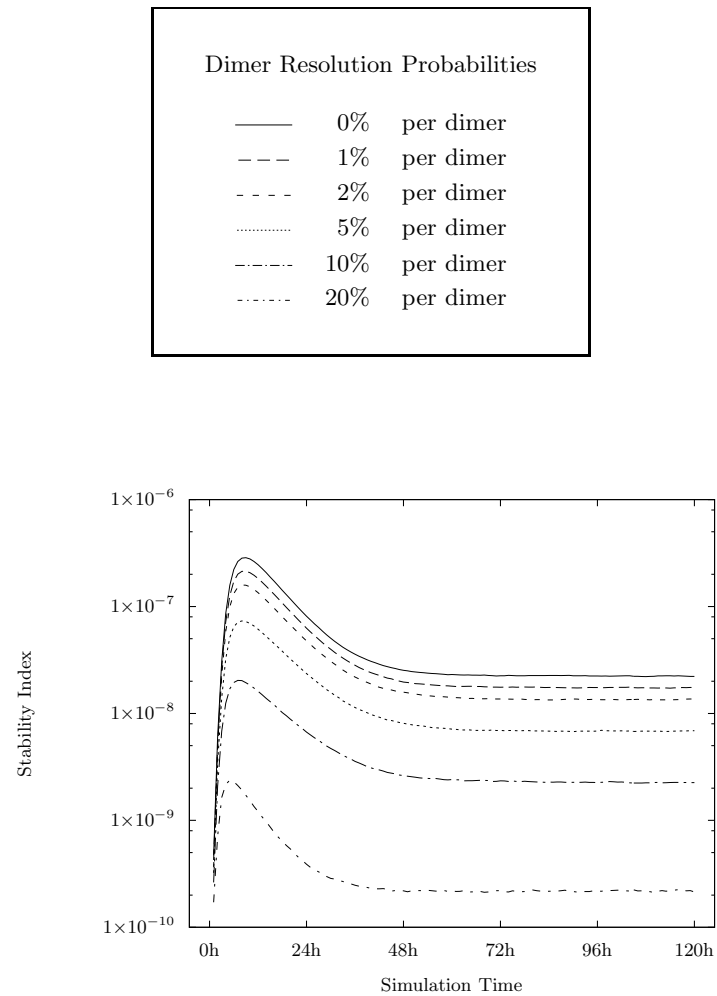


Figure 3.10: Plasmid stability with different dimer resolution probabilities.

- The time taken to reach steady-state was determined by the dimer formation rate and the division time penalty associated with dimer-containing cells.
- A dimer resolution chance of around 3% per dimer per generation has the same effect on plasmid stability as copy number increase of 1.
- The system was able to reach the same steady-state from a variety of initial conditions, indicating its reliability.

This simulation is useful for the study of an ideal system. However its plasmid replication model is inadequate. As the plasmid copy number is always restored to full between cell divisions, an accurate parametric study of plasmid stability cannot be made and it is difficult to assess the efficiency of the dimer resolution system.

In the following chapter, a significantly more accurate plasmid replication model is developed for inclusion into the overall cell population model. Variable copy number will change the relationship between the stabilities of dimer-only and monomer-only cells and thus how the different simulation parameters affect the system.

Modelling ColE1 Copy Number Distribution

4.1 Introduction

The number of plasmids in the cell immediately prior to division is dependent on how many it inherited at birth and the number of replication events during the subsequent growth period. The control system that regulates these events is stochastic in nature, involving just a few hundred molecules in the cell and creating a degree of variability in plasmid replication that is not currently captured in the cell population model of Chapter 3. In order to more accurately simulate plasmid behaviour, some attempt must be made to include this statistical variation.

As detailed in Section 1.3.2, ColE1 replication is initiated by the transcription and cleavage of a non-coding RNA, RNAII. As transcription proceeds, a second small RNA, RNAI, can interfere with the folding of RNAII, preventing replication initiation. Both RNAI and RNAII are transcribed constitutively from the plasmid, along with a protein, Rom, that strengthens the RNAI-RNAII interaction. Thus, the frequency of replication is dependent on the transcription rates of the two RNAs, their stabilities and their interaction.

There have been a number of mathematical models of ColE1 plasmid replication (see Section 1.7). The most useful models to date are those of Paulsson and Ehrenberg, in which the stochastic nature of plasmid replication is considered. The complex process is distilled down to a number of key molecules, parameters and assumptions, with a view to analysing the effects of those key parameters on the control of the replication system. Using their models as a basis, the work described in this chapter aims to simulate plasmid replication behaviour for incorporation into the cell population model. The replication model will include dimer formation, dimer resolution by XerCD, and copy number variation using realistic parameters drawn from *in vivo* work.

4.2 Model Design

The model functions as a generalised Gillespie algorithm (Gillespie, 1977). The different numbers of important molecules are counted and comprise the state of the system. There are a number of possible events that can change the state of the system, each with an associated rate of reaction which can be constant or based on the current state. It is assumed that there are no local effects, that is, all of the molecules under consideration are diffused randomly throughout the cell.

4.2.1 System State

The key molecules in this model are M , the number of plasmid monomers in the cell, D , the number of plasmid dimers in the cell and I , the number of replication-inhibiting RNAI molecules in the cell. The state of an individual cell is summarised by these three parameters and referred to in this work by the shorthand (M, D, I) . It is assumed that there is always a saturating concentration of Rom protein (Keasling and Palsson, 1989) and that higher-order multimers are sufficiently rare to ignore. The simulation time and cell volume are also tracked. Cells are assumed to be growing exponentially such that cell volume, $v(t)$, doubles between birth and division (Schaechter et al., 1962).

4.2.2 Transitions

The transition events between states and their reaction rates are shown in Figure 4.1. They are detailed as follows:

RNAI Transcription

Increases the number of RNAI molecules in the cell at a rate proportional to the number of plasmid origins.

$$k_1 \times (M + 2D)$$

RNAI Degradation

Decreases the number of RNAI molecules in the cell at a rate proportional to the number of existing RNAI molecules.

$$e_1 \times I$$

RNAII Transcription

Triggers a potential plasmid replication event at a rate proportional to the number of plasmid origins. It is assumed that the eventual fate of all RNAII primers is to form a duplex with an RNAI molecule, whether they initiate replication or not (Tomizawa, 1986).

$$k_2 \times (M + 2D)$$

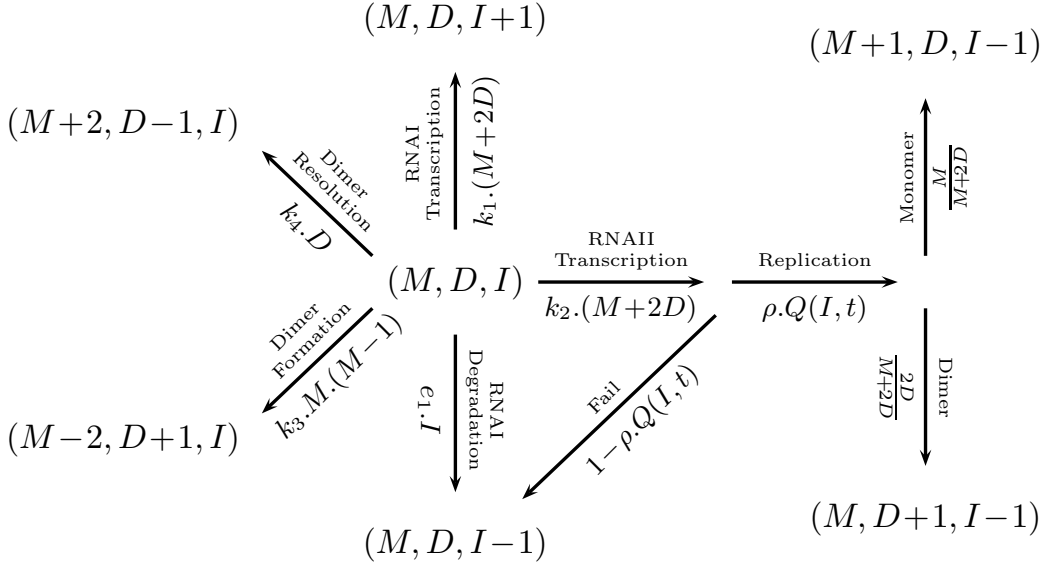


Figure 4.1: State transitions and their rates in the plasmid replication model.

Replication

After RNAII transcription, the probability of successfully initiating replication is $\rho \times Q(I, t)$. ρ is a factor to account for the fact that not all uninhibited RNAII transcripts are able to initiate replication. The function $Q(I, t)$ depends on the nature of the RNAI-RNAII inhibition process (Ehrenberg, 1996). If it is dominated by a single rate-limiting step then the function is hyperbolic:

$$Q(I, t) = \frac{1}{1 + \frac{I}{K_i \times v(t)}}$$

This is close to the case of plasmid R1 (Nordström, 1985). If instead it is governed by multiple rate-limiting steps then the function is exponential:

$$Q(I, t) = e^{-\frac{I}{K_i \times v(t)}}$$

This is an approximation which becomes more accurate as the true number of rate-limiting steps tends to infinity. There are at least 251 steps of nucleotide addition in the transcription of RNAII during which it is susceptible to RNAI interference, so this model assumes exponential inhibition.

If replication is initiated, a random origin is chosen to replicate such that the number of monomers is increased by one, with probability $\frac{M}{M+2D}$, or the number of dimers is increased by one, with probability $\frac{2D}{M+2D}$. Whether the plasmid replication event is successful or not, the number of RNAI molecules is decreased due to RNAI-RNAII duplex formation.

Dimer Formation

Homologous recombination converts two monomers into a dimer with a rate proportional to the number of possible monomer pairs.

$$k_3 \times M \times (M - 1)$$

Dimer Resolution

Converts a dimer into two monomers with a rate proportional to the number of dimers. Whilst the dimer resolution rate may be limited *in vivo* by the availability of XerCD or DNA replication disrupting the complex, these factors have been ignored for simplicity.

$$k_4 \times D$$

4.2.3 Algorithm

The algorithm for this model proceeds as follows:

1. Calculate the reaction rates (r_1 , r_2 , etc..) of the possible events based on the current system state.
2. Choose a time through which the simulation will discretely ‘jump’, based on an exponentially distributed random number with parameter r_T (the sum of all reaction rates).
3. Choose an event to occur with a uniformly distributed random number, where each event has a probability of occurring equal to $\frac{r_n}{r_T}$.
4. Change the system state based on the chosen event and jump time.
5. Return to step 1 unless the algorithm has reached an ending condition, such as a time or volume limit.

Thus the simulation proceeds by skipping through the continuous activity associated with a single event, such as the base-by-base transcription of RNA, and looking at the immediate results. The time taken to do this is determined by the rates of all possible reactions in the system, and the probability of a specific event occurring is calculated relative to all other events. Reactions with fast rates contribute little to the total of all reaction rates, so the average jump time is dominated by the slowest process in the system. However, such fast reactions are the most likely to occur and constitute the bulk of events in the system.

Parameter	Value	Description
$v(t)$	$0.45\text{-}0.9\ \mu\text{m}^3$	Cell volume
k_h	$0.0234\ \text{min}^{-1}$	Exponential growth rate of the cell
k_1	$4.56\ \text{min}^{-1}$	Rate of RNAI transcription initiation
k_2	$0.80\ \text{min}^{-1}$	Rate of RNAII transcription initiation
k_3	Varies	Rate of dimer formation by homologous recombination
k_4	Varies	Dimer resolution rate
e_1	$0.93\ \text{min}^{-1}$	Rate of RNAI degradation
ρ	0.5	Probability that a mature RNAII primer initiates replication
K_i	4.44×10^{19}	RNAI-RNAII interaction constant

Table 4.1: Replication model parameters with typical values drawn from various sources (see text).

4.3 Parameters

The behaviour of the model depends on several different parameters, summarised in Table 4.1. Previous models have investigated the effects of varying these parameters (Paulsson et al., 1998) or attempted to discern them *in vivo* (Brenner and Tomizawa, 1991; Lin-Chao and Bremer, 1986, 1987). Paulsson and Ehrenberg (2001) suggest a set of parameter values normalised to the growth rate of the cell such that the average plasmid copy number is the same for all growth rates, but note that those measured *vivo* are likely to vary considerably. Here, parameters have been normalised, though the average plasmid copy number is known to be dependent on growth rate (Lin-Chao and Bremer, 1986).

4.3.1 Cell Parameters

The generation time for *E. coli* depends on its growth conditions, and can be as fast as 24 minutes. For compatibility with the population models discussed in Section 1.7 and the culturing methods discussed in Chapter 2, the generation time here is 30 minutes. Assuming that the rate of cell growth is proportional to its length, the growth rate is thus:

$$2^{\frac{1}{30}} - 1 = 0.0234\ \text{min}^{-1}$$

The cell volume of *E. coli*, $v(t)$, also depends its growth conditions. A widely quoted average cell size in exponential phase is between $0.6\ \mu\text{m}^3$ and $0.7\ \mu\text{m}^3$ (Kubitschek, 1990). With the distribution of cell sizes in exponential phase for a culture growing asynchronously, the average cell is approximately 69% of maximum size. This model therefore uses a convenient starting volume of $0.45\ \mu\text{m}^3$ and a pre-division volume of $0.9\ \mu\text{m}^3$.

4.3.2 Plasmid Parameters

The plasmid replication behaviour is dependent on the parameters discussed below, and any credible model must aim to generate a plasmid copy number which is consistent with experimental observation. ColE1 has a copy number per chromosome of around 10 (Som and Tomizawa, 1983), and cells growing rapidly in exponential phase are known to have multiple chromosomes (Cooper and Helmstetter, 1968). An average copy number of anywhere between 15 and 40 would be acceptable for this model.

The transcription rates of RNAI and RNAII are reported to be in a ratio between 2.5:1 and 5:1 (Lin-Chao and Bremer, 1987). The same work estimates the absolute rates to be 2 min^{-1} for RNAI and 0.35 min^{-1} for RNAII, strangely outside of the reported ratio range. Brenner and Tomizawa (1991) measured the absolute number of RNAI and RNAII molecules *in vivo*, but derived rate estimates inconsistent with the reported ratios. In other plasmid replication models, the available parameter space has been searched to obtain these rates, or they have been set relative to other parameters (Brendel and Perelson, 1993; Paulsson et al., 1998). Thus, it is somewhat a matter of preference, and here the rates of Lin-Chao and Bremer (1987) have been normalised to the chosen growth rate.

The degradation rate of RNAI is the only other pertinent parameter, as it is assumed that RNAII will always be degraded after its transcription, whether replication was primed or not. Brenner and Tomizawa (1991) estimate the half-life of RNAI to be 2 minutes, whereas Lin-Chao and Bremer (1986) estimate it at 0.55 minutes. These correspond to degradation rates of 0.35 min^{-1} and 1.26 min^{-1} respectively. This model uses the 2 minute estimate, normalised to the growth rate, as it is the more recent and generally favoured by other models (Brendel and Perelson, 1993; Paulsson et al., 1998).

4.3.3 Replication Parameters

The probability that a primer finally initiates replication, ρ , has been estimated to be 0.5 (Brendel and Perelson, 1993).

To determine the RNAI-RNAII interaction constant K_i , some derivation is required. The exponential term in the function $Q(I, t)$ is of the form:

$$\frac{-I}{K_i \times v(t)} \quad (4.1)$$

This is a rearranged version of equation 7 from Paulsson et al. (1998):

$$\frac{-R_i}{K_j} \quad (4.2)$$

Therein, R_i represents the intracellular concentration of RNAI and K_j is defined as:

$$\frac{1}{K_{ai} \times T_{tr}}$$

T_{tr} is the transcription time of the 110-360 base region of RNAII, in which it is vulnerable to RNAI inhibition. K_{ai} is the association rate constant between RNAI and RNAII. The transcription time for each of the 251 nucleotides and the association rate constant are known to vary considerably across the vulnerability window (Tomizawa, 1986), but for this analysis it is simpler to consider a constant transcription rate and an average association rate constant. Taking the transcription rate to be 50 nt s^{-1} (Hippel et al., 1984) and K_{ai} to $2.7 \times 10^6 \text{ mol}^{-1} \text{ s}^{-1}$ (Brenner and Tomizawa, 1991):

$$K_j = \frac{1}{250 \times 0.02 \times 2.7 \times 10^6} = 7.41 \times 10^{-8}$$

Including a term to convert cell volume to litres, R_i , is:

$$\frac{I}{1000 \times v(t) \times 6.022 \times 10^{23}}$$

Thus, using equations 4.1 and 4.2, K_i for this model is:

$$1000 \times 6.022 \times 10^{23} \times K_j = 4.44 \times 10^{19}$$

4.4 Results

The software implementation of this model requires specification of the initial cell state, (M, D, I) , the number of successive generations to simulate, and the number of repeats (each starting from the initial state). The state of the cell is recorded at the end of each generation for later statistical analysis. This allows for repeated simulation of a few generations to obtain copy number averages, or many generations to determine copy number distribution.

Each successive generation requires the selection of one of two daughter cells to be the next simulated. Their initial states can be determined in two ways: the plasmids (and inhibitors) could be divided equally, akin to a plasmid partitioning system, or they could be divided at random between the two potential cells, akin to multicopy plasmid distribution. The latter is preferable for later inclusion in the population model, but a comparison of the two is a useful exercise.

4.4.1 Average Copy Number Estimates

To determine the steady-state average plasmid copy number for monomer-only cells and dimer-only cells, the simulation was initialised with states $(1, 0, 0)$ and $(0, 1, 0)$ respectively. Both the partitioning and the random distribution feedback mechanisms were tested for each initial state. The simulation was run 100,000 times, through 25 cell generations in each run, with dimer formation and resolution rates set to 0.

For each combination, the system reached a steady-state average plasmid copy number in less than five generations; summarised in Table 4.2. Two interesting observations can be made. Firstly, the average dimer copy number is approximately 62.5% of the average monomer copy number for both feedback mechanisms, rather than 50%. This is consistent with previous observations *in vivo* (Chiang and Bremer, 1988) and the consequence that dimer-containing cells therefore contain more plasmid DNA may go some way to explaining the dimer-associated metabolic load observed in the work of the original model (Summers et al., 1993).

Secondly, there is a small difference in the average final state between the two different distribution methods. Counter-intuitively, the random distribution feedback mechanism results in 0.14 more monomers and 0.11 more dimers on average. However, this difference is small and although the standard deviation is also slightly higher, this demonstrates that partitioning is not required for good copy number control.

Also informative, the steady-state average inhibitor numbers, \bar{I} , are summarised in Table 4.3. For both monomers and dimers and both feedback mechanisms there are around 3.95 inhibitors for each plasmid origin.

To explain the increased average copy numbers under random distribution, consider that in a new cell, the ratio of plasmids to inhibitor molecules is unlikely to be at the mean. If the inhibitor to plasmid ratio was lower than average, there would be an increased likelihood of replication occurring, whereas if it were higher than average, there would be a decreased likelihood of replication occurring. The key difference between the two situations lies in the kinetics of RNAI-RNAII interaction, as defined by equation 4.1. As this interaction is exponential in form, the likelihood of replication in the lower than average situation is increased by *more* than it is decreased in the counter-situation. Overall, this results in increased replication with the random distribution feedback mechanism.

4.4.2 Plasmid Copy Number Distributions

In the population model, the stability of the plasmid will depend not on the average plasmid copy number, but the copy number distribution. Those cells with fewer plasmids are significantly more likely to suffer plasmid loss than the average, and strongly influence

	Partitioning	Random Distribution
Monomers	$\bar{M} = 29.48$ $\sigma = 2.77$	$\bar{M} = 29.62$ $\sigma = 2.85$
Dimers	$\bar{D} = 18.41$ $\sigma = 1.91$	$\bar{D} = 18.52$ $\sigma = 1.93$

Table 4.2: Steady-state average plasmid copy numbers for monomer-only (\bar{M}) and dimer-only (\bar{D}) cells with standard deviations (σ), subject to either equal partitioning or random distribution of key molecules at cell division.

	Partitioning	Random Distribution
Monomers	$\bar{I} = 116.29$ $\sigma = 15.65$	$\bar{I} = 116.90$ $\sigma = 15.92$
Dimers	$\bar{I} = 145.28$ $\sigma = 19.42$	$\bar{I} = 146.22$ $\sigma = 19.59$

Table 4.3: Summary of steady-state average inhibitor levels (\bar{I}) with standard deviations (σ) for monomer-only and dimer-only cells, subject to either equal partitioning or random distribution of key molecules at cell division.

	Average	Standard Deviation	Skew
Monomers	$\bar{M} = 29.61$	$\sigma = 2.85$	$\gamma_1 = 0.04$
Dimers	$\bar{D} = 18.52$	$\sigma = 1.93$	$\gamma_1 = 0.09$

Table 4.4: Plasmid distribution statistics for monomer-only and dimer-only cells.

the stability index calculation for a population. To estimate the distributions for monomer-only and dimer-only cells in this model, the simulation was initialised with an average example of each and then run through 1,000,000 successive generations. Plasmids and inhibitors were randomly distributed at the end of each generation.

The results can be seen in Figure 4.2 and Figure 4.3. In Figure 4.3, a log scale is used to make the extremes of the distributions more visible. The distribution statistics are summarised in Table 4.4. As expected, the distributions are roughly normal in form, with a slight positive skew, suggesting that it is a little more likely for the final plasmid copy number to be above the mean than below.

The stability index (the average probability of plasmid loss across the whole cell population, see Chapter 2) was calculated to be 1.64×10^{-8} for monomer-only cells and 1.26×10^{-5} for dimer-only cells, which is approximately 750 times less stable. If a mixed population consisted of 1% dimer-only cells and 99% monomer-only cells, the stability index would be 1.42×10^{-7} , which is still nearly 10 times worse than a dimer-free population. As the mixed population was over 10,000 times less stable than a dimer-free population in the model of Chapter 3, introducing variable copy number clearly has a significant impact on the behaviour of the system.

4.4.3 Dimer Formation and Resolution

In the population model of Chapter 3, dimer formation and resolution events occur, unrealistically, after plasmid replication for a given cell is already complete. In each generation, a cell is given a certain probability of experiencing a dimer formation event due to homologous recombination, and each dimer is given a certain probability of resolving due to XerCD activity. *In vivo*, the emergence or disappearance of a dimer can occur at any time and therefore has a much greater impact on the final plasmid content of the cell. Stochastic simulation can take this into account.

In this model, the rate of dimer formation is dependent on the number of matching sequence pairs that might recombine, that is, the number of plasmid monomer pairs, and the parameter k_3 , which represents the host's overall recombination rate. In published work, the recombination rate is often calculated as a frequency within a cell population, therefore there is no obvious value available for k_3 . However, once incorporated into the population model, the recombination rate can be tuned to give the same steady-state level of dimer-only cells as the existing model.

For XerCD-mediated dimer resolution, even less information is available. In this model, the rate is dependent on the number of dimers in the cell and the parameter k_4 , which represents the XerCD recombination rate.

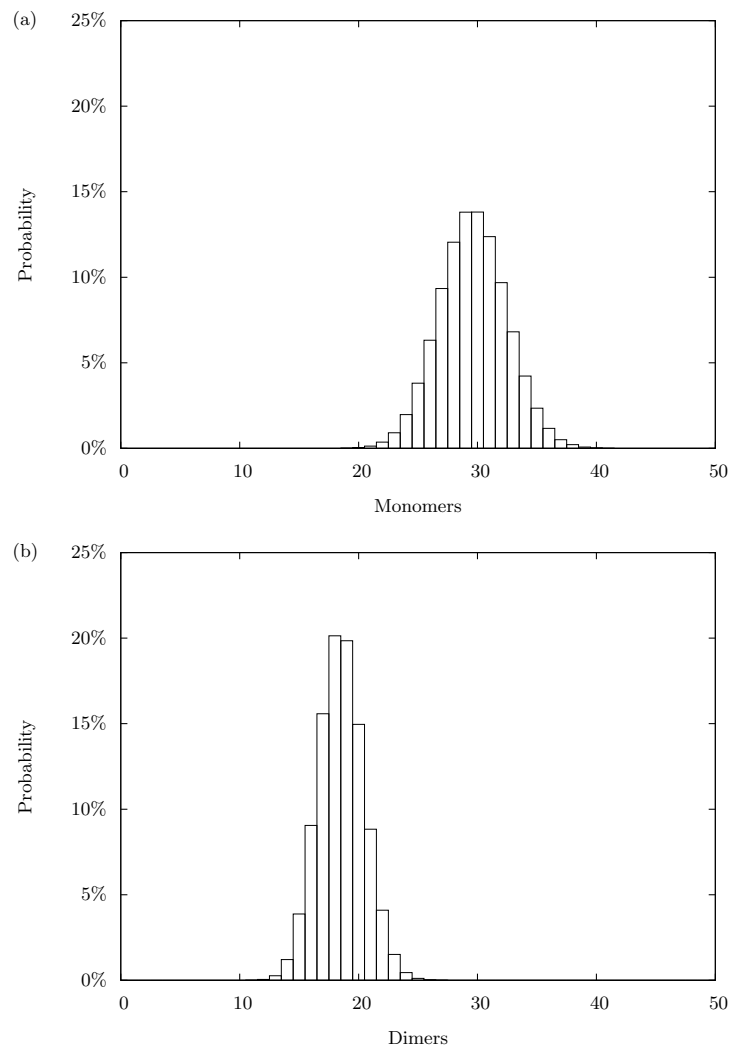


Figure 4.2: Probability distribution of plasmid copy number; (a) monomer-only cells, (b) dimer-only cells

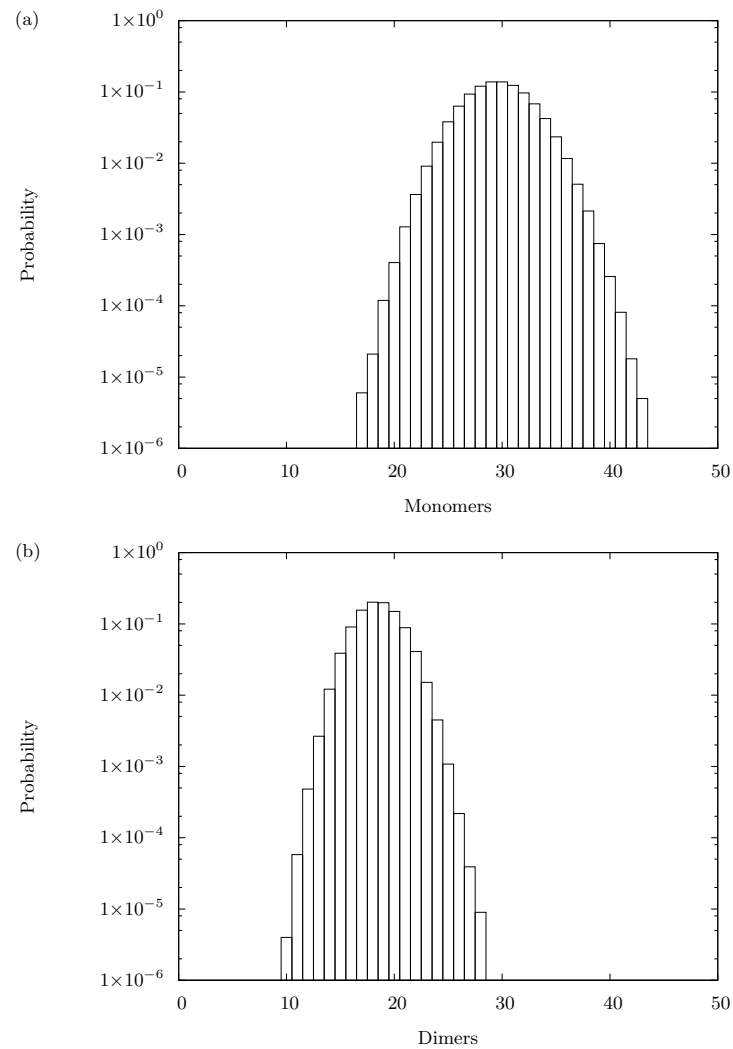


Figure 4.3: Probability distribution of plasmid copy number on a logarithmic scale; (a) Monomer-only cells, (b) Dimer-only cells

This simulation can provide useful information on the relative effects of dimer formation and resolution by looking at the change in average plasmid copy number over many successive generations. This was done initially with a dimer formation rate of $1 \times 10^{-6} \text{min}^{-1}$ and no resolution, for 1,000 generations and 1,000 repeats. This caused the average copy number to drop over time from around 30 to around 20 as the number of monomers fell and number of dimers increased (Figure 4.4(a)). The curve is characteristic of exponential decay, with monomers having an effective half-life of around 250 generations, which is just over 5 days. Of course, it is probable that in the absence of dimer resolution, many cells ended up dimer-only over the long period, but not all repeats necessarily experienced recombination, and some dimers that emerged were not necessarily inherited by the subsequent generation. Averaging the results of many repeats is similar to modelling a small population, though without the metabolic penalty associated with dimers.

The proliferation of ColE1 dimers is opposed by *Xer-cer* dimer resolution *in vivo*. Keeping the recombination rate at $1 \times 10^{-6} \text{min}^{-1}$, the resolution rate was set initially to $1 \times 10^{-6} \text{min}^{-1}$, but this had no effect on the monomer rate of decay (Figure 4.4(b)). The resolution rate was increased until an effect was observed at $1 \times 10^{-3} \text{min}^{-1}$, which increased the monomer half-life to around 500 generations, though the curve was less similar to ideal exponential decay (Figure 4.4(c)). A resolution rate of $1 \times 10^{-2} \text{min}^{-1}$ appeared to prevent decay altogether, though the average was still slightly reduced, likely due to some repeats still ending in a dimer-only state (Figure 4.4(d)).

4.5 Conclusions

In this model of ColE1 plasmid replication, the state of the cell is reduced to the number of monomer plasmids, dimer plasmids and RNAI inhibitors it contains. In the simulation, the cell experiences events that cause it to move from one state to another. The probability of an event occurring is based on the rate at which it occurs in the cell, and each rate is dependent on the state of the cell itself. Thus, the lifetime of the cell is simply a series of events chosen to occur with random numbers. The computer simulation that enacts this model is a definitive improvement upon the plasmid replication method used by the population model of Chapter 3.

The simulation suggested values for the steady-state number of monomers in a monomer-only cell and dimers in a dimer-only cell. With this model and set of parameters, the total number of origins in a dimer-only cell is significantly higher than the number in a monomer-only cell, and the increased quantity of plasmid DNA goes some way to explain the observed metabolic penalty for dimer-only cells. Additionally there was a difference in the steady-state number of monomers and dimers when different feedback mechanisms were used to tune the system. Random distribution of the cell contents at division seemed to slightly increase the plasmid count of the pre-divisional cell.

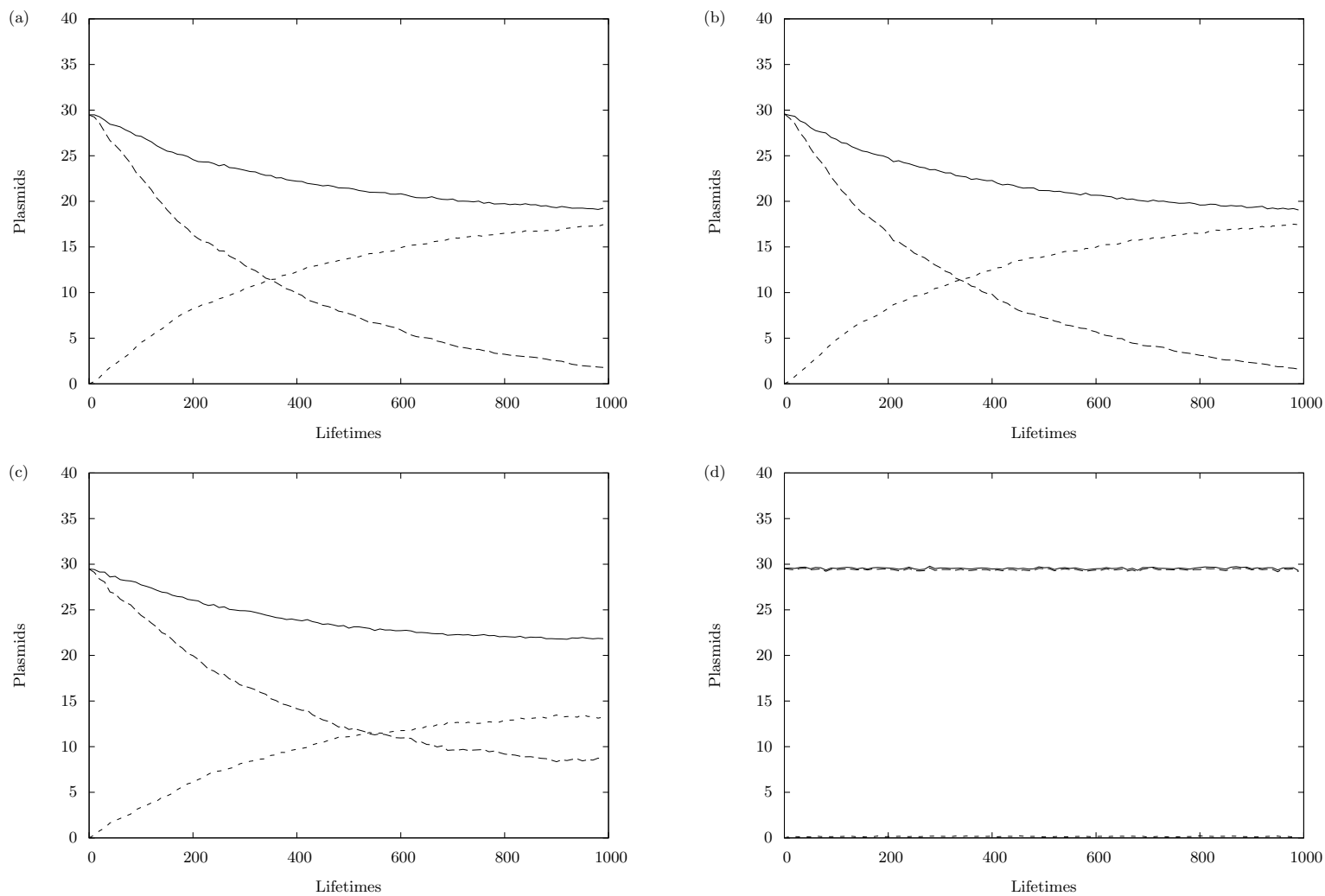


Figure 4.4: Plasmid copy number (—) as the sum of the number of monomers (---) and dimers (-.-.-) over 1,000 successive generations, averaged over 1,000 repeats; dimer formation rate is $1 \times 10^{-6} \text{ min}^{-1}$ for each, dimer resolution rate is (a) no resolution, (b) $1 \times 10^{-6} \text{ min}^{-1}$, (c) $1 \times 10^{-3} \text{ min}^{-1}$, (d) $1 \times 10^{-2} \text{ min}^{-1}$.

Whilst random distribution runs a higher risk of plasmid loss, an increased copy number is also advantageous. Indeed, such a positive feedback loop would encourage a new plasmid with an unreliable partitioning system to increase its copy number and rely less on equal distribution. In turn, an increased copy number would impose a higher metabolic load on host cells, selecting for smaller plasmids, encouraging the loss of the partitioning system altogether, and further amplifying the positive feedback loop to increase copy number.

In comparison to the original model, and that of Chapter 3, variable copy number introduces significant changes. Firstly, the ratio of plasmids in average monomer-only and dimer-only cells is no longer 2:1; the difference between the stabilities of average monomer-only and dimer-only cells is therefore less than before. Overall, due to there being a copy number distribution rather than fixed value, plasmid stability is reduced in this model.

To what extent the simulation accurately reflects plasmid replication *in vivo* is difficult to determine. The average plasmid copy number is within the 15 to 40 range aimed for, but this could have been achieved by chance, especially when parameters were selected from different pieces of work. Whilst the parameters of the model could be retuned for a copy number of 40, the value used in the original model, that number has always been at the upper end of estimates and is not used by other models. The average number of inhibitor molecules is significantly less than the measurements of [Brenner and Tomizawa \(1991\)](#), but transcription rates derived from their results do not fit the ratios reported by [Lin-Chao and Bremer \(1987\)](#).

In the following chapter, this model will be incorporated into the framework of the existing cell population model, allowing observation of the impact of variable copy number on the overall stability of the plasmid.

An Improved Model of Plasmid Behaviour in a Cell Population

5.1 Introduction

In Chapter 3, the simulation of plasmid behaviour in a cell population provided useful insight into the importance of different model parameters. The distribution of monomers and dimers, the stability of the plasmid and the steady-state level of monomer-only and dimer-only cells were all affected by the rate of dimer formation, the increased metabolic load associated with dimers and the dimer resolution rate. The model upon which the simulation was based ensured that the plasmid always replicated up to its full, exact copy number prior to cell division. Given that the copy number control system of ColE1 cannot be so accurate, and that those cells with the fewest plasmids will have the greatest probability of suffering plasmid loss, this aspect of the model is highly unrealistic and potentially misleading.

In Chapter 4, a model of ColE1 replication was developed to consider the variability in plasmid copy number prior to cell division. It expanded upon previous models of the system to include plasmid dimers, the homologous recombination events that create them and the XerCD resolution system that eliminates them. Using parameters derived from *in vivo* work, it was found that dimers over-replicate compared to monomers, which is a possible cause of the metabolic penalty with which they are associated.

In this chapter, the population model of Chapter 3 is improved by incorporating the plasmid replication model. After testing and adjustments, the behaviour of the system is observed through rigorous simulation. The simulation tests a range of values for the parameters explored previously to produce data on the distribution of monomers and dimers in the population, the steady-state level of monomer-only and dimer-only cells and the stability of the plasmid itself.

5.2 Simulation Design and Operation

Unlike the software of the previous two chapters, there was no need for extensive design of this simulation. The cell population is again modelled as a priority queue, with each cell summarised by its number of monomers, number of dimers, time at which it is due to divide, and also, now, its number of RNAI inhibitors. When a cell divides, its monomers, dimers and inhibitors are divided randomly between two daughter cells. The contents of each daughter cell are then used as the initial state for the model of Chapter 4, and each lifetime is simulated in full before the daughters are inserted back into the queue. Daughter cells that inherit no plasmids are rejected before this stage, as previously. The simulation proceeds until the maximum population size is reached, at which point a cell is removed at random from the queue for each new cell inserted into the queue. The population is sampled periodically and the various data analysed and output as per Chapter 3, with the addition of tracking the exact distribution of plasmids in each sample for producing two-dimensional monomer-dimer frequency plots.

There are however, some notable differences in the operation of this model compared to the population model of Chapter 3. Certain compromises were required in order to combine the population and plasmid replication models and these are detailed below.

5.2.1 Copy Number

The plasmid copy number in the models of [Summers et al. \(1993\)](#) and Chapter 3 was fixed at 40 for monomers and consequently 20 for dimers. Whilst estimates for the copy number of ColE1 vary, 40 is towards the upper limit. The parameters of the plasmid replication model of Chapter 4, derived from *in vivo* work, produced an average copy number of around 30, which is still within sensible estimates for ColE1. Whilst this means that the two population models cannot be easily compared, particularly with regards to plasmid stability, such a comparison provides little insight into the biological system anyway.

5.2.2 Metabolic Load

In the models of [Summers et al. \(1993\)](#) and Chapter 3, each dimer in a cell imposed a small but significant metabolic load on the cell. In reality, any plasmid, irrespective of whether it is a monomer or dimer, slows the growth of the host cell. The magnitude of the load is dependent upon plasmid size, copy number and whether or not it expresses costly gene products ([Cheah et al., 1987](#); [Seo and Bailey, 1985](#)). The copy number of a plasmid varies with the host and conditions in which it is grown ([Nordström et al., 1984](#)).

In Chapter 4, the model suggested that the average dimer copy number is higher than half the average monomer copy number, which is consistent with the experimental data of [Chiang and Bremer \(1988\)](#). Typically, dimer-containing cells therefore contain more origins and more plasmid DNA than monomer-only cells. Thus, if growth is penalised on the basis of the number of plasmid origins the cell contains, dimer-containing cells should grow more slowly. This offers an opportunity to move away from the perhaps arbitrary penalty imposed only on dimer-containing cells in the earlier models.

There is further complication however. In the plasmid replication model of Chapter 4, if the growth rate is reduced, the resultant average copy number increases. In reality, of course, the other rates (RNA transcription, degradation, recombination) would also change with the growth rate. However it is not apparent what determines the precise relationships between the different aspects of metabolic load and plasmid content, even from studies of macromolecular concentrations in cells of different generation times ([Churchward et al., 1982](#)). If the growth rate were to decrease without modification of these other rates, then each additional plasmid would slow growth more, allowing for the replication of further plasmids to further slow the growth, and so on, until the cell eventually divides. As there is not enough information on this potential feedback loop, it would be pure conjecture to model the metabolic penalties imposed by plasmids in this way. So, even though it is not necessarily accurate, this model assumes that all of the pertinent rates are proportionally related, such that it does not matter what the growth rate is, the plasmid will achieve the same mean copy number before cell division.

Plasmid replication is simulated in each cell as if it had a generation time of 30 minutes. To account for the metabolic load imposed by increasing plasmid copy number, this generation time is scaled post-simulation, according to the plasmid content of the cell. A small penalty is added to the time at which the cell is due to divide per plasmid origin, which results, on average, in a longer generation time for dimer-containing cells, due to their over-replication. The number of origins in an average dimer-only cell is only 7.42 higher than in an average monomer-only cell. Imposing the 10% difference between the generation times of the two in earlier models would result in an unrealistic 20 minute penalty for an average monomer-only cell compared to a plasmid-free cell. Instead, this model adopts the 3% difference measured *in vivo* ([Summers et al., 1993](#)), corresponding to 0.14 minutes per plasmid origin, which results in a more reasonable 4 minute penalty for an average monomer-only cell compared to a plasmid-free cell.

5.2.3 Dimer Formation Rate

In Chapter 4, the dimer formation rate that resulted in approximately the same number of recombination events per cell per generation as the earlier models was found to be 1×10^{-6} . After initial testing, this rate was found to be too high to imitate the steady-state level of dimer-only cells observed *in vivo* (around 2.3%; [Summers et al., 1993](#)), and the rate

was reduced to 2×10^{-7} . The model of Chapter 4 made no attempt to reconcile itself with *in vivo* data, instead focusing on improving the earlier model. Selecting this rate and the growth rate penalty to fit *in vivo* data significantly improves the accuracy of the simulation.

5.2.4 Initial Conditions

As the model now incorporates variable copy number by considering the number of RNAI inhibitors in the cell, and the average copy number is lower, new initial conditions are required. The average monomer-only cell has approximately 30 monomers, 0 dimers and 116 inhibitors prior to cell division, so to follow the same principle as the original model, applying a single recombination event makes the state of the starting cell (28, 1, 116).

5.2.5 Simulation and Sampling Time

The simulation of Chapter 3 had a time limit of 7,200 minutes and sampled the population every 60 minutes. By reducing the difference in cell growth times between average monomer-only and dimer-only cells to only 3%, the simulation did not reach a steady-state in the available time. Therefore, both the time limit and sample period were doubled to 14,400 minutes and 120 minutes respectively.

5.2.6 Parallelisation

With the complete model of plasmid replication included in the program, a simulated 120 hour run for a population of 10,000 cells takes in the region of 12 hours. In order to collate statistics for 1,000 runs, the software was separated into two programs; the first as a single run of the simulation and the second to process the raw numbers produced by multiple instances of the first. The simulation could then be run across a computing grid and reduce the overall run time from 12,000 hours to approximately 24 hours.

5.3 Basic operation

After initial testing, the simulation was run 1,000 times with the default parameter values, listed in Table 5.1. In a system with variable copy number, there are far more ways of observing the resulting behaviour of the system.

Parameter	Typical Value	Description
n_{pop}	10000	Maximum population size
t_{ori}	0.14 minutes	Division time penalty per plasmid origin
T_{run}	14,400 minutes	Total simulation time
t_{sample}	120 minutes	Sample period
n_{run}	1000	Number of repeats
k_h	0.0234 min^{-1}	Exponential growth rate of the cell
k_1	4.56 min^{-1}	Rate of RNAI transcription initiation
k_2	0.80 min^{-1}	Rate of RNAII transcription initiation
k_3	$2 \times 10^{-7} \text{ min}^{-1}$	Dimer formation rate
k_4	-	Dimer resolution rate
e_1	0.93 min^{-1}	Rate of RNAI degradation
ρ	0.5	Probability that a mature RNAII primer initiates replication
K_i	4.44×10^{19}	RNAI-RNAII interaction constant

Table 5.1: Combined model parameters, taken from the models of Chapters 3 and 4, adjusted to represent *in vivo* data.

5.3.1 Average Steady-State Behaviour

Figure 5.1 shows the average percentage of monomer-only and dimer-only cells in the population over time with standard deviations. The steady-state percentage of monomer-only cells is 96.8% and the steady-state level of dimer-only cells is 2.4%. Figure 5.2 shows the average stability index of the population over time. It settles to a steady-state value of approximately 3.3×10^{-7} with a standard deviation of 1×10^{-7} . This is equivalent to 1 cell in around 3×10^6 emerging plasmid-free, compared to 1 in around 6×10^7 for the monomer-only distribution calculated in Chapter 4.

If the average copy number of 30 is used as the uniform copy number in the model of Chapter 3, with the same steady-state levels of monomer-only and dimer-only cells, the plasmid stability index works out to be approximately 1.5×10^{-6} . Plasmid stability in this simulation is 4 times greater due to the over-replication of dimers, suggesting that variation in copy number may be advantageous for ColE1 when in a host that forms dimers.

5.3.2 Plasmid Distribution

Figure 5.3 shows the plasmid copy number distribution in the population, whether monomer or dimer, at various time points. Figures 5.4 and 5.5 show the probability distribution of cells by total number of monomers and dimers respectively, at the same time points. Finally, Figure 5.6 shows a heat map of the probability distribution of cells by total num-

ber of monomers and dimers; the darker the area, the more cells have that number of monomers and dimers in an average population.

The number of plasmids per cell is clearly distributed bimodally in Figure 5.3. The two peaks are centered around the mean copy numbers for monomers and dimers, and Figure 5.6 shows that the population consists almost entirely of monomer-only (left edge) and dimer-only (bottom edge) cells, with a few cells of mixed population (diagonal trail). Time point (a) corresponds to the peak of the dimer-only cell curve seen in Figure 5.1. Over time, as seen in Figure 5.5 and the leftmost bar of Figure 5.4, the percentage of dimer-only cells decreases from around 20% to 2.4%. The percentage of monomer-only cells increases over time, as seen in Figure 5.4 and the leftmost bar of Figure 5.5.

Monomer-only cells are kept from achieving hegemony by relatively infrequent dimer formation events. The heat maps of Figure 5.6 show a thin trail between the established monomer-only and dimer-only populations, which are some of the progeny of cells in which dimer formation events have occurred. Their progeny will, in turn, either become dimer-only due to the replicative advantage of dimers, or be removed from the simulation before they reach that stage. Some will escape to the monomer-only population if they happen to inherit none of their parent cell's dimers, though this drives their twin faster towards the dimer-only population.

5.4 Dimer Formation Rate

The rate at which homologous recombination converts monomer pairs into dimers was modified to observe the effects on the behaviour of the system. Values for k_3 between 0.25 and 5 times the default rate of $2 \times 10^{-7} \text{min}^{-1}$ were used. It can be seen in Figure 5.7 that the steady-state level of monomer-only and dimer-only cells is dependent on the dimer formation rate. The results are similar to those of Figure 3.5 in Chapter 3, although the change in steady-state levels for both types of cell is more severe as the dimer formation rate increases. Also as before, the time taken to reach steady-state is increased when the rate of dimer formation is decreased.

Figure 5.8 shows the effects of these different dimer formation rates on the stability index of the population. There is a logarithmic relationship between the dimer formation rate and the stability index, as with the model of Chapter 3. The change in stability here is less severe, however, despite the more severe changes in the steady-state levels of monomer-only and dimer-only cells. This is probably because the calculation of the stability index is largely determined by those cells with the fewest plasmids, which in this simulation are produced at the tail end of the probability distribution for dimer-only cells. Increasing the number of dimer-only cells will make it more likely for cells with a low dimer copy number to appear in the population, but they are not the only type of dimer-only cell (as was the case in Chapter 3).

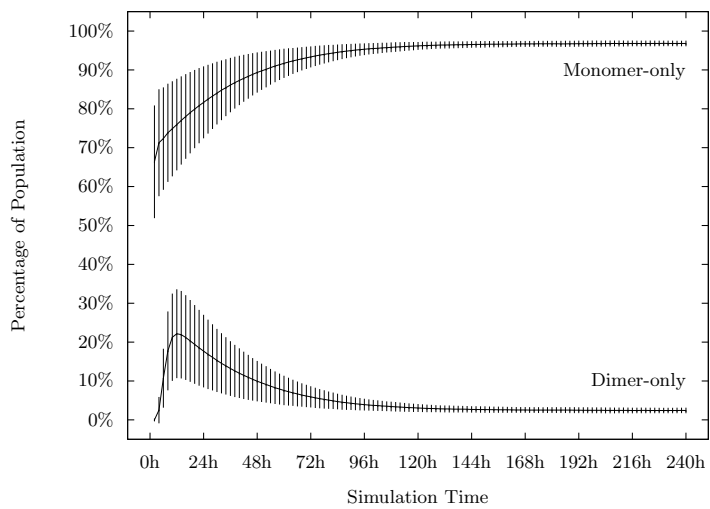


Figure 5.1: Percentage of monomer-only and dimer-only cells in the simulated population over time.

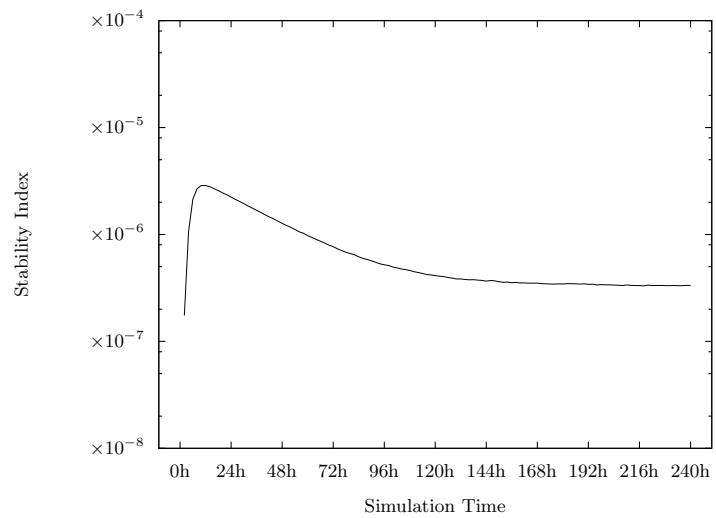


Figure 5.2: Plasmid stability index of the simulated population over time.

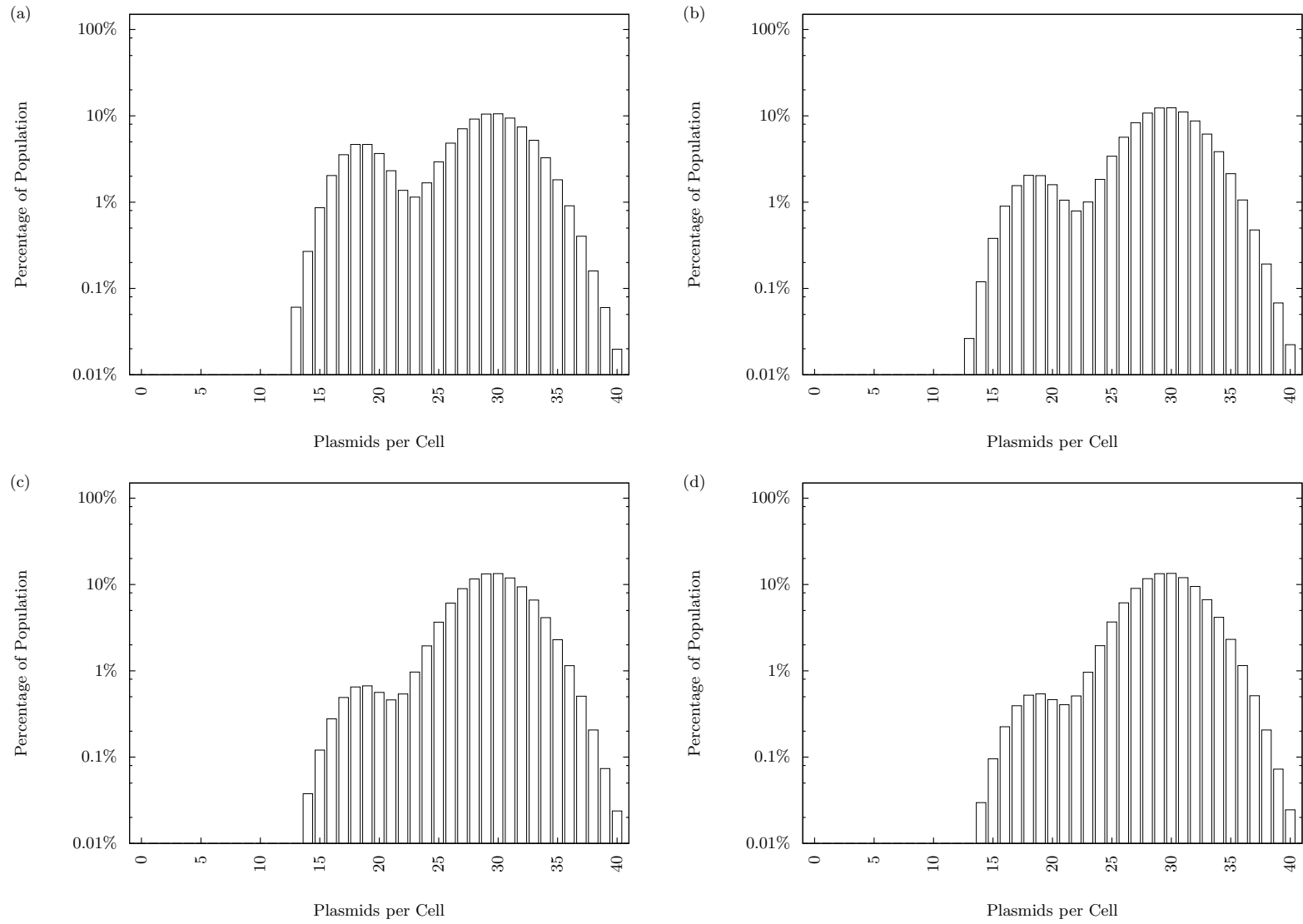


Figure 5.3: Distribution of number of plasmids per cell at various time points; (a) 12 hours, (b) 48 hours, (c) 120 hours, (d) 240 hours.

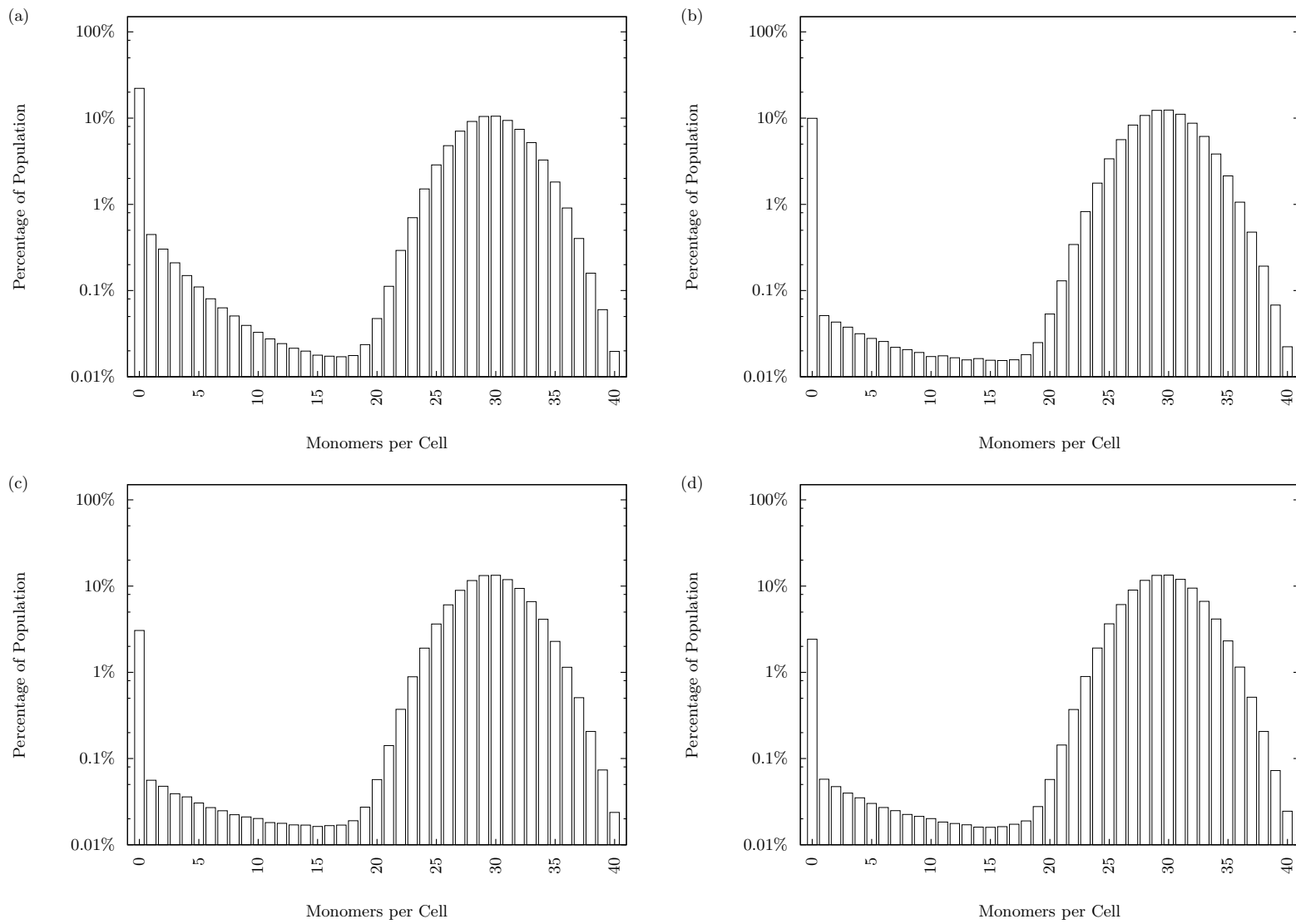


Figure 5.4: Distribution of number of monomers per cell at various time points; (a) 12 hours, (b) 48 hours, (c) 120 hours, (d) 240 hours.

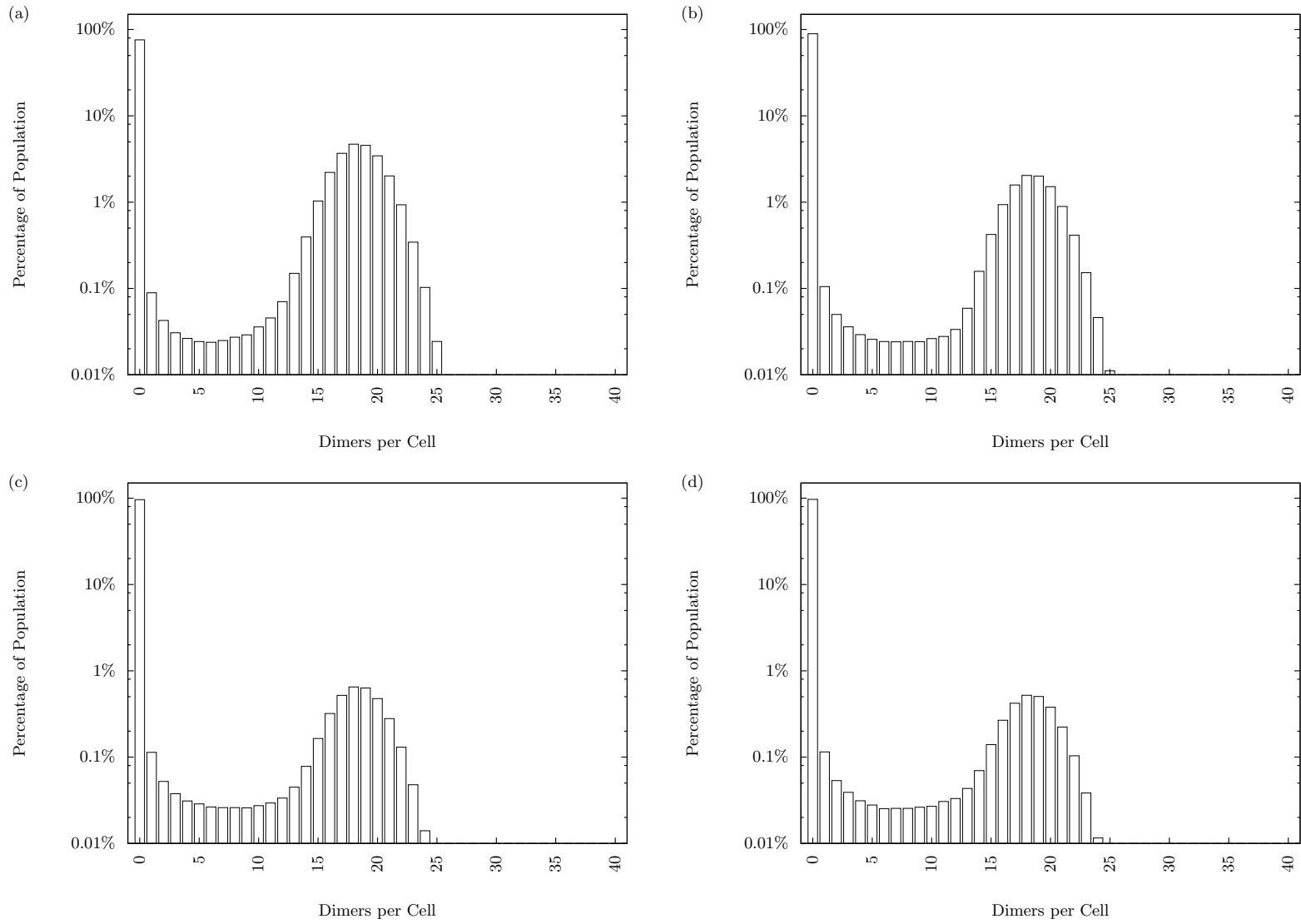


Figure 5.5: Distribution of number of dimers per cell at various time points; (a) 12 hours, (b) 48 hours, (c) 120 hours, (d) 240 hours.

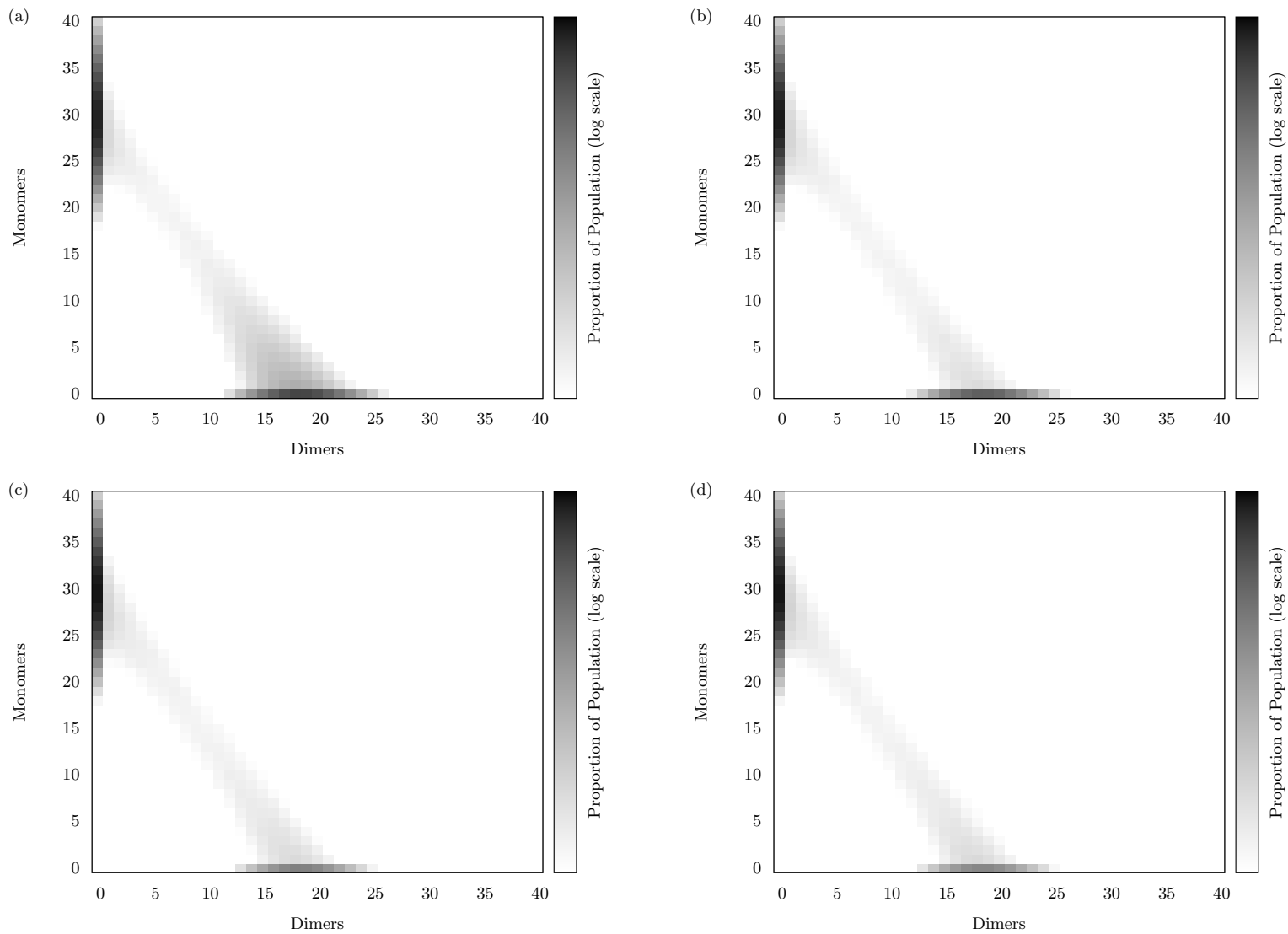


Figure 5.6: Colour map representation of the plasmid probability distribution at various time points; (a) 12 hours, (b) 48 hours, (c) 120 hours, (d) 240 hours.

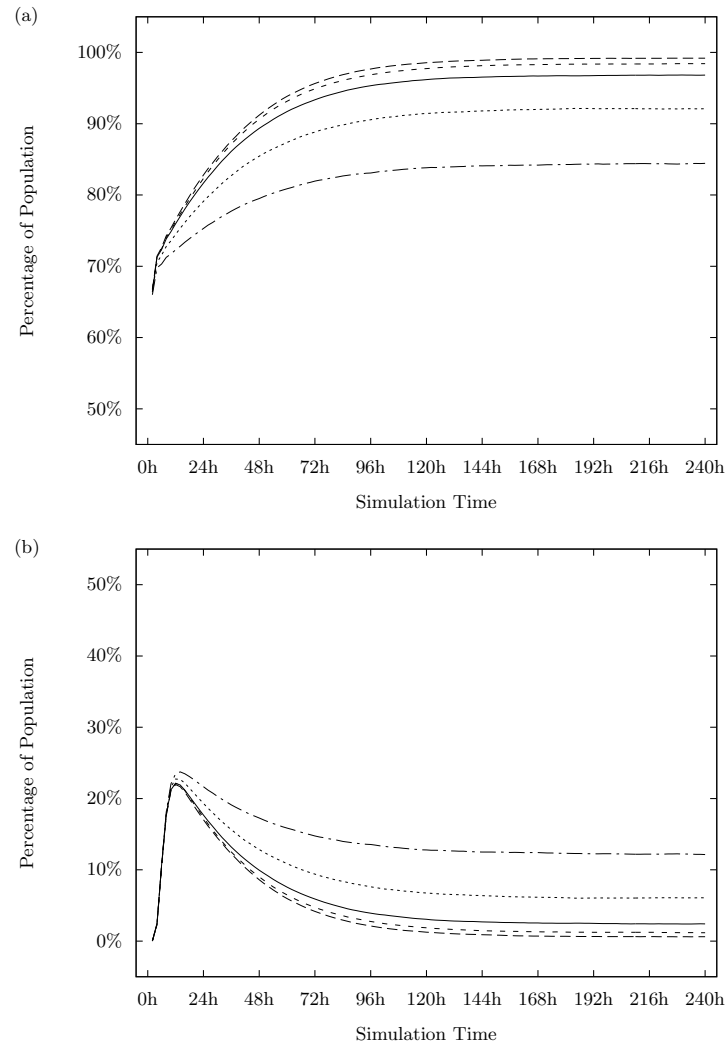


Figure 5.7: Simulation behaviour with different dimer formation rates; (a) Monomer-only cells, (b) Dimer-only cells.

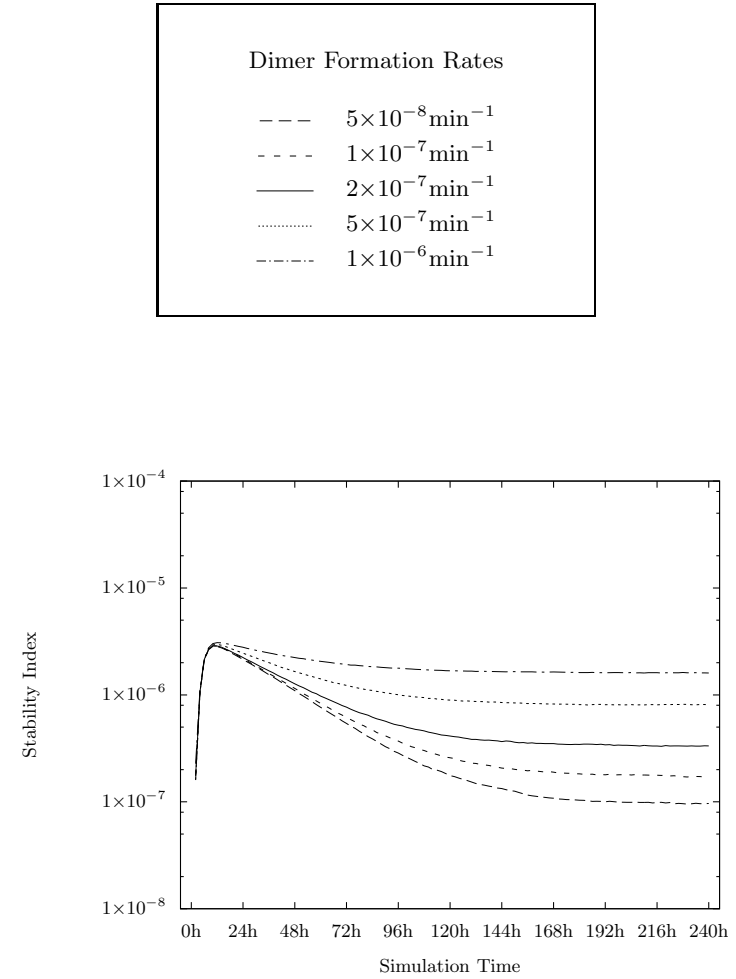


Figure 5.8: Plasmid stability with different dimer formation rates.

5.5 Division Time Penalty

The division time penalty imposed on cells for each plasmid origin they contain was varied to observe the effects on the behaviour of the system. Penalties between 0.2 and 5 times the 0.14 minutes calculated as the default were used. As in the model of Chapter 3, an increased penalty on dimer-containing cells results in a lower steady-state level of dimer-only cells and a higher steady-state level of monomer-only cells (Figure 5.9). Although a higher penalty in this case also increases the generation times of monomer-only cells, they do not have to compete with plasmid-free cells, which are not allowed to survive.

Figure 5.10 shows that the relationship between the metabolic penalty and plasmid stability is again logarithmic. The changes in plasmid stability are less severe than for the model of Chapter 3. The speed with which the system reaches steady-state is considerably increased when the penalty is increased, and the plasmid stability is higher. This contrasts with the effects of changing the dimer formation rate, where steady-states with increased plasmid stability take longer to reach.

5.6 Dimer Resolution

The rate of dimer resolution, controlled by the parameter k_4 , was varied to observe the effects on the behaviour of the system. Values for k_4 between $1 \times 10^{-6} \text{min}^{-1}$ and $1 \times 10^{-2} \text{min}^{-1}$ were tested in a system using otherwise default parameter values. There was no discernable effect on the behaviour of the system until k_4 was raised to around $1 \times 10^{-4} \text{min}^{-1}$ or higher; only results using values at or above this point are discussed here. Figure 5.11 shows the steady-state levels of monomer-only and dimer-only cells in the population. Whilst the level of monomer-only cells is unaffected until k_4 reaches $1 \times 10^{-2} \text{min}^{-1}$, the level of dimer-only cells is reduced a little at $1 \times 10^{-4} \text{min}^{-1}$, considerably more at $1 \times 10^{-3} \text{min}^{-1}$ and almost completely at $1 \times 10^{-2} \text{min}^{-1}$.

Figure 5.12 demonstrates that the plasmid stability index is heavily influenced by the level of dimer-only cells in the population. A k_4 value of $1 \times 10^{-4} \text{min}^{-1}$ has a marginal effect on plasmid stability. A dimer resolution rate of $1 \times 10^{-3} \text{min}^{-1}$ improves on this, but is only as effective as increasing the average plasmid copy number by 1. Increasing the dimer resolution rate further to $1 \times 10^{-2} \text{min}^{-1}$ brings plasmid stability up to the level seen in a population of monomer-only cells.

$1 \times 10^{-2} \text{min}^{-1}$ is 50,000 times the dimer formation rate, demonstrating how significant their replicative advantage is. Colloms et al. (1996) incubated plasmids containing 2 *cer* sites with purified XerC, XerD, ArgR and PepA for 1 hour *in vitro*, and observed that 30% had formed a Holliday junction (a conformation adopted after exchange of the first DNA strand) in this time. Assuming that the rate of Holliday junction formation was

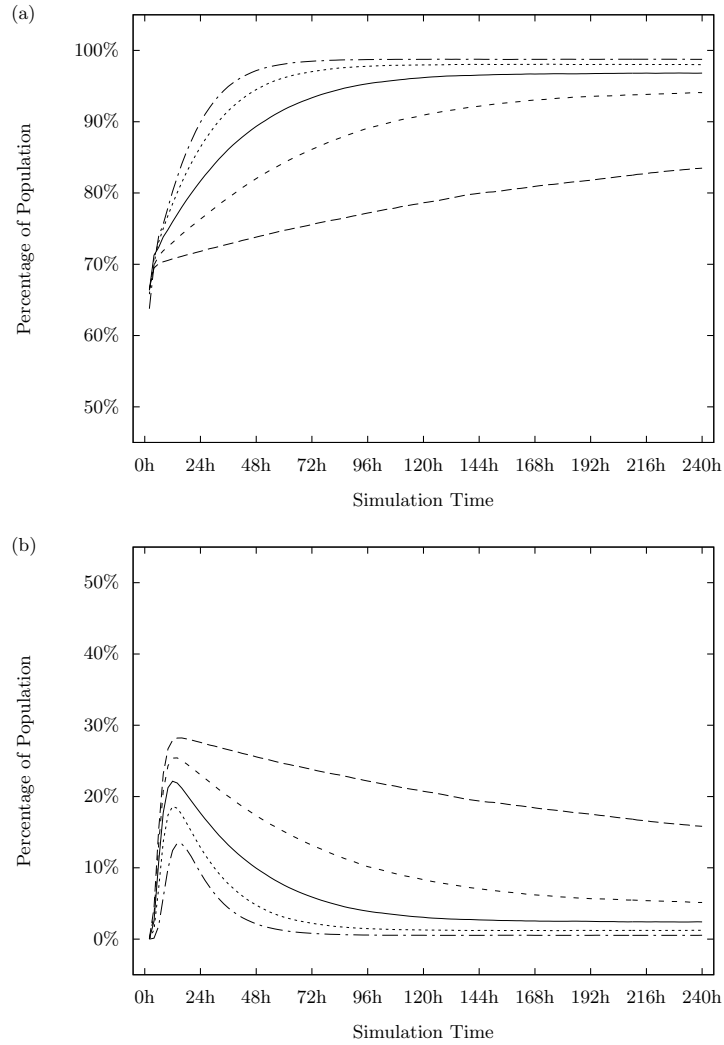


Figure 5.9: Simulation behaviour with different division time penalties per plasmid origin; (a) Monomer-only cells, (b) Dimer-only cells.

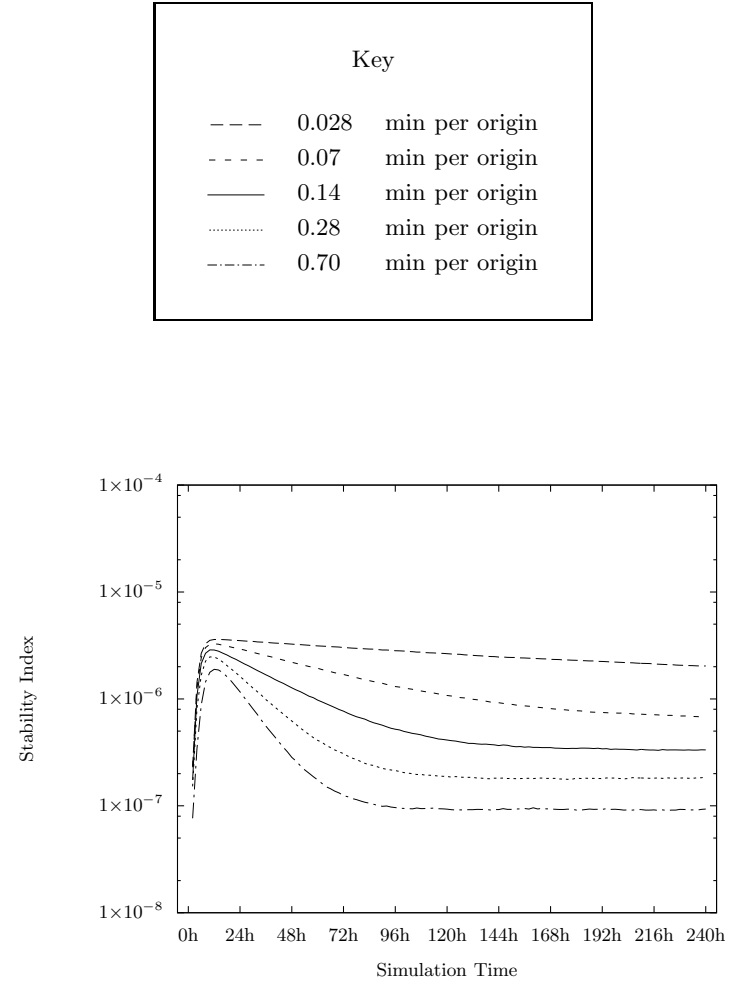


Figure 5.10: Plasmid stability with different division time penalties per plasmid origin.

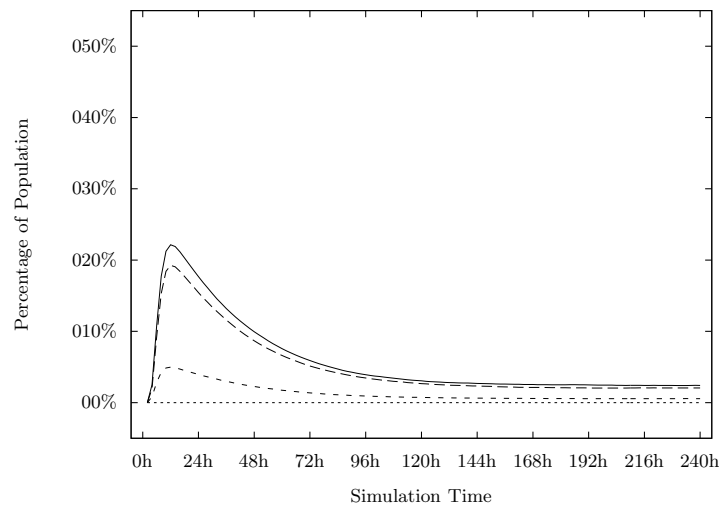
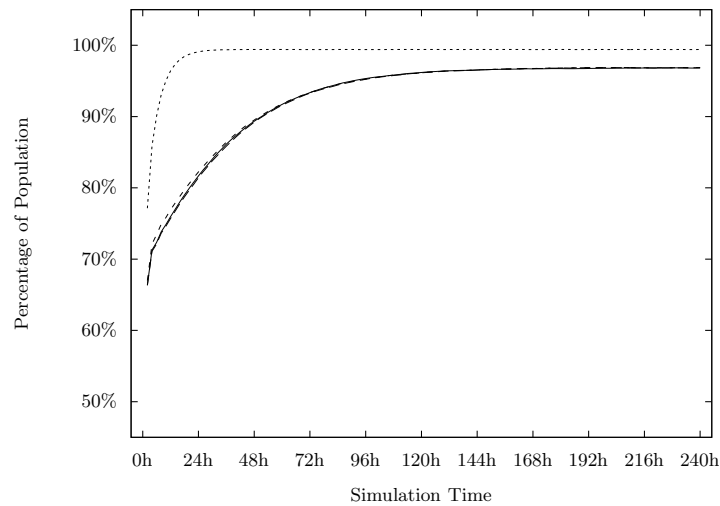


Figure 5.11: Simulation behaviour with different dimer resolution rates; (a) Monomer-only cells, (b) Dimer-only cells.

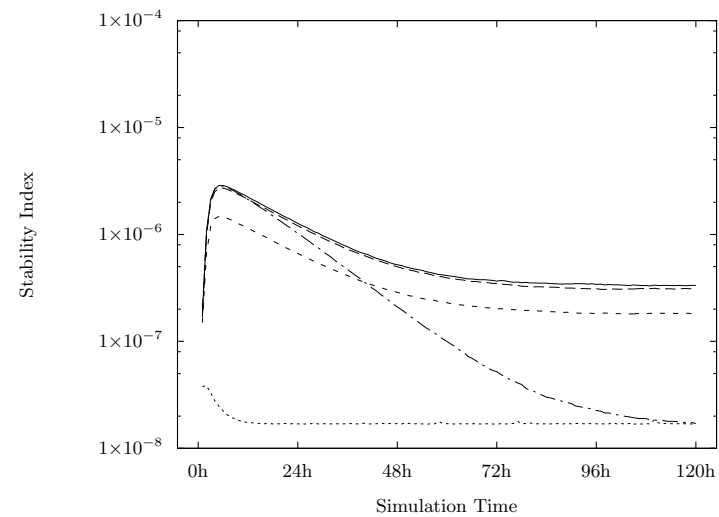
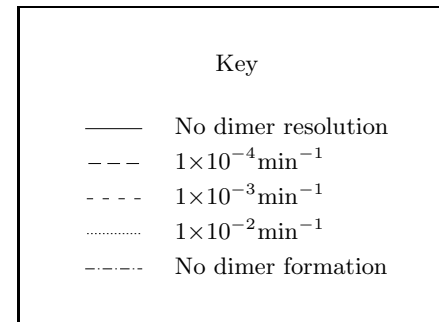


Figure 5.12: Plasmid stability with different dimer resolution rates; stability of a monomer-only population is shown for reference.

proportional to the number of unchanged plasmids, this corresponds to a rate of initial strand exchange of $0.6 \times 10^{-3} \text{min}^{-1}$. Given that this was with saturating quantities of the necessary proteins, and represents only the first step in dimer resolution, it is unlikely that a dimer resolution rate of even $1 \times 10^{-3} \text{min}^{-1}$ can be achieved *in vivo*, let alone $1 \times 10^{-2} \text{min}^{-1}$. This suggests an explanation for the existence of the Rcd checkpoint.

5.7 Conclusions

This model combined the framework of the population model (Chapter 3) with a more accurate model of plasmid replication (Chapter 4). Some compromise was required to unite the two models, mostly in the choice of parameter values. The software itself was modified to run on multiple processors, taking full advantage of the available computing resources.

The behaviour of the system was observed with various parameter values, similarly to the work of Chapter 3. Though the response of the system to a parameter change was less dramatic, the results in this model were similar to those of the previous model:

- Imposing a division time penalty per plasmid origin contained the dimer catastrophe, as the over-replication of dimers penalised their hosts' growth rate.
- The majority of dimers in the population were hosted by dimer-only cells.
- Plasmid stability was logarithmically related to both the dimer formation rate and the division time penalty imposed per plasmid origin.
- Whilst increasing the dimer formation rate made the system take longer to reach steady-state, increasing the division time penalty made it take less time.
- The rate of dimer resolution required to significantly affect plasmid stability was at least 5,000 times greater than the rate of dimer formation; comparing this rate to *in vitro* data suggested an explanation for the existence of the Rcd checkpoint.
- The system was still very robust, reaching the same steady-state from a variety of initial conditions.

The most striking effect of introducing variable copy into this model is the change in the relative stability of plasmids in monomer-only and dimer-only cells. The plasmid stability in a population without dimer formation is only 8 times that in a population with it, which equates to a copy number increase of just 3. In the model of Chapter 3, with a uniform copy number equal to the average copy number here, this difference is around 750 times. Variable copy number worsens the stability of a monomer-only population, but improves the stability of a population with dimers due to their over-replication. This means that the effectiveness of a dimer resolution system is lessened, however efficiently it may act.

The Effect of Indole on Plasmid Replication

6.1 Introduction

Homologous recombination *in vivo* creates plasmid dimers, with independently acting origins of replication that give them a replicative advantage over monomers (Summers et al., 1993). Simulations of the plasmid copy number control system in Chapter 4 suggest that whilst dimers replicate to a higher average copy number than half that of monomers, there are still fewer independent plasmids in a dimer-containing cell, threatening stability.

Xer-*cer* site-specific recombination resolves dimers into monomer pairs, providing a way to restore plasmid stability. It is assisted in this by the Rcd checkpoint. Rcd is a small RNA transcribed from the *cer*-sites of plasmid dimers (Patient and Summers, 1993). Rcd interacts with tryptophanase, enhancing the enzyme's affinity for its substrate, tryptophan, and increasing the production of indole (Chant and Summers, 2007). In broth culture and on plates, cell division and growth are severely inhibited by overexpression of Rcd, or by the addition of exogenous indole (Chant and Summers, 2007; Patient and Summers, 1993). The Rcd checkpoint hypothesis is based on these observations and the fact that plasmid loss can only occur at cell division. It proposes that, in response to dimer accumulation, Rcd expression increases indole production, delaying cell division and allowing time for Xer-*cer* dimer resolution to restore plasmid stability.

In this chapter, the Rcd checkpoint hypothesis is subjected to critical examination and found to be incomplete; the proposed mechanism would not be enough to prevent plasmid instability. The effects of indole on plasmid replication are studied to look for an as yet undetected aspect of the checkpoint and an expanded hypothesis is proposed.

6.2 Examination of the Rcd Checkpoint Hypothesis

6.2.1 Inhibition of Cell Division is Insufficient to Ensure Plasmid Stability

The Rcd checkpoint hypothesis suffers from a critical flaw. If cell division is prevented in an ideal manner, then all other cellular processes would continue, including growth. Continued cell growth would dilute RNAI and so plasmid replication would continue at the same rate as before. In a cell with both monomers and dimers, dimers would continue to out-replicate monomers. If the rate at which dimers are being created remains the same, whether the cell has divided or not, then dimer resolution is made no more effective by delaying division.

Figure 6.1 illustrates this point; the fates of two identical cells, A_0 and B_0 , are shown with and without delayed cell division due to the Rcd checkpoint, respectively. At the outset, each contains three monomers and a single dimer of a plasmid with an imagined copy number of 5. A_0 , whose cell division is delayed, eventually grows to a volume equal to the combined volume of B_1 and B_2 (daughters of B_0 , which divided normally). The plasmid in both cases has maintained its concentration, but the dimer has had greater opportunity to replicate in cell line A. Plasmid loss is avoided by not dividing, but this is a temporary solution since the cell must eventually divide. Whilst plasmid loss is less likely when cell A_0 divides than when cell B_0 divides, its daughters will be worse off in the long run, with fewer total plasmids and more dimers. Delayed cell division can therefore only be effective if the accumulation of dimers is simultaneously prevented. A plausible way to do this would be by inhibiting plasmid replication.

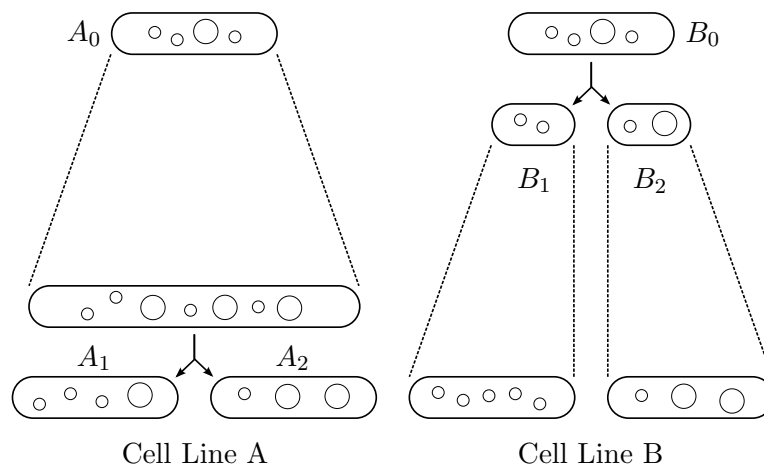


Figure 6.1: The fates of cells A_0 and B_0 , with delayed and normal cell division respectively. Delayed cell division in A does not improve the stability of the plasmid in the long run. Plasmid monomers are depicted as small circles; dimers as larger circles. Dashed lines show cell growth; solid lines show cell division.

6.2.2 Inhibition of Cell Growth is Insufficient to Prevent Plasmid Replication

Indole does more than just prevent cell division; it also inhibits cell growth at a concentration of 3 mM or higher ([Chant and Summers, 2007](#)). If slower growth were to indirectly reduce the rate of plasmid replication, then the checkpoint would be effective.

When cell growth is inhibited, the subsequent plasmid replication behaviour will depend on changes in the transcription rates of RNAI and RNAII and the degradation rate of RNAI. If the transcription, degradation and growth rates in the plasmid replication model of Chapter 4 are all increased or decreased proportionally to one another, then the plasmid replication rate matches the cell growth rate. This means that the average copy number prior to cell division remains constant, at around 30.

If the growth rate is decreased but the synthesis and degradation rates of RNAI and RNAII remain unaltered, the average plasmid copy number is increased. In this case, a higher concentration of RNAI results in a decreased chance of successful replication priming. However, over the increased generation time, there are more replication attempts made by RNAII transcription and, despite the reduced chance of success for each RNAII transcript, there are more plasmid replication events overall. If the transcription and degradation rates were decreased more severely than the growth rate, the average plasmid copy number is decreased, which would make the checkpoint effective.

There is empirical evidence supporting the hypothesis that the plasmid replication rate is only marginally affected by a reduced growth rate. Although R1 has control kinetics closer to hyperbolic than exponential, if the plasmid is given more time to replicate before cell division, copy number increases. [Engberg and Nordström \(1975\)](#) report that the copy number of plasmid R1 is increased from two to six when the growth rate is decreased from 1.8 to 0.4 doublings per hour. [Atlung et al. \(1999\)](#) similarly report that the copy number of plasmid pBR322, a ColE1 derivative, increases 3- to 4-fold when the generation time is increased from 20 to 80 minutes.

Increased plasmid copy number is also associated with slowed growth as broth culture enters stationary phase. This was quantified by [Stueber and Bujard \(1982\)](#), who report that the copy number of pBR322 increases 4-fold from exponential to stationary phase in correlation with the reduced growth rate. Stationary phase cultures are therefore preferred for high yield plasmid extraction. The addition of chloramphenicol to a culture just entering stationary phase further increases the yield of ColE1-like plasmids ([Clewelly, 1972](#); [Frenkel and Bremer, 1986](#)). In the presence of chloramphenicol, protein synthesis is inhibited but RNA and DNA synthesis continue for as long as the relevant proteins remain intact. This leads to continued expression of the RNAs responsible for plasmid copy number control and, as seen at slower growth rates, continued plasmid replication.

These pieces of evidence all indicate that reducing the cell growth rate will not prevent continued plasmid replication. At best it might reduce the rate of dimer proliferation, but the difference is unlikely to be dramatic enough for the dimer resolution system to become significantly more effective. If the inhibitory effects of indole on cell division and growth are not sufficient to make the resolution system effective, how does the Rcd-indole mechanism stabilise the plasmid? Indole must have an additional, as yet unknown, effect on the cell. The dimer catastrophe can be mitigated either by increasing the rate of dimer resolution or decreasing the rate of dimer creation. Stimulation of site-specific recombination by indole seemed less plausible, so the effect of indole on plasmid replication was investigated.

6.3 The Effect of Indole on Plasmid Replication

In order to investigate the effect of indole on plasmid replication it is necessary to elevate its intra-cellular concentration in a controlled manner. One way to achieve this is to stimulate endogenous indole synthesis *via* a temperature-sensitive Rcd expression system on a plasmid (Rowe and Summers, 1999). However, indole production by this method is hard to quantify, and introducing a temperature change creates an additional variable. Alternatively, indole can be introduced exogenously by addition of indole stock solution, in ethanol, to broth culture. This simpler option was used here.

The major difficulty in establishing whether indole affects plasmid replication directly is that it also impedes cell division and growth. A way to control for the growth and division effects is needed. There are a number of antibiotics, including chloramphenicol, that will prevent cell division and growth entirely, but allow for continued plasmid replication. Therefore, to observe the direct effect of indole on plasmid replication, chloramphenicol and indole could be added to broth culture together, and the effect compared with the addition of chloramphenicol alone.

6.3.1 Plasmid Replication in Chloramphenicol-Treated Cells

To confirm that plasmid replication continues in chloramphenicol-treated cells, a culture of *E. coli* W3110 that had been transformed with the plasmid pBR322 was grown in LB medium at 37°C. When the culture reached an OD₆₀₀ of around 0.2, chloramphenicol (34 µg ml⁻¹) was added to stop cell division and growth. At this time point, and every 1.5 hours thereafter for the next 7.5 hours, plasmid DNA was prepared from a fixed volume of culture and the OD₆₀₀ recorded (Figure 6.2). The samples were then normalised to correct for variation in OD₆₀₀ and visualised by gel electrophoresis (Figure 6.4(a)). The gel image was analysed with Quantity One software to produce densitometry data, quantifying the change in plasmid DNA over time (Figure 6.3).

The gel image shows that after the addition of chloramphenicol, the brightness of the plasmid bands grows consistently over time, reaching maximum brightness at 6 hours. This is confirmed by the densitometry data. As the OD_{600} of this culture increases only slightly after the addition of chloramphenicol, this indicates that plasmid replication continues despite the inhibition of cell division and growth. The apparent cessation of plasmid replication after 6 hours is because either replication has stopped due to degradation of the relevant machinery, or band brightness is saturated by visualisation of the gel itself.

6.3.2 Indole Prevents Plasmid Amplification

The experiment described in the previous section was repeated with the addition of both chloramphenicol and 5 mM indole. Figure 6.4(b) shows that whilst the brightness of the plasmid bands increases slightly in this culture, it does so far less than in a culture treated with chloramphenicol alone (Figure 6.4(a)). The densitometry data (Figure 6.3) for these samples confirms this conclusion. Thus the plasmid amplification seen in the culture treated with chloramphenicol alone is prevented by simultaneous treatment with indole. As the OD_{600} of the culture containing both chloramphenicol and indole decreases over time (Figure 6.2), and the samples were normalised by this measure, the total plasmid DNA in the culture is increasing even less than Figure 6.3 might suggest, perhaps even remaining the same over time.

As a positive control (Figure 6.4(c)), a culture was treated with both chloramphenicol and nalidixic acid ($30\text{ }\mu\text{g ml}^{-1}$), as nalidixic acid is known to inhibit growth, cell division and plasmid replication (Uhlen and Nordström, 1985). The effect on plasmid replication was similar to that seen in the culture treated with indole and chloramphenicol (Figure 6.4(b)). Densitometry confirms that the two sample sets are similar, though normalisation obscures the fact that there is an increase in the total plasmid DNA over time for the culture treated with nalidixic acid, as its OD_{600} increases 3- to 4-fold (Figure 6.2). It is possible that this concentration of nalidixic acid is insufficient to completely prevent plasmid replication, but is inhibiting the amplification seen in culture treated with chloramphenicol alone.

So for cultures treated with chloramphenicol and either indole or nalidixic acid, plasmid replication does not continue as in the culture treated with chloramphenicol alone. Nalidixic acid is known to inhibit plasmid replication, so this is no surprise. However, it appears that indole is also directly inhibiting plasmid replication.

6.3.3 Indole Alone Inhibits Plasmid Replication

It is not certain that the mechanism of plasmid replication in chloramphenicol-treated cells is the same as in the absence of chloramphenicol. It could be argued that the inhibition of plasmid replication by indole in chloramphenicol-treated cells might not occur in the

Key					
—	Cm	(a)	---	Cm + Indole	(b)
----	Cm + NA	(c)	----	Indole	(e)
.....	Cm technical repeat	(d)	NA	(f)

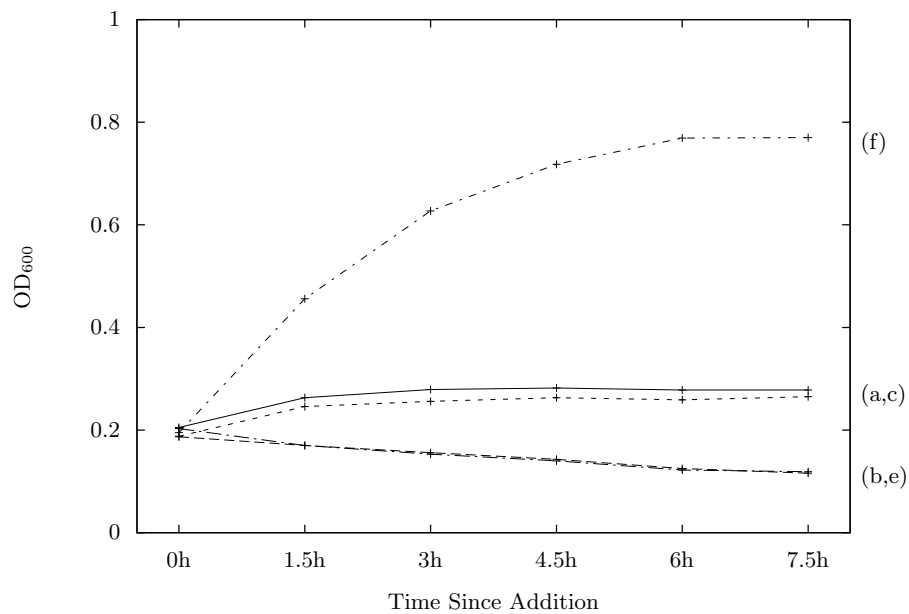


Figure 6.2: OD₆₀₀ of broth cultures treated with combinations of chloramphenicol (Cm), indole and nalidixic acid (NA (see Key).

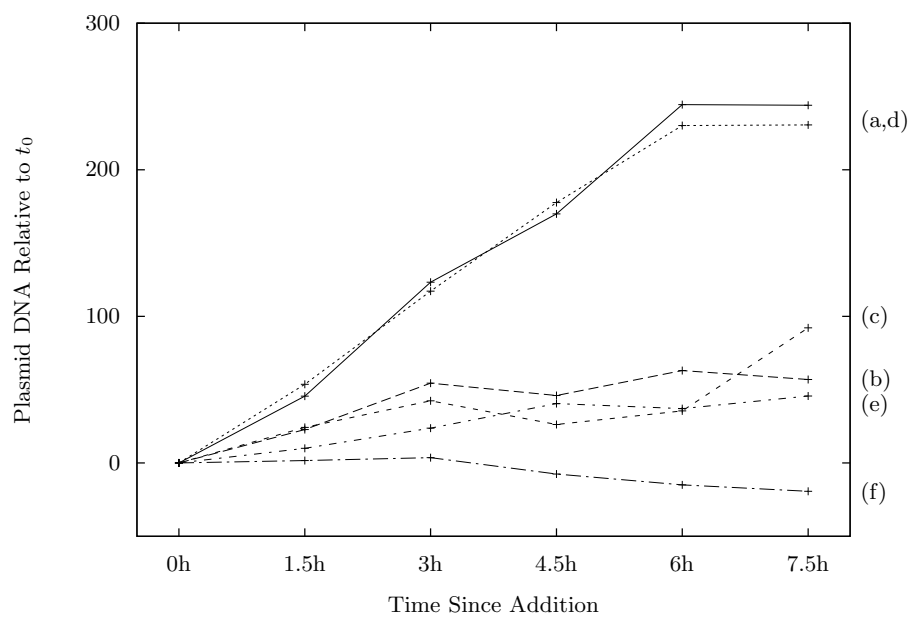


Figure 6.3: Change in plasmid DNA content of cells treated with combinations of chloramphenicol (Cm), indole and nalidixic acid (NA (see Key). Values are in arbitrary units of brightness measured relative to t_0 .

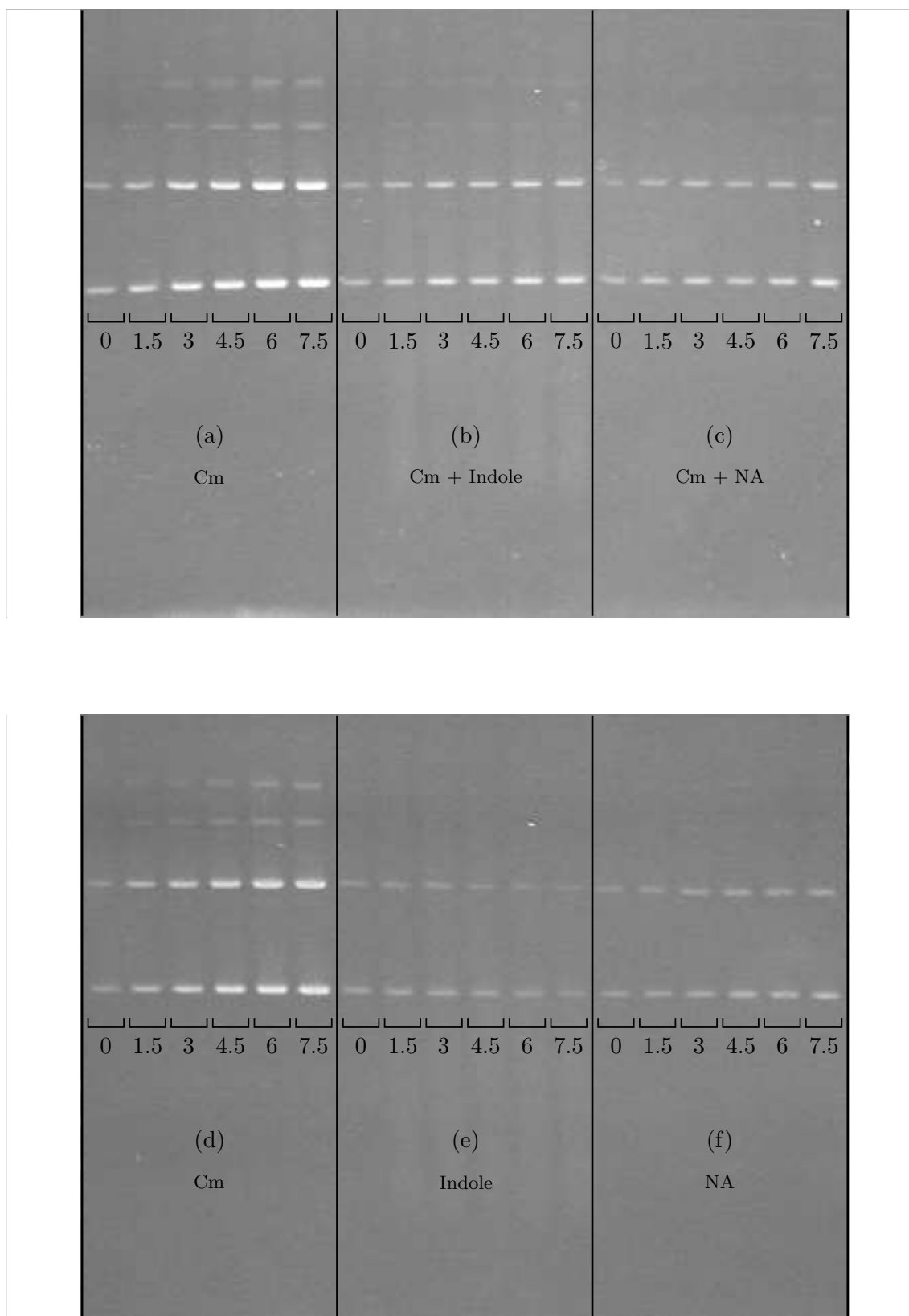


Figure 6.4: Plasmid DNA extracted from cells treated with different chemical combinations; (a,d) chloramphenicol, (b) chloramphenicol and indole, (c) chloramphenicol and nalidixic acid, (e) indole, (f) nalidixic acid. The time at which each sample was taken is given beneath it, in hours.

absence of chloramphenicol. To test this possibility, 5 mM indole was added to a culture in the absence of chloramphenicol and the experiment repeated. A control with only nalidixic acid ($30 \mu\text{g ml}^{-1}$) was also tested. The results were again visualised with gel electrophoresis, and the subsequent images analysed by densitometry.

For the culture treated with indole alone, the brightness of the plasmid bands decreases slightly over time (Figure 6.4(e)), confirmed by the densitometry data (Figure 6.3). This suggests that plasmid DNA is actually being lost from the culture. The OD_{600} of this culture decreased in the same way as the culture with both indole and chloramphenicol (Figure 6.2). This is likely due to cell lysis, which will cause plasmid loss into the supernatant, where the DNA can be sheared or lost during the extraction procedure. However, the most important observation is that there is no continued plasmid replication, confirming the conclusion of the first experiment that indole is inhibiting plasmid replication and demonstrating that the inhibition is independent of the presence of chloramphenicol.

For the culture treated with nalidixic acid alone, the increase in brightness of the plasmid bands was similar to that of the culture with both nalidixic acid and chloramphenicol (Figure 6.4(f)). The densitometry data for these two cultures is also very similar (Figure 6.3). The OD_{600} of the culture treated with nalidixic acid, however, rises over time; cells treated with nalidixic acid are known to elongate for a period after exposure (Kantor and Deering, 1968). So overall, the total plasmid DNA in the culture is increasing, such that the plasmid copy number is being maintained. This is still evidence of inhibition, however, as the slower growth of the culture would ordinarily result in increased copy number.

This set of experiments was repeated with plasmid pUC18, which is a very high copy number ColE1 derivative. The results were the same; 5 mM indole inhibited plasmid replication with or without chloramphenicol (data not shown).

6.4 Concentration Dependence of Plasmid Replication Inhibition by Indole

Using chloramphenicol treated cells, different concentrations of indole were tested to find the concentration dependence of the plasmid replication inhibition. Cultures were treated with chloramphenicol ($34 \mu\text{g ml}^{-1}$) as well as indole at 1 mM, 2 mM, 3 mM, 4 mM and 5 mM. As indole was introduced to each culture in ethanol solution, a 0 mM indole, 0.5% ethanol control was also tested. In each case, the OD_{600} of the culture remained approximately the same over time (Figure 6.5). Figure 6.7 shows the results of the subsequent plasmid preparations. The gels were analysed with the same densitometry methods as before and the results are shown in Figure 6.6.

A concentration of 2 mM indole or higher is enough to have an effect on the subsequent accumulation of plasmid DNA. The difference is significant at 3 mM indole or higher and 5 mM prevents replication almost entirely. In experiments on the effects of indole on cell

division and growth, 3 mM is the point at which colonies begin to appear smaller on agar plates, with no visible growth at all at higher concentrations. This perhaps indicates that the same intracellular target is involved in the inhibition of plasmid replication, cell division and growth.

In the cells treated with 1 mM indole, there appears to be a greater accumulation of plasmid DNA than in the ethanol control. This is a surprising observation, as 1 mM indole is approximately the concentration seen in the supernatant of stationary phase cultures, implying that plasmid replication could be enhanced at a point where the plasmid ought to be more considerate of the metabolic load it is imposing on its host.

6.5 Replication Inhibition for Monomers and Dimers

With strong evidence that indole inhibits plasmid replication, it was considered that there may be stronger inhibition of dimer replication than monomer replication. This is an attractive idea, as it would allow monomers to continue to replicate whilst the dimers cease replicating and undergo resolution, all the while maintaining high copy number.

To test this possibility, plasmid pBR322 was transformed into *E. coli* strain JC8679; a *recBC* mutant that has high recombination activity, resulting in a ‘ladder’ of plasmid multimers. After extraction of plasmid DNA, the monomers and dimers were separated by gel electrophoresis and their bands cut from an agarose gel for purification. They were then transformed separately into *E. coli* strain DS941, a *recF* mutant with minimal recombination activity, in which the incoming plasmid species (monomer or dimer) would be maintained.

Identical cultures containing either monomers or dimers were treated with chloramphenicol ($34 \mu\text{g ml}^{-1}$) and indole at 0 mM (ethanol control), 1 mM, 2 mM, 3 mM, 4 mM and 5 mM. Qualitatively, the effect of indole on plasmid replication appeared to be the same for both monomers and dimers (data not shown). However, it is worth noting that they were tested separately. It remains possible that when both are present in the same cell, indole inhibits the replication of plasmid dimers preferentially.

6.6 The Effects of Indole Analogs on Plasmid Replication

In previous studies in this laboratory, it was found that certain *hns* mutants of *E. coli* exhibited an unusual indole response. The addition of 3 mM indole to a broth culture of this strain causes it to enter a quiescent state, in which cell division and growth are suspended, but the rate of protein synthesis remains high (C. Chen and D. Summers, pers.

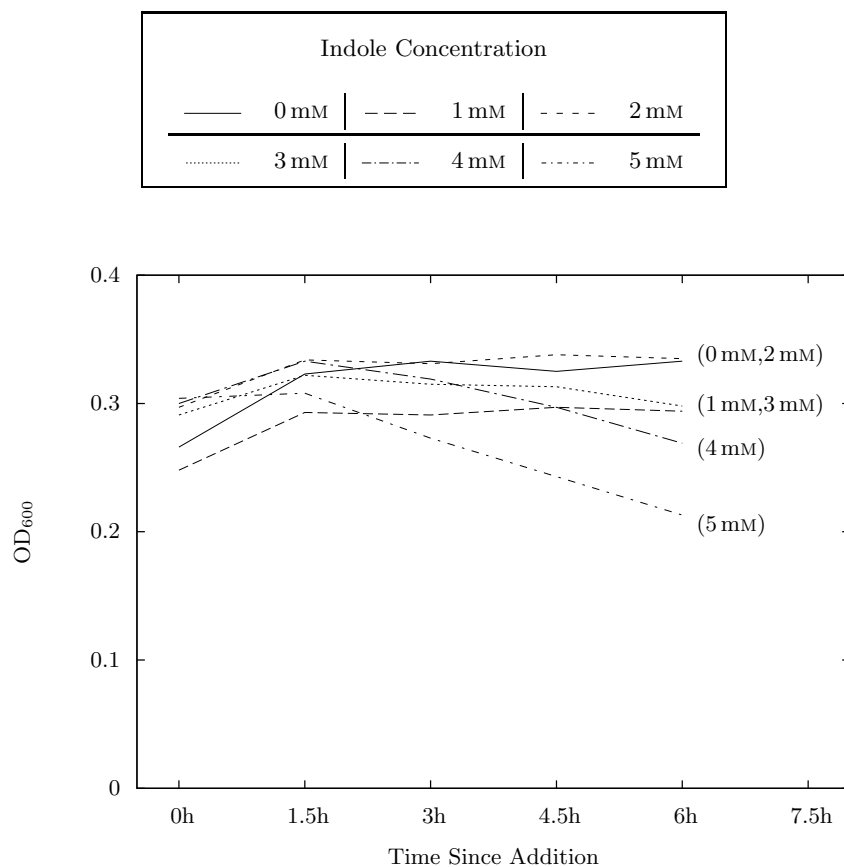


Figure 6.5: OD₆₀₀ of the broth cultures after addition of the chloramphenicol at $34 \mu\text{g ml}^{-1}$ and indole at different concentrations (see Key).

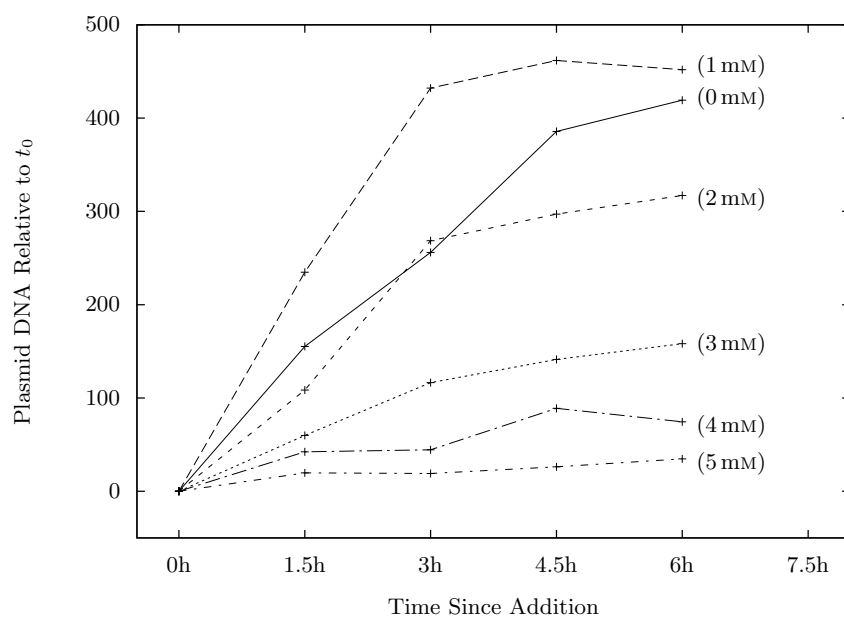


Figure 6.6: Change in DNA content of cells treated with chloramphenicol at $34 \mu\text{g ml}^{-1}$ and indole at different concentrations (see Key). Values are in arbitrary units of brightness measured relative to t_0 .

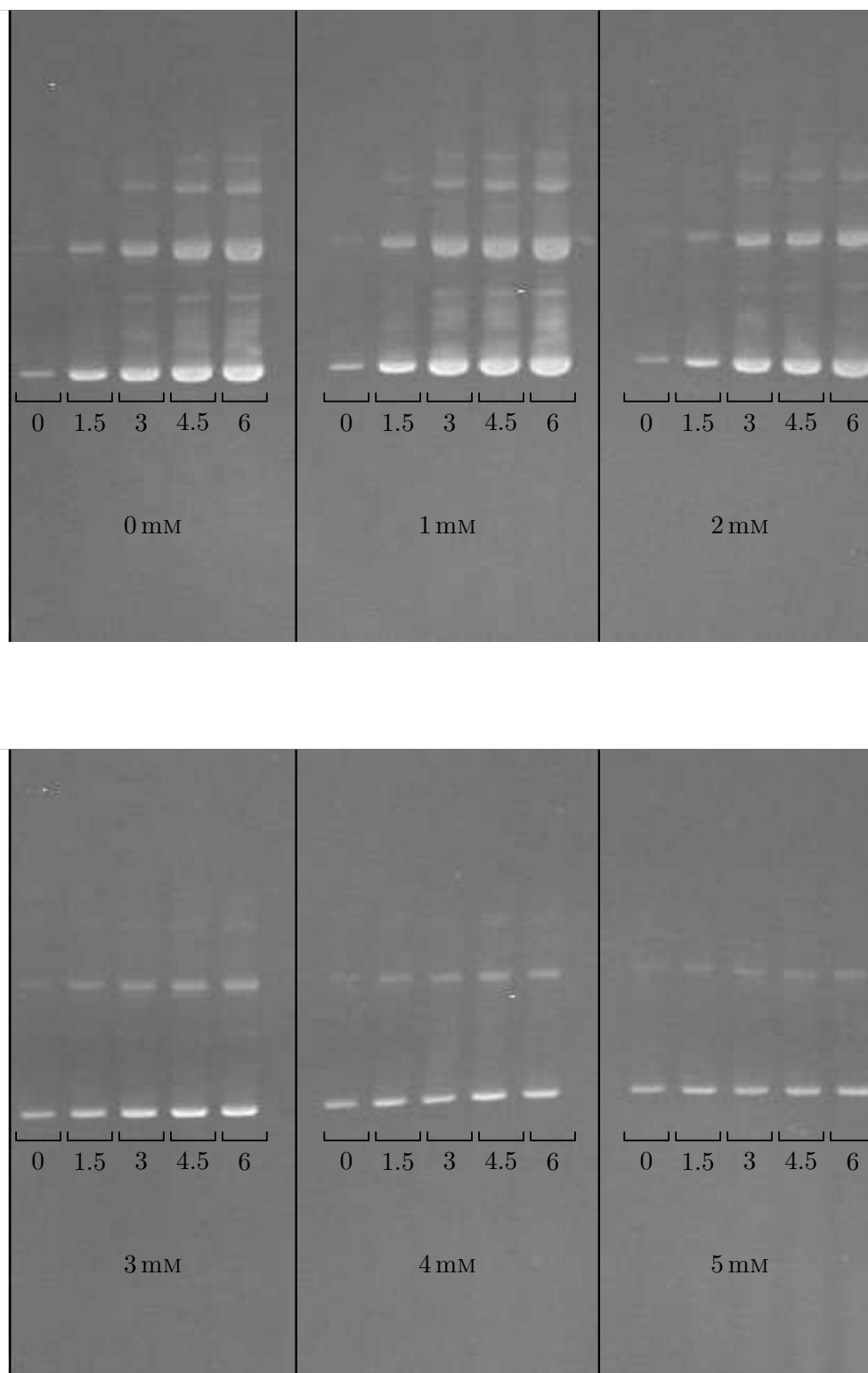


Figure 6.7: Plasmid DNA extracted from cells treated with chloramphenicol at $34 \mu\text{g ml}^{-1}$ and indole at different concentrations. The time at which each sample was taken is given beneath it, in hours.

comm. 2010). This is desirable for use in industry, though there is still a degree of cell lysis due to indole toxicity.

In order to find alternatives to indole for the induction of quiescence, a number of chemical analogs of indole (see Table 6.1) were tested (C. Chen, pers. comm., 2010). These analogs were screened initially for their effects on growth of the wild-type *E. coli* strain W3110. It was found that isoquinoline, quinoline, 3- β -indoleacrylic acid and 1-acetylintoline inhibited growth in a similar way to indole. Indoline, tryptamine, indole-3-acetic acid and pyrrole had considerably less or no effect on growth.

To investigate whether these analogs also affect plasmid replication, chloramphenicol-treated cells were exposed to 5 mM of each and assayed as before. Figure 6.8 shows the results of the plasmid preparations. The gel was analysed with the same densitometry methods as before and the results are shown in Figure 6.9. Isoquinoline, indoline and 3- β -indoleacrylic acid appear to inhibit plasmid replication strongly. Indole appears to inhibit plasmid replication less than these three according to the densitometry data, however the higher level of background fluorescence in those tracks makes the assay less reliable – it is symptomatic of cell lysis, which makes normalisation, to correct for differences in OD₆₀₀, inaccurate.

Whilst isoquinoline and 3- β -indoleacrylic acid inhibit growth in addition to plasmid replication, indoline does not. Quinoline inhibits growth but does not appear to affect plasmid replication. 1-Acetylintoline also inhibits growth, but seems to have no effect on plasmid replication for the first two hours of the assay, after which it appears inhibitory. This may be due to slow uptake by the cell, but the transport mechanism for this chemical is unknown. It has been demonstrated that the same indole concentration is required to inhibit plasmid replication, cell division and growth, suggesting that a single target may be responsible for all three phenotypes. In contrast, this experiment suggests that different indole analogs can inhibit different processes, suggesting that indole has multiple targets.

6.7 Conclusions

The Rcd checkpoint hypothesis proposes that the role of indole is to prevent the division of dimer-containing cells, granting the resolution system the time to restore plasmids to the monomeric state. However, careful inspection of the hypothesis revealed that inhibition of cell division is not sufficient to achieve this, as plasmid replication would continue. Dimers would continue to out-replicate monomers, and possibly overwhelm the *Xer-cer* resolution system.

Indole, however, also slows growth. Whilst this reduces the dilution rate of RNAI, decreasing the chance that RNAII will successfully initiate plasmid replication, there is more time available for RNAII transcription. This means that plasmid copy number increases

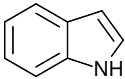
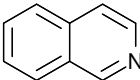
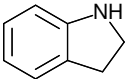
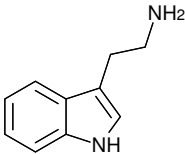
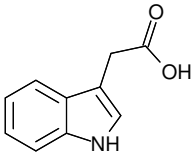
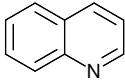

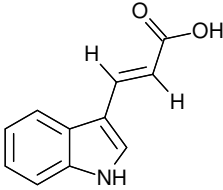
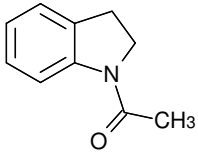
Chemical Name	Structure	Growth Inhibition	Plasmid Replication Inhibition
(b) Indole		✓	✓
(c) Isoquinoline		✓	✓
(d) Indoline		—	✓
(e) Tryptamine		—	—
(f) Indole-3-Acetic Acid		—	—
(g) Quinoline		✓	—
(h) Pyrrole		—	—
(i) 3-β-Indoleacrylic Acid		✓	✓
(j) 1-Acetylindoline		✓	?

Table 6.1: Chemical analogs of indole tested for their effects on cell growth and plasmid replication. Letters in the leftmost column refer to Figure 6.8. Solutions were prepared according to Table 2.5.4.

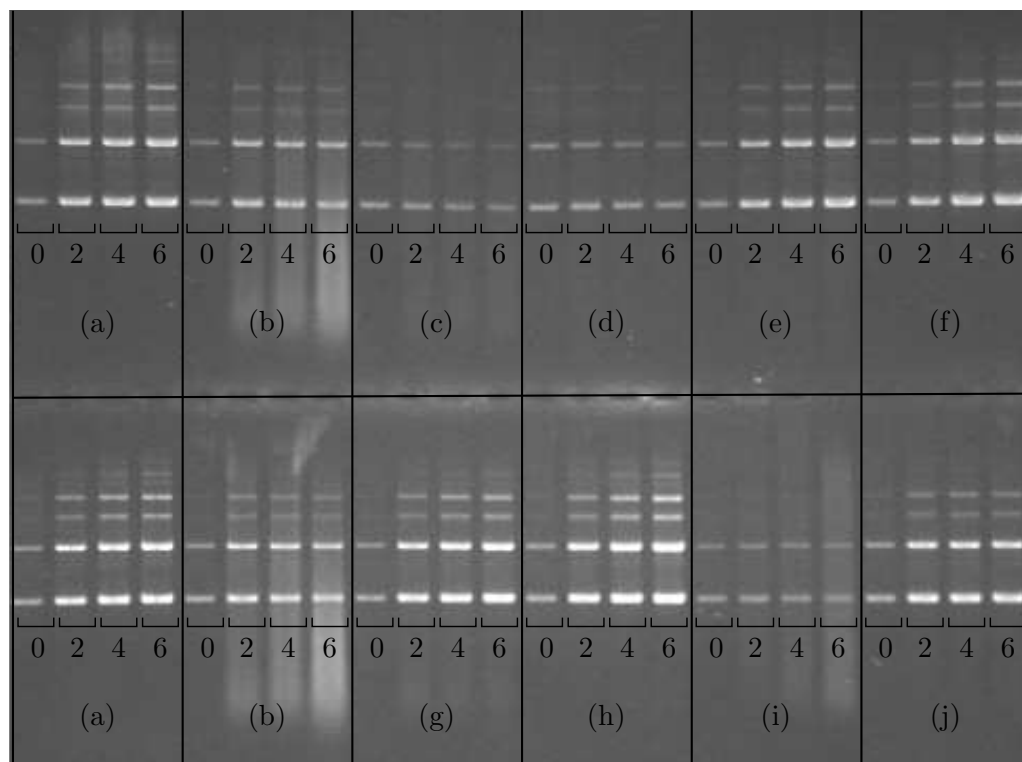


Figure 6.8: Plasmid DNA extracted from cells treated with chloramphenicol and different chemical analogs of indole; (a) ethanol, (b) indole, (c) isoquinoline, (d) indoline, (e) tryptamine, (f) indole-acetic-acid, (g) quinoline, (h) pyrrole, (i) 3- β -indoleacrylic acid, (j) 1-acetylindoline. The time at which each sample was taken is given beneath it, in hours.

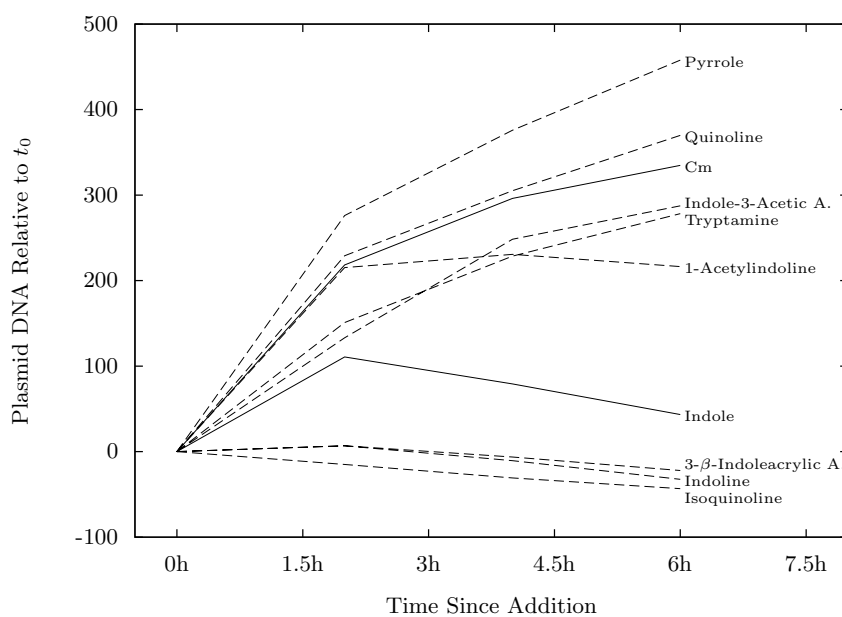


Figure 6.9: Change in DNA content of cells treated with different indole analogs. Values are in arbitrary units of brightness measured relative to t_0 .

in slower growing cells. Observations of this, as well as cells in stationary phase and those treated with chloramphenicol, indicate that plasmid replication is largely unaffected by slowed growth. This indirect effect cannot, therefore, assist dimer resolution, and the hypothesis that indole may have a direct effect on plasmid replication was investigated.

To separate the effect of indole on plasmid replication from its effects on cell division and growth, a broth culture of plasmid-carrying *E. coli* was treated with chloramphenicol. This caused amplification of the plasmid over time, as replication continued and plasmid DNA accumulated. However, in the presence of 5 mM indole, this effect was negated, such that the total amount of plasmid DNA remained roughly constant. This effect was also observed with the addition of 5 mM indole alone, eliminating the possibility that it might have inhibited a replication process specific to chloramphenicol-treated cells. Although it is not rigorously demonstrated here, preliminary experiments suggest that cells will recover from indole treatment with their plasmid replication restored (unpublished data).

A range of indole concentrations was tested with the same assay. The inhibition of plasmid replication was significant from a concentration of 3 mM or higher, though 1 mM seemed to actively enhance plasmid replication. The effect of indole on plasmid replication was also shown to favour neither monomers or dimers.

6.7.1 Indole Has Multiple Targets

Indole analogs were tested to observe their effects on plasmid replication, and to compare this with their effects on cell growth. Though isoquinoline and 3- β -indoleacrylic acid inhibited both plasmid replication and growth, indoline only affected plasmid replication, quinoline only affected cell growth and the results for 1-acetylintoline were unclear. Tryptamine, indole-3-acetic acid and pyrrole appeared to affect neither plasmid replication nor cell growth, though they were not necessarily taken up by the cell. That there are two different sets of responses for these structurally similar chemicals suggests that there might be two different targets involved, which are both inhibited by indole.

In fact, at least two targets for indole have been recently identified. Firstly, it has been shown to disrupt the ring formation of protein FtsZ (Pinero and Summers, in prep.), which could be responsible for inhibition of cell division. Secondly, *in vitro* experiments with a synthetic cell membrane suggest that the free diffusion of indole through the membrane increases its permeability to positive ions (Chimerel and Keyser, in prep.). If the membrane was made permeable to H^+ ions in this way, then the loss of membrane potential could severely disrupt ATP synthesis, which would explain the observed inhibition of cell growth.

There are a number of ways in which indole could interfere with the replication of ColE1. Perhaps it has DNA binding properties which modify the rate of RNAI and RNA II

transcription, or perhaps the RNAs themselves are being interfered with. There are also a number of potential targets in the replication machinery, and there is always the possibility that this effect is indirect, a symptom of indole interacting with an altogether different system.

6.7.2 A Modified Rcd Checkpoint Hypothesis

A modified hypothesis can be formed from the observations reported in this chapter. Increased indole production in the cell, in response to Rcd expression from plasmid dimers, directly inhibits plasmid replication, cell division and growth. The plasmid cannot be lost if the cell does not divide, which is made ‘safe’ by significantly reducing the rate at which new dimers are created. This makes it more likely that existing dimers will be resolved by *Xer-cer* recombination, and that the plasmid will be stable when cell division eventually occurs.

Without careful analysis of the system *in vivo*, it is difficult to estimate the time period over which this mechanism would operate. It may be that Rcd expression from just a single dimer causes a sufficient build-up of indole inside the cell to induce enough inhibition to aid resolution of that dimer. Alternatively, it may take several dimers to cause an arrest, preventing their accumulation or resolving all of them. Perhaps the system is designed only for dimer-only cells, reducing the probability of a sub-population establishing itself with substantially reduced plasmid stability.

There is also the question of whether an intracellular concentration of 3 mM indole (the point at which there is significant inhibition) is achievable. Measurement of indole concentration is very difficult at the intracellular level, but it is known that a broth culture has an extracellular concentration of 1 mM indole by the time it has reached stationary phase. Indole diffuses freely through the cell membrane (Chimerel and Keyser, in prep.), but cell volume is very small, such that the intracellular concentration of indole will be closely related to its rate of production. Thus a high intracellular concentration can be achieved quickly with a high production rate, but the indole will remain near undetectable in the large volume of the continuously agitated supernatant. This is much like a campfire producing smoke; close to the fire, the smoke is dense, but move away and it is unnoticeable in a breezy sky. Determining how much indole is produced by dimer-containing cell, and how significantly this affects the cell’s behaviour is an important focus for future work.

Having demonstrated that indole inhibits plasmid replication, the mechanism by which it does so will be investigated in the next chapter.

The Mechanism of Plasmid Replication Inhibition by Indole

7.1 Introduction

Analysis of the Rcd checkpoint in Chapter 6 suggested that the inhibition of cell growth and division would only improve plasmid stability if plasmid replication was also inhibited. It was demonstrated that indole was able to inhibit plasmid replication independently of its effects on cell division and growth.

This chapter concerns itself with the mechanism by which indole inhibits plasmid replication. The effect may or may not be specific to plasmid replication. Inhibition of all DNA replication could be achieved by targetting DNA polymerase I or III, as some rifamycin derivatives do (Frolova et al., 1977), or by allosteric inhibition of the process through DNA-binding, perhaps in the manner of chloroquine (O'Brien et al., 1966). Alternatively, inhibition of DNA replication could be more severe for a plasmid than for the chromosome, as caused by the inhibition of DNA gyrase by oxolinic acid (Gellert et al., 1977) or novobiocin (Hooper et al., 1984). Finally, it could be highly specific to the replication of ColE1-like plasmids, perhaps by modulating the conformation of RNAII *via* a riboswitch (reviewed in Winkler and Breaker, 2005).

For the purposes of the Rcd checkpoint, the preferential inhibition of plasmid replication, or the specific inhibition of ColE1-like plasmid replication, would seem preferable. The chemical structure of indole provides some clues as to how it might achieve this. Indole consists of a benzene ring fused with a nitrogen-containing pyrrole ring (Figure 7.1). Indole derivatives have been investigated for antibiotic activity – specifically, for their potential as inhibitors of DNA gyrase (Hubschwerlen et al., 1992; Oblak et al., 2005).

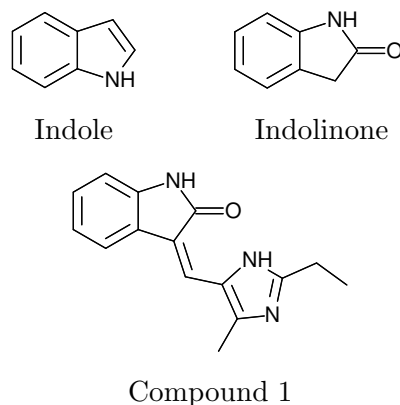


Figure 7.1: The chemical structures of indole, indolinone and compound 1 (Oblak et al., 2006).

Oblak et al. (2005) discovered, *in silico*, an indolinone derivative (compound 1; Figure 7.1) that was believed to be a potent inhibitor of DNA gyrase. A model for binding between this compound and the GyrB subunit of DNA gyrase was proposed (Oblak et al., 2006). It showed that the indole substructure of compound 1 fits into a hydrophobic pocket in the ATP-binding site, suggesting an ATP-competitive mechanism of inhibition. It is possible that indole alone could bind in the same way, though perhaps with a much weaker affinity. In this chapter, therefore, the effect of indole on DNA gyrase activity is investigated *in vitro*.

7.2 Indole Inhibits DNA Gyrase *in vitro*

DNA gyrase introduces negative supercoils to closed circular DNA in a reaction that requires ATP and Mg^{2+} and is further stimulated by spermidine (Gellert et al., 1976). Supercoiled DNA migrates through agarose gel by electrophoresis faster than relaxed DNA. A simple *in vitro* assay was thus designed to detect DNA gyrase activity by monitoring the increase in negative superhelicity of closed circular DNA (see Chapter 2 for details).

Initial assays were performed to determine the optimal quantity of enzyme and relaxed plasmid DNA to use such that the supercoiling reaction was completed over 1 to 2 hours and the resolution of the plasmid topoisomers in agarose gel was sufficient. Further, it was determined that the best way to stop the reaction, to allow subsequent electrophoresis of all samples simultaneously, was by heat inactivation of the enzyme at 65 °C for 20 minutes. The assay was then performed with the addition of 1 mM, 2 mM, 5 mM and 10 mM indole. The indole was dissolved in ethanol, and so a 1.3% ethanol (no indole) control was also tested, along with positive (no indole or ethanol), negative (no enzyme) and positive-for-inhibition (120 μM nalidixic acid) controls. The reaction mixture was sampled every 20 minutes for 100 minutes and the samples analysed by gel electrophoresis (Figure 7.2).

The positive control (b) shows that more of the DNA shifts towards the fastest-moving band over time, indicative of supercoiling activity. Each of the bands between the relaxed and fastest band represents an additional complete supercoil in the circular molecule, each increasing its speed in the gel until additional supercoils provide no further speed advantage. The negative control (a) shows that there is no supercoiling activity in the absence of DNA gyrase. The ethanol control (c) demonstrates that the indole solvent does not significantly inhibit the enzyme by itself. Finally the positive-for-inhibition control (h) shows no supercoiling activity, as nalidixic acid is known to inhibit the enzyme ([Sugino et al., 1977](#)).

For 1 mM indole (d), there is little difference in comparison to the positive and ethanol controls. However, from 2 mM indole (e) upwards, there is visibly less supercoiling activity. At 5 mM indole (f), it takes 60 minutes for the fastest band to emerge distinctly, and for 10 mM indole (g) there is no visible activity at all. Thus, indole appears to inhibit the supercoiling activity of DNA gyrase in this assay.

7.3 Indole Does Not React with ATP or the DNA Substrate

The inhibition of supercoiling activity in the *in vitro* assay above could be a result of an interaction between indole and either ATP or the DNA substrate, rather than gyrase, as was first proposed. To exclude the former possibilities, the supercoiling activity assay was performed with an additional pre-incubation step. A solution of 35.7 mM ATP and 10 mM indole was incubated at 37 °C for 1 hour. Similarly, 1 µg of relaxed, closed-circular pUC19 DNA was incubated in 10 mM indole under the same conditions. The remaining components of the reaction mix were then added to each, diluting the ATP to a standard concentration of 1.67 mM and the indole to a concentration of 0.47 mM, which is not high enough to inhibit the supercoiling activity of DNA gyrase. The reaction mixtures, along with a positive control, were sampled every 20 minutes for 100 minutes as before, and the samples analysed by gel electrophoresis (Figure [7.3](#)).

Although the positive control appears to have slightly greater supercoiling activity compared to the assays where ATP or DNA were pre-incubated with indole, there is nothing like the inhibition seen with 10 mM indole in the previous assay. Indole does not appear to be chemically reacting with either of these components, and if there is binding between them, then it is rapidly reversed by dilution. The most likely explanation for inhibition of supercoiling activity remains that which was suggested originally: indole is directly inhibiting DNA gyrase.

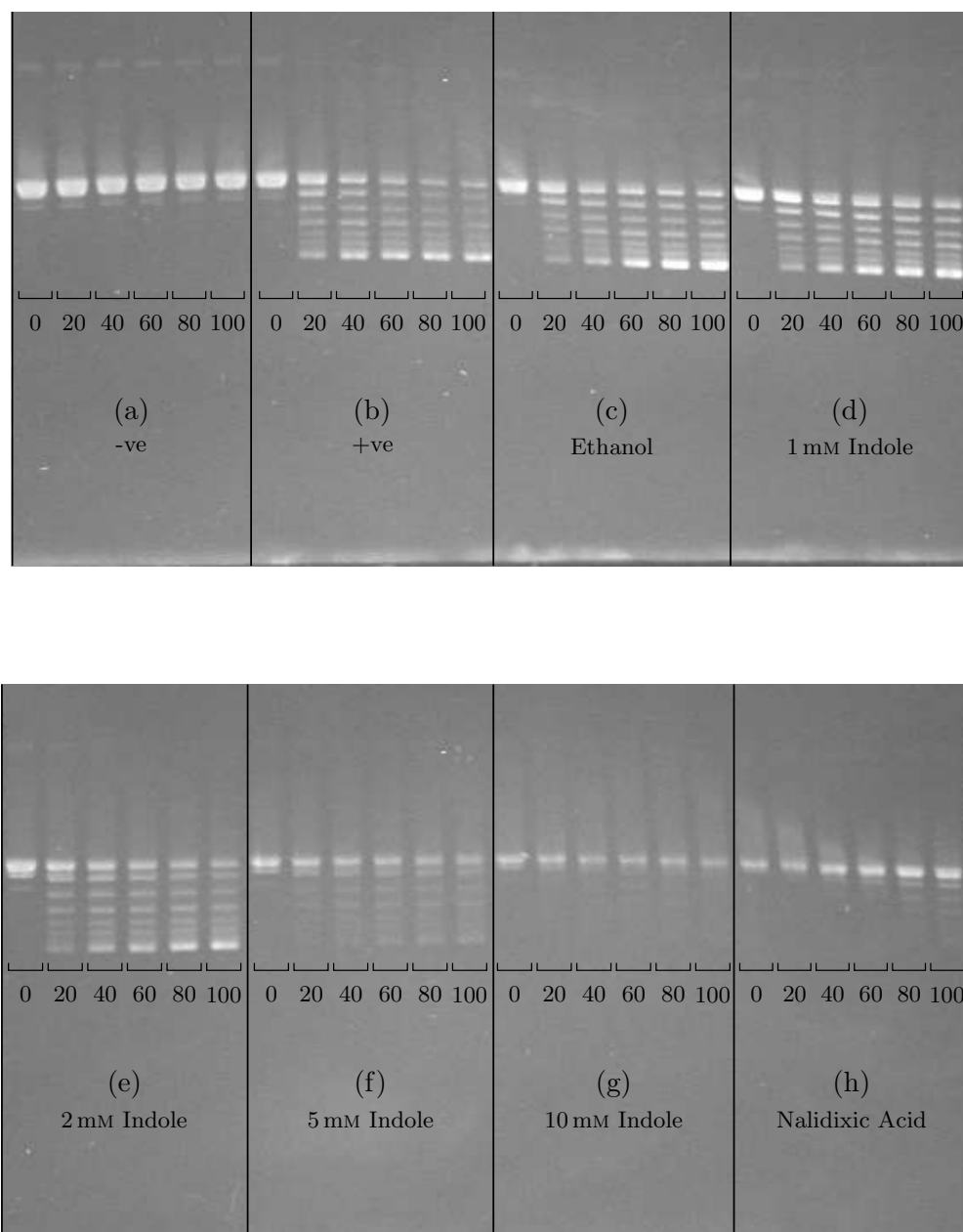


Figure 7.2: Assay for DNA gyrase supercoiling activity *in vitro*; (a) no enzyme, (b) no indole, (c) 1.3% ethanol, (d) 1 mM indole, (e) 2 mM indole, (f) 5 mM indole, (g) 10 mM indole, (h) 120 μ M nalidixic acid. The time at which each sample was taken is given beneath it, in minutes.

7.4 Concentration Dependence of DNA Gyrase Inhibition

The inhibition of DNA gyrase-mediated supercoiling by indole showed a concentration dependence in the initial assay. To further illuminate the nature of this dependence, an endpoint assay was performed with indole concentrations ranging from 0 mM to 10 mM, with a separate assay for every 1 mM increment. Positive (no indole), negative (no enzyme) and positive-for-inhibition (120 μ M nalidixic acid) controls were also tested. A single sample was taken from each reaction after 60 minutes and analysed by gel electrophoresis (Figure 7.4).

As the indole concentration is increased, there is evidence of reduced supercoiling activity, as the slowest band (relaxed substrate DNA) is brighter and the faster bands (supercoiled DNA) are less bright. It has hard to discern much of a difference among the assays with 1 mM, 2 mM and 3 mM indole, but the fastest band is distinctly thinner at 4 mM indole and barely visible at 5 mM indole. At around 7 mM or 8 mM indole, the fastest band disappears altogether, and the intermediate bands begin to weaken and disappear at higher concentrations. For 10 mM indole and the nalidixic acid control, only the first supercoiled band is distinct, so there is still slight gyrase activity.

The kinetics of DNA gyrase inhibition by indole could be discerned if supercoiling activity could be quantified in this assay. This is theoretically possible, by counting the number of discrete supercoils introduced to the substrate DNA over time. Each additional supercoil increases the migration speed of an individual closed-circular molecule such that it appears in a separate band on the agarose gel. However, more supercoils provide diminishing increases in speed; the bands become closer together and eventually indistinguishable as molecules reach a maximum migration speed. Attempts to measure band density, and hence supercoiling activity, did not produce reliable results for this reason.

7.5 The Inhibition of DNA Gyrase by Indole is Reversible

7.5.1 Pre-Incubation of Indole with DNA Gyrase

The inhibition of DNA gyrase by indole could be due to either permanent modification of the enzyme or a reversible interaction with it. To investigate the reversibility of the inhibition, DNA gyrase was incubated at 37 °C for 1 hour in sterilised distilled water with (a) 1.3% ethanol, (b) 1.3% ethanol and 50 mM ATP, (c) 5 mM indole and finally (d) 5 mM indole and 50 mM ATP. The remaining components of the reaction were then added to each, diluting the ATP to a standard concentration of 1.67 mM and the indole, where present, to a concentration of 0.17 mM, low enough to have no inhibitory effect in the assay. The reaction mixtures were sampled every 20 minutes for 100 minutes as before, and the samples analysed by gel electrophoresis (Figure 7.5).

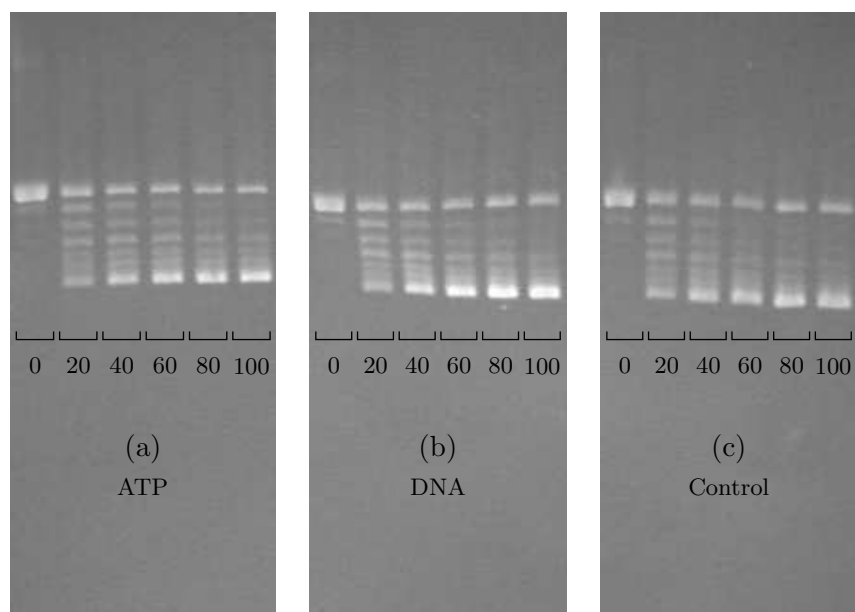


Figure 7.3: Supercoiling activity assay with pre-incubation step; (a) indole pre-incubated with ATP, (b) indole pre-incubated with the DNA substrate, (c) no pre-incubation control. The time at which each sample was taken is given beneath it, in minutes.

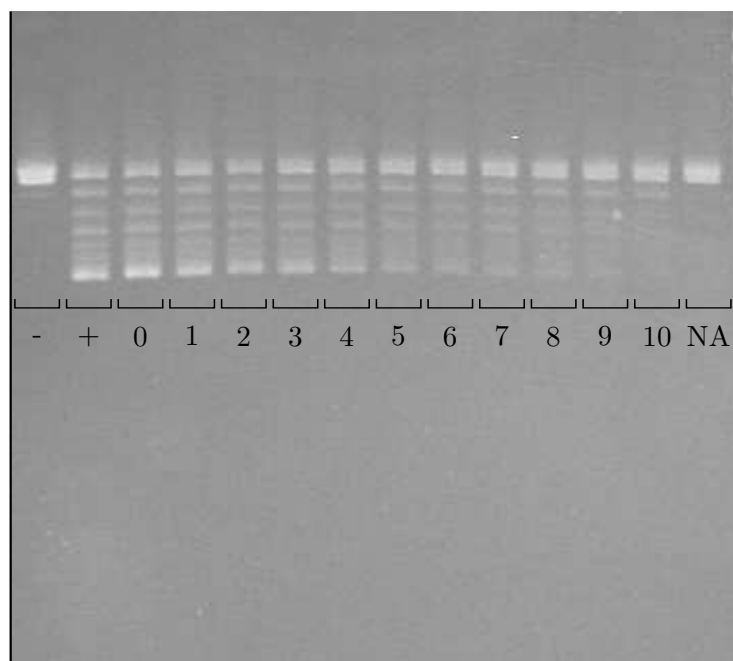


Figure 7.4: Concentration dependence of DNA gyrase inhibition by indole; from left to right: no enzyme, no indole, 0 mM to 10 mM indole in 1 mM increments, 120 μ M nalidixic acid.

There is little difference among the reactions, though each shows less activity than seen in previous assays, likely due to some denaturation of the enzyme during the incubation period. The reaction in which DNA gyrase was pre-mixed with 5 mM indole (c) appears to have the least supercoiling activity. However, the difference is small in comparison to the reaction in which DNA gyrase was mixed with ethanol alone (a). The presence of ATP, with (b) or without (d) indole, does not change this. This suggests that the inhibition is not due to permanent damage of the enzyme by indole, at least in the absence of DNA.

7.5.2 Recovery of Supercoiling Activity After Indole Treatment

To demonstrate more directly that the inhibition of DNA gyrase by indole is reversible, a modified version of the supercoiling assay was performed. Three reactions were set up; one as a positive control (a) and two with 10 mM indole (b, c). Samples were taken at the start of the assay and after 30 minutes, at which point reactions (a) and (b) were diluted five-fold with enzyme buffer, and reaction (c) was diluted five times with enzyme buffer containing 10 mM indole. Reaction (a) was therefore still a positive control, (b) was the test case in which indole had been diluted to 2 mM, and (c) the negative control in which the indole concentration remained at 10 mM. Larger volume samples, to compensate for the dilution of the DNA substrate, were taken after 60 minutes and 120 minutes and all samples were concentrated to 5 μ l using a heated vacuum centrifuge before analysis by gel electrophoresis (Figure 7.6).

The positive control (a) demonstrates that there has been no drastic effect on the reaction due to dilution or sample concentration. The second samples of reactions (b) and (c) show that 10 mM indole is inhibiting the supercoiling activity as expected. The third and fourth samples of reaction (b) show greater supercoiling activity than those of reaction (c), suggesting that the dilution of indole from 10 mM to 2 mM has restored supercoiling activity. Along with the pre-incubation assay, this indicates that the inhibition of DNA gyrase by indole is reversible.

7.6 Indole Does Not Cause Double-Strand Breaks

The supercoiling activity of DNA gyrase involves the breakage and reunion of both polynucleotide strands, mediated by the GyrA subunits (Sugino et al., 1977). One class of gyrase inhibitors, which includes oxolinic acid and nalidixic acid, causes double-strand breaks (Gellert et al., 1977) and is thought to target the GyrA subunit (Snyder and Drlica, 1979). The GyrB subunit is responsible for the ATPase activity required for supercoiling (Sugino et al., 1978). A second class of gyrase inhibitors, which includes novobiocin and coumermycin, does not cause double-strand breaks, but instead inhibits ATPase activity, and is therefore thought to target the GyrB subunit (Sugino et al., 1978).

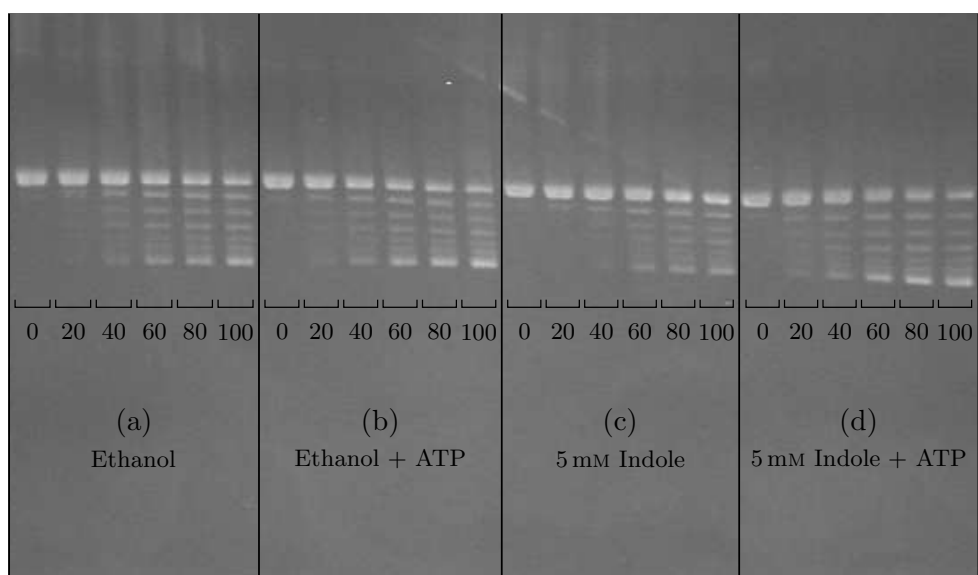


Figure 7.5: Supercoiling assay in which DNA gyrase is pre-incubated with: (a) 1.3% ethanol, (b) 1.3% ethanol and 50 mM ATP, (c) 5 mM indole, (d) 5 mM indole and 50 mM ATP. The time at which each sample was taken is given beneath it, in minutes.

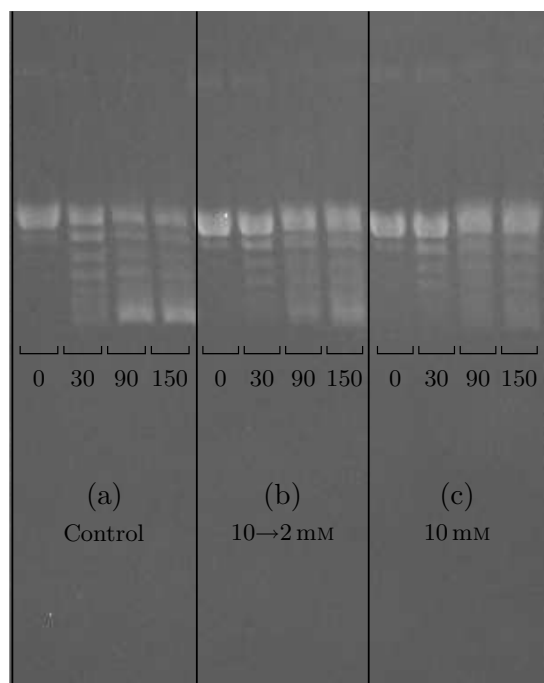


Figure 7.6: Recovery of DNA gyrase supercoiling activity; (a) no indole, (b) 10 mM indole diluted down to 2 mM after 30 minutes, (c) 10 mM indole. The time at which each sample was taken is given beneath it, in minutes.

To illuminate the mechanism of DNA gyrase inhibition by indole, the potential for double-strand breakage was investigated. If inhibition by indole introduced double-strand breaks into the DNA substrate, they would be located more-or-less randomly. Any plasmid molecules that had been broken by the activity of gyrase in the presence of indole could be detected by digesting the plasmid DNA with a restriction enzyme at a unique location. If they were cut into two or more pieces, then multiple distinct bands or a smear of DNA would be detected after gel electrophoresis, suggesting that indole interacts with the GyrA subunit. If instead, there was no double-strand breakage, then only a single band would be detected, and it would be more likely that indole was interacting with the GyrB subunit.

The supercoiling activity assay was performed as before, with and without 10 mM indole. Prior to gel electrophoresis, the samples were treated with restriction enzyme PvuII at 37°C for 60 minutes. For both reactions, just a single band of the correct size for the plasmid was found in every sample (data not shown). This suggests that indole does not cause double-strand breaks and indicates that it is more likely to be interacting with the GyrB subunit, consistent with the original proposal that indole might inhibit DNA gyrase.

7.7 The Kinetics of DNA Gyrase Inhibition by Indole

It has been demonstrated that the inhibition of DNA gyrase by indole is reversible. There are three major types of reversible inhibition: competitive, uncompetitive and noncompetitive (Lehninger, 1970). A competitive inhibitor competes with the normal substrate for free enzyme and the two are typically similar in structure. An uncompetitive inhibitor can only bind to the enzyme-substrate complex to inactivate it. A noncompetitive inhibitor can bind the enzyme at the same time as the substrate, and hence inactivates both free enzyme and the enzyme-substrate complex. Competitive inhibition can be overcome by the addition of excess substrate, unlike the other two forms of inhibition.

With the assumption that indole targets the GyrB subunit, the kinetics of its inhibition of DNA gyrase were investigated. The supercoiling assay was performed with 0.13 mM, 0.67 mM, 1.3 mM and 6.7 mM ATP in place of the normal 1.75 mM ATP. 5 mM indole was used to inhibit the supercoiling activity in each case. Samples were taken as usual and analysed by gel electrophoresis (Figure 7.7).

As the amount of ATP in the reaction increases, the level of supercoiling activity increases, for the first three reactions. However, at 6.7 mM ATP, there appears to be complete inhibition of the enzyme. This is surprising, as excess substrate should sustain maximum enzyme activity for longer, according to Michaelis-Menten kinetics. The result was confirmed with a different stock solution of ATP and the pH of the reaction mixture was found to be unchanged. Further, a repeat of the assay with 6.7 mM ATP and no indole still showed inhibition of supercoiling activity (data not shown).

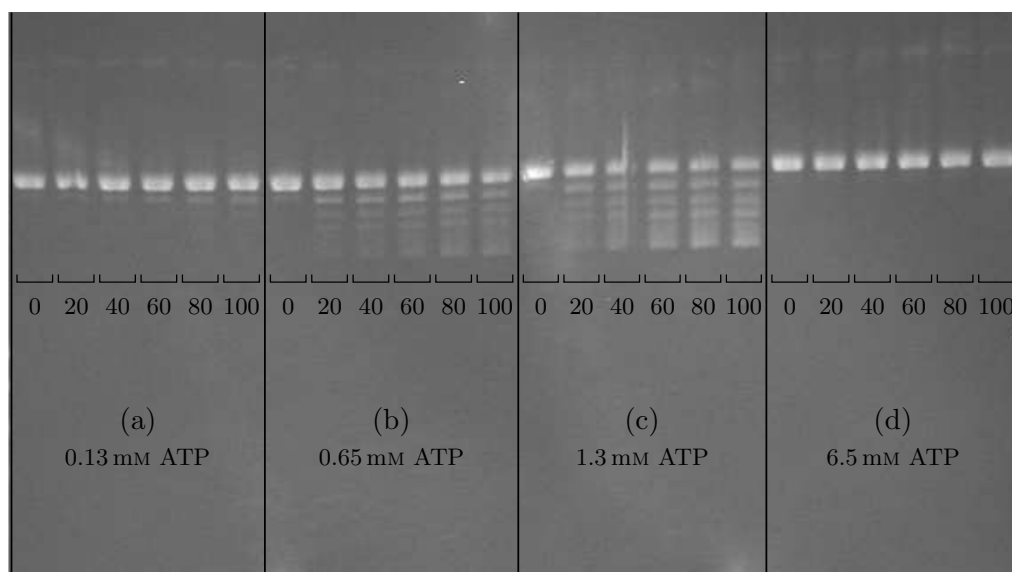


Figure 7.7: Supercoiling activity assay in the presence of 5 mM indole and increasing concentrations of ATP; (a) 0.13 mM ATP, (b) 0.65 mM ATP, (c), 1.3 mM ATP, (d) 6.5 mM ATP. The time at which each sample was taken is given beneath it, in minutes.

[Ali et al. \(1993\)](#) demonstrated that ATP hydrolysis by DNA gyrase does not follow Michaelis-Menten kinetics, which may explain the result for the reactions with 6.7 mM ATP. Unfortunately, this makes it very difficult to determine what form of inhibition indole imposes on the enzyme. A more accurate method for measuring supercoiling activity, perhaps by detection of ATPase activity, could help in characterising the behaviour of DNA gyrase, even if the kinetics are unusual.

7.8 Conclusions

7.8.1 Indole Inhibits DNA Gyrase

In the previous chapter, it was demonstrated that indole inhibits plasmid replication, independently of its effects on cell division and growth. The mechanism by which this might be achieved was considered. Indole is structurally similar to indolinone, a chemical from which DNA gyrase inhibitors have been derived. A model of the interaction between such an inhibitor and the GyrB subunit of the enzyme suggested that an indole structure could fit into a hydrophobic pocket in the ATPase site *in silico* ([Oblak et al., 2005](#)).

Therefore, the effects on indole on the supercoiling activity of DNA gyrase were investigated *in vitro*. It was found that increasing concentrations of indole reduced the rate of supercoiling of a closed-circular DNA substrate by the enzyme. In subsequent experi-

ments, indole did not appear to damage either the ATP necessary for enzyme activity or the DNA substrate itself, suggesting that DNA gyrase itself was the target for inhibition. The concentration dependence of DNA gyrase inhibition by indole was more fully demonstrated in a further experiment. Whilst 10 mM indole was required to completely inhibit supercoiling over the 100 minute time period of the assay, there was visible inhibition with a concentration of just 2 mM indole.

From the data presented here, DNA gyrase inhibition is the strongest candidate for the cause of plasmid replication inhibition by indole. Certainly, plasmid replication appears more sensitive to gyrase inhibition than chromosomal replication and cell survival (Hooper et al., 1984; Wolfson et al., 1982) and unpublished work in this laboratory correlates with this, suggesting that replication of the chromosome is less inhibited, if at all, by indole treatment. However, it is difficult to compare the *in vitro* concentration dependence of inhibition in this chapter with the *in vivo* data of the previous chapter. Future work in this area should look to identify a gyrase mutant in which plasmid replication continues in the presence of indole, as this would provide a formal demonstration of the mechanism.

7.8.2 Indole Probably Targets GyrB

The mechanism of DNA gyrase requires double-strand breakage to allow the passing of one strand through another, decreasing the linking number of a closed-circular molecule by 2 (see Chapter 1). Those inhibitors that target the GyrA subunit prevent reunion of the broken strand (Gellert et al., 1977). In contrast, those inhibitors that target GyrB are thought to inhibit ATPase activity (Sugino et al., 1978). The original reasoning that led to the investigation of DNA gyrase as a target for indole suggested that it might bind a hydrophobic pocket in the GyrB subunit. So, as the inhibition of DNA gyrase supercoiling activity by indole was not found to cause double-strand breaks in the substrate DNA, it is probable that indole does indeed target GyrB.

Inhibitors of DNA gyrase that target GyrB compete for binding with ATP. The hallmark of competitive inhibition is that enzyme activity can be restored with the addition of excess substrate. Experiments to investigate whether DNA gyrase supercoiling activity in the presence of indole could be restored by the addition of ATP produced interesting results. When the concentration of ATP was increased five- or ten-fold in a supercoiling assay inhibited by 5 mM indole, there was increased enzyme activity. However, with fifty times the ATP there was no activity at all. Ali et al. (1993) demonstrated that ATP hydrolysis by DNA gyrase does not follow Michaelis-Menten kinetics, which may explain this result. A method to better quantify the activity of DNA gyrase would help characterise the behaviour of the enzyme in the presence of indole, but the evidence so far suggests competitive inhibition over any other form.

Discussion

8.1 Computer Modelling

At the outset of this work, a principal aim was to update and improve the computer model that simulates a cell population undergoing a dimer catastrophe. Technical improvements, such as maintaining a constant population instead of regularly subculturing it, really only improved accuracy and made the results more robust to external inspection. The most significant change by far was the introduction of realistic copy number distribution, incorporated with a stochastic plasmid replication model.

8.1.1 The Role of the Rcd Checkpoint

In previous models ([Summers et al., 1993](#), MEng thesis, Field, 2006), and that of Chapter [3](#), it was assumed that dimers would have half the copy number of monomers, and for convenience, that the control system would always achieve the ideal copy number for each. So a plasmid in a monomer-only cell with a nominal copy number of 30 would be over 30,000 times more stable than a plasmid in a dimer-only cell with a copy number of 15. A population containing only 2.3% dimer-only cells (the proportion estimated *in vivo* by [Summers et al. \(1993\)](#)) suffers a plasmid stability over 750 times less than that of a monomer-only population.

In contrast, when copy number is variable, as in the model of Chapter [5](#), the difference in plasmid stability between a monomer-only and a dimer-only cell is less than 1,000 times. A plasmid in a population with no dimers is less than 10 times more stable than in a population containing 2.3% dimer-only cells. It is worth acknowledging that adjusting the parameters of the plasmid replication model, which were based upon data taken from just a small set of *in vivo* experiments in slow-growing cells, can change the ratio of average

copy numbers in monomer-only and dimer only-cells. Nonetheless, the impact of dimers is significantly reduced compared to the previous, simpler model.

As the presence of dimers in a population reduces the average plasmid stability by only a small amount, then perhaps the Rcd checkpoint solves problems beyond just copy number depression. In this sense, it may be that the presence of dimers in a population is not catastrophic in itself, but could lead to a worse situation. The current computer model does not take into account higher-order multimers. One would expect trimers and tetramers to over-replicate compared to dimers, perhaps imposing an intolerably high metabolic load on the host. The Rcd checkpoint may assist in the resolution of dimers to prevent the emergence of such plasmid species, as well as to slightly improve plasmid stability.

It is not known whether the Rcd checkpoint is activated upon the emergence of a single dimer, or in the presence of many. From the point of view of a single cell, a single dimer poses little immediate threat to the stability of the plasmid. Further, it is not clear whether or not a single dimer can effect a significant enough change in the production rate of indole to invoke the checkpoint. It may be that the checkpoint only exists as an emergency system for cells that contain mostly, or all, dimers. This mechanism, the dimer contingency, would effectively act to prevent segregation between dimers and monomers. Whilst a few dimers in the population are tolerable, allowing them to persist in their own cell line would be bad for both plasmid and host in the long run, despite any metabolic penalties a dimer-only cell might suffer.

8.1.2 Population-Wide Effects

As the computer model simulates plasmid behaviour in a population of cells, with results averaged across multiple runs, the effects of changing various parameters are considered in a very wide context. From the model, it is possible to increase the stability of the plasmid in the whole population, and the time it takes to reach this improved state, by increasing the metabolic penalty associated with dimers. Whether this penalty is due to their over-replication relative to monomers, the Rcd checkpoint, or some other, unknown mechanism, does not matter. However, the effect on an individual cell must be considered. If dimers were heavily penalised, such that the plasmid stability of the population is high and achieved in just a few hours, then an individual dimer-only cell may not divide for several monomer-only cell generations. Is it reasonable to expect such a disadvantageous system to have been selected for? Even if the penalty for plasmid loss is death, it is still a rare occurrence, and surely the penalty for delaying the division of a percentage of the cell population, some for many generations, is more severe in the long run? Having said this, there exist systems in which bacterial altruism has been shown to exist, so perhaps a cell with many dimers prevents future competition from its potential plasmid-free descendants by inhibiting its own division. Given enough time, and enough dimer resolution, it could even rejoin the dividing population.

It is interesting that this scenario might be achieved by the use of indole, as recent work by Lee et al. (2010) suggests that it may play a role in an altruistic system that increases the overall antibiotic-resistance of a bacterial cell population. There may be some additional advantage to removing a particular cell from the dividing population if the indole it produces also protects its neighbours from environmental hazards. On the other hand, perhaps the indole it produces also afflicts a cell's neighbours; if a cell has to stop dividing to resolve its plasmid dimers, it could reduce the competitive disadvantage of doing so by also slowing its immediate competitors. Turning that around once again, the neighbours of a dimer-containing cell are more likely to contain dimers themselves, at least in non-broth culture, so a small group of cells may act in unison to achieve the indole levels required to allow for efficient dimer resolution. Again, characterisation of the relationship between the number of dimers in a cell and the subsequent impact on its division and growth is required.

8.1.3 Future Work

In terms of the computer model, there are plenty of minor improvements that could be made. From the technical side, making an individual simulation able to run across multiple processors and machines would save CPU time. So far, each simulation requires a separate machine, though the results from many machines can be analysed afterwards. There is also more information about the behaviour of the cell population that could be extracted from the simulation, such as the number of plasmid-free cells generated or the average lifetime of monomer-only and dimer-only cells before they are removed by the population-limiting algorithm.

In terms of improving the accuracy of the model, the most important future work would be to obtain more accurate estimates of the rate parameters involved in plasmid replication, particularly in fast-growing cells. Knowing how those parameters vary with generation time would also be useful. These parameters determine not only the average plasmid copy number, but also the copy number distribution and the copy number ratio for monomer-only to dimer-only cells. Further, it would be interesting to know if the metabolic load imposed by increasing copy number inhibits plasmid replication in the same way as it does growth. In other words, is it a global effect or does it just affect the activity of certain enzymes. Some considerable, but ultimately unsuccessful, effort was made as part of this work to determine the rate of dimer resolution *in vivo*; determination of this parameter would allow for characterisation of the Rcd checkpoint *in silico*. With some of this information, it would be possible to include the Rcd checkpoint in the plasmid replication model itself, and an improved estimate of its effectiveness could be made. The model could also be adapted to estimate the rate of plasmid-free cell accumulation rather than rejecting them from the population. The framework could also be easily modified to simulate a system with two incompatible plasmids, or to include higher-order multimers beyond dimers.

8.2 Indole and the Rcd Checkpoint

The role of indole in the Rcd checkpoint, and in bacterial control generally, has been illuminated by this and other recent work. A picture has emerged of a small molecule with multiple cellular targets and a subsequently wide range of effects. Here, it has been shown that indole inhibits plasmid replication, probably by inhibiting the supercoiling activity of DNA gyrase. Chimere and Keyser (in prep.) have demonstrated that indole can diffuse freely through a cell membrane in such a way as to increase its permeability to positive ions. This mechanism could destroy a cell's membrane potential, disrupting ATP synthase activity, with far-reaching consequences. Pinero and Summers (in prep.) have identified a second protein target for indole, FtsZ, showing that it disrupts the ring formation required for successful cell division. The involvement of all these processes in a plasmid stability function is novel, and perhaps extreme, but inhibition of all three of cell division, growth and plasmid replication is required for the checkpoint to increase plasmid stability. Of course, indole also plays a role during stationary phase growth, and all these effects may be useful to plasmid-free cells in this context, though it is not clear how they modulate their indole production.

8.2.1 The Modified Rcd Checkpoint Hypothesis

Our present view of the Rcd checkpoint can be summarised as follows. When a dimer is formed through homologous recombination in a cell, it will transcribe Rcd. As the Rcd concentration builds up, it will modify the behaviour of the enzyme tryptophanase, increasing intracellular indole production. A raised level of indole production will increase its local concentration, inhibiting the activity of DNA gyrase and FtsZ, and disrupting the cell membrane. This will in turn inhibit plasmid replication, cell division, and cell growth itself. With the cell effectively in stasis, only dimer resolution will reduce the rate of Rcd production, reducing indole production and alleviating the three forms of inhibition. Built-up indole will diffuse freely away from the cell, and remaining Rcd will be diluted as the cell begins to grow again. The plasmid will ideally be entirely monomeric at this stage, such that when the cell divides it can be distributed to daughter cells with a low risk of plasmid loss.

An interesting consequence of inhibiting plasmid replication is that it may assist the dimer resolution system. During the passage of a replication fork, the XerCD-PepA-ArgR complex will be disassembled from the *cer* site. This will delay the formation of a synaptic complex between two *cer* sites until the nucleoprotein complex can reassemble. If plasmid replication is stalled, then the complex will not have to deal with this interference and dimer resolution may proceed faster. Additionally, inhibition of DNA gyrase allows for preferential inhibition of plasmid replication. As discussed in Chapter 7, an appropriate concentration of DNA gyrase inhibitor can cure a plasmid without killing the host strain.

Allowing the chromosome to continue to replicate minimises the long-term damage caused by the checkpoint; the cell will be in a position to divide soon after the checkpoint is cleared. Unpublished work in this laboratory has demonstrated that cells can divide faster during recovery from indole treatment than when allowed to grow normally. It is thought that continued chromosome replication whilst cell division is delayed, and the slow filamentation of indole-treated cells allow for these subsequent rapid division cycles.

8.2.2 Indole Production and Concentration

Characterisation of the relationships between the number of dimers in a cell, the consequent Rcd concentration, the production rate of indole and the inhibition of cell division, growth and plasmid replication is essential to a full understanding of the behaviour of the checkpoint. One of the remaining key questions is exactly what concentration of indole is achieved in a cell undergoing a dimer catastrophe. This has been difficult to estimate to-date as indole is difficult to detect, can freely diffuse from the cell and because the volume of a single cell is so small compared to that of the broth the indole diffuses into. It is hoped that the indole production rate of a wild-type strain of *E. coli* with and without Rcd expression can be estimated by measuring the gross level in broth culture at a high enough cell density for accurate measurement, although expression of tryptophanase is likely modulated by the growth phase of such a culture. From this, the internal indole concentration for a single cell can be estimated.

A rate estimate for indole production in the presence of Rcd is critical to demonstrating that a high enough indole concentration can be achieved inside the cell, so as to cause inhibition of plasmid replication, cell division and growth. A improved method of inducing Rcd expression in the cell would be useful in this regard, as current methods involve an undesirable temperature shift. The addition of exogenous indole, though it has been an effective technique, may be amplifying the membrane effects over those on FtsZ or gyrase.

8.2.3 Future Work

Most important for future work in this area is establishing a formal connection between the DNA gyrase inhibition seen in Chapter 7 with the inhibition of plasmid replication seen in Chapter 6. A strain with a mutant, indole-resistant gyrase would be ideal for this. The *in vitro* work on the inhibition of DNA gyrase by indole can also be expanded upon. Experiments to characterise the kinetics of the enzyme in the presence of indole would be useful, even though these kinetics may not be Michaelis-Menten. Also, though it is unlikely, it would be good to definitively demonstrate that indole is not binding to the DNA substrate; a simple assay with the addition of more DNA could achieve this.

As many inhibitors of DNA gyrase also target topoisomerase IV, there exists the possibility that indole does the same. Experiments in the style of the supercoiling assays of Chapter 7 could be used to determine whether indole inhibits the decatenation and relaxation activity of topoisomerase IV. Determining the indole concentration required for inhibition of each enzyme would give insight into the topological arrangement of plasmids in a cell that has reached the Rcd checkpoint. Do plasmids stop replicating before or after the products of replication are prevented from decatenating? Is this a necessary component of the checkpoint, or a side effect of targetting DNA gyrase?

With regards to the inhibition of plasmid replication, it would be interesting to observe the effect of indole on plasmids that do not use a ColE1-like mechanism. This may require a redesign of the chloramphenicol assay however, as many plasmids require protein expression to replicate. However, if DNA gyrase inhibition is responsible, it will likely affect other plasmids equally. A more careful assay to determine whether there is preferential inhibition of dimer over monomer replication, perhaps when in the same cell, would also be interesting. Finally, a more accurate assessment of plasmid stability would be exceptionally useful. With current methods, it is hard to distinguish between the rate of plasmid-free cell occurrence and the accumulation of plasmid-free cells due to their metabolic advantage. Perhaps a system could be calibrated such that there is a non-lethal level of antibiotic to inhibit the growth of plasmid-free cells, and the metabolic load imposed by a plasmid is perfectly counteracted by the antibiotic resistance it provides. The rate of plasmid-free cell occurrence could then be measured independently, and the copy number distribution estimated from the model of Chapter 5 – though only for ColE1-like plasmids which have been well characterised.

References

- Abeles, A. L., Friedman, S. A., and Austin, S. J. (1985). Partition of unit-copy mini-plasmids to daughter cells III: the DNA sequence and functional organization of the P1 partition region. *Journal of Molecular Biology*, 185:261–272.
- Achtman, M., Schwuchow, S., Helmuth, R., Morelli, G., and Manning, P. A. (1978). Cell-cell interactions in conjugating *Escherichia coli*: Con^- mutants and stabilization of mating aggregates. *Molecular & General Genetics*, 164:171–183.
- Adams, D. E., Shekhtman, E. M., Zechiedrich, E. L., Schmid, M. B., and Cozzarelli, N. R. (1992). The role of topoisomerase IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication. *Cell*, 71:277–288.
- Alén, C., Sherratt, D. J., and Colloms, S. D. (1997). Direct interaction of aminopeptidase A with recombination site DNA in Xer site-specific recombination. *The EMBO journal*, 16:5188–5197.
- Ali, J. A., Jackson, A. P., Howells, A. J., and Maxwell, A. (1993). The 43-kilodalton N-terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds coumarin drugs. *Biochemistry*, 32:2717–2724.
- Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999). Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *Journal of Bacteriology*, 181:6361–6370.
- Arciszewska, L. K. and Sherratt, D. J. (1995). Xer site-specific recombination *in vitro*. *The EMBO Journal*, 14:2112–2120.
- Ataai, M. M. and Shuler, M. L. (1986). Mathematical Model for the Control of ColEI Type Plasmid Replication. *Plasmid*, 16:204–212.

- Atlung, T., Christensen, B. B., and Hansen, F. G. (1999). Role of the rom protein in copy number control of plasmid pBR322 at different growth rates in *Escherichia coli* K-12. *Plasmid*, 41:110–119.
- Austin, S. J., Ziese, M., and Sternberg, N. (1981). A novel role for site-specific recombination in maintenance of bacterial replicons. *Cell*, 25:729–736.
- Bachmann, B. J. (1972). Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriological Reviews*, 36:525–557.
- Baird, C. L., Harkins, T. T., Morris, S. K., and Lindsley, J. E. (1999). Topoisomerase II drives DNA transport by hydrolyzing one ATP. *Proceedings of the National Academy of Sciences*, 96:13685–13690.
- Balding, C., Blaby, I. K., and Summers, D. K. (2006). A mutational analysis of the ColE1-encoded cell cycle regulator Rcd confirms its role in plasmid stability. *Plasmid*, 56:68–73.
- Barillà, D., Rosenberg, M. F., Nobbmann, U., and Hayes, F. (2005). Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *The EMBO Journal*, 24:1453–1464.
- Bastia, D. (1978). Determination of restriction sites and the nucleotide sequence surrounding the relaxation site of ColE1. *Journal of Molecular Biology*, 124:601–639.
- Bellon, S., Parsons, J. D., Wei, Y., Hayakawa, K., Swenson, L. L., Charifson, P. S., Lippke, J. A., Aldape, R., and Gross, C. H. (2004). Crystal structures of *Escherichia coli* topoisomerase IV ParE subunit (24 and 43 kilodaltons): a single residue dictates differences in novobiocin potency against topoisomerase IV and DNA gyrase. *Antimicrobial Agents*, 48:1856–1864.
- Betlach, M. C., Hershfield, V., Chow, L., Brown, W., Goodman, H., and Boyer, H. W. (1976). A restriction endonuclease analysis of the bacterial plasmid controlling the EcoRI restriction and modification of DNA. *Federation Proceedings*, 35:2037–2043.
- Birnboim, H. and Doly, J. (1979). A rapid alkaline procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7:1513–1523.
- Blaby, I. K. and Summers, D. K. (2009). The role of FIS in the Rcd checkpoint and stable maintenance of plasmid ColE1. *Microbiology*, 155:2676–2682.
- Blakely, G., Colloms, S. D., May, G., Burke, M., and Sherratt, D. J. (1991). *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. *The New Biologist*, 3:789–798.
- Blakely, G., May, G., McCulloch, R., Arciszewska, L. K., Burke, M., Lovett, S. T., and Sherratt, D. J. (1993). Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. *Cell*, 75:351–361.

- Boe, L., Gerdes, K., and Molin, S. (1987). Effects of genes exerting growth inhibition and plasmid stability on plasmid maintenance. *Journal of Bacteriology*, 169:4646–4650.
- Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heynker, H. L., Boyer, H. W., Crosa, J. H., and Falkow, S. (1977). Construction and characterization of new cloning vehicles II: a multipurpose cloning system. *Gene*, 2:95–113.
- Botchan, P., Wang, J. C., and Echols, H. (1973). Effect of circularity and superhelicity on transcription from bacteriophage lambda DNA. *Proceedings of the National Academy of Sciences*, 70:3077–3081.
- Boyd, A. C., Archer, J. A., and Sherratt, D. J. (1989). Characterization of the ColE1 mobilization region and its protein products. *Molecular & General Genetics*, 217:488–498.
- Boyer, H. W. (1971). DNA restriction and modification mechanisms in bacteria. *Annual Review of Microbiology*, 25:153–176.
- Bradley, D. E. (1980). Morphological and serological relationships of conjugative pili. *Plasmid*, 4:155–169.
- Bradley, D. E. (1984). Characteristics and function of thick and thin conjugative pili determined by transfer-derepressed plasmids of incompatibility groups I1, I2, I5, B, K and Z. *Journal of General Microbiology*, 130:1489–1502.
- Bremer, H. and Lin-Chao, S. (1986). Analysis of the physiological control of replication of ColE1-type plasmids. *Journal of Theoretical Biology*, 123:453–470.
- Brendel, V. and Perelson, A. S. (1993). Quantitative model of ColE1 plasmid copy number control. *Journal of Molecular Biology*, 229:860–872.
- Brenner, M. and Tomizawa, J.-I. (1991). Quantitation of ColE1-encoded replication elements. *Proceedings of the National Academy of Sciences*, 88:405–409.
- Brown, P. O. and Cozzarelli, N. R. (1979). A sign inversion mechanism for enzymatic supercoiling of DNA. *Science*, 206:1081–1083.
- Buchanan-Wollaston, V., Passiatore, J. E., and Cannon, F. (1987). The *mob* and *oriT* mobilization functions of a bacterial plasmid promote its transfer to plants. *Nature*, 328:172–175.
- Cabral, J. H. M., Jackson, A. P., Smith, C. V., Shikotra, N., Maxwell, A., and Liddington, R. C. (1997). Crystal structure of the breakage-reunion domain of DNA gyrase. *Nature*, 388:903–906.
- Carattoli, A., Bertini, A., Villa, L., Falbo, V., Hopkins, K. L., and Threlfall, E. J. (2005). Identification of plasmids by PCR-based replicon typing. *Journal of Microbiological Methods*, 63:219–228.

- Catlin, B. W. and Cunningham, L. S. (1961). Transforming activities and base contents of deoxyribo-nucleate preparations from various *Neisseriae*. *Microbiology*, 26:303–312.
- Cesareni, G., Cornelissen, M., Lacatena, R. M., and Castagnoli, L. (1984). Control of pMB1 replication: inhibition of primer formation by Rop requires RNAI. *The EMBO Journal*, 3:1365–1369.
- Champoux, J. J. (2001). DNA topoisomerases: structure, function, and mechanism. *Annual Review of Biochemistry*, 70:369–413.
- Chant, E. L. and Summers, D. K. (2007). Indole signalling contributes to the stable maintenance of *Escherichia coli* multicopy plasmids. *Molecular Microbiology*, 63:35–43.
- Charlier, D., Hassanzadeh, G., Kholti, A., Gigot, D., Piérard, A., and Glansdorff, N. (1995). *carP*, involved in pyrimidine regulation of the *Escherichia coli* carbamoylphosphate synthetase operon encodes a sequence-specific DNA-binding protein identical to XerB and PepA, also required for resolution of ColE1 multimers. *Journal of Molecular Biology*, 250:392–406.
- Charlier, D., Roovers, M., Van Vliet, F., Boyen, A., Cunin, R., Nakamura, Y., Glansdorff, N., and Pierard, A. (1992). Arginine regulon of *Escherichia coli* K-12: a study of repressor-operator interactions and of *in vitro* binding affinities versus *in vivo* repression. *Journal of Molecular Biology*, 226:367–386.
- Chatwin, H. M. and Summers, D. K. (2001). Monomer-dimer control of the ColE1 P_{cer} promoter. *Microbiology*, 147:3071–3081.
- Cheah, U. E., Weigand, W. A., and Stark, B. C. (1987). Effects of recombinant plasmid size on cellular processes in *Escherichia coli*. *Plasmid*, 18:127–134.
- Chen, I. and Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nature Reviews: Microbiology*, 2:241–249.
- Chiang, C. S. and Bremer, H. (1988). Stability of pBR322-derived plasmids. *Plasmid*, 220:207–220.
- Cho, B.-K., Knight, E. M., Barrett, C. L., and Palsson, B. O. (2008). Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Research*, 18:900–910.
- Churchward, G., Bremer, H., and Young, R. (1982). Macromolecular composition of bacteria. *Journal of Theoretical Biology*, 94:651–670.
- Clewell, D. B. (1972). Nature of ColE1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. *Journal of Bacteriology*, 110:667–676.
- Clewell, D. B. and Helinski, D. R. (1969). Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and induced conversion to an open circular DNA form. *Proceedings of the National Academy of Sciences*, 62:1159–1166.

- Colloms, S. D., Alén, C., and Sherratt, D. J. (1998). The ArcA/ArcB two-component regulatory system of *Escherichia coli* is essential for Xer site-specific recombination at *psi*. *Molecular Microbiology*, 28:521–530.
- Colloms, S. D., Bath, J., and Sherratt, D. J. (1997). Topological selectivity in Xer site-specific recombination. *Cell*, 88:855–864.
- Colloms, S. D., McCulloch, R., Grant, K., Neilson, L., and Sherratt, D. J. (1996). Xer-mediated site-specific recombination *in vitro*. *The EMBO Journal*, 15:1172–1181.
- Colloms, S. D., Sykora, P., Szatmari, G., and Sherratt, D. J. (1990). Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. *Journal of Bacteriology*, 172:6973–6980.
- Cooper, S. and Helmstetter, C. E. (1968). Chromosome replication and the division cycle of *Escherichia coli* B/r. *Journal of Molecular Biology*, 31:519–540.
- Corbett, K. D., Schoeffler, A. J., Thomsen, N. D., and Berger, J. M. (2005). The structural basis for substrate specificity in DNA topoisomerase IV. *Journal of Molecular Biology*, 351:545–561.
- Cornet, F., Mortier, I., Patte, J., and Louarn, J.-M. (1994). Plasmid pSC101 harbors a recombination site, *psi*, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site *dif*. *Journal of Bacteriology*, 176:3188–3195.
- Costenaro, L., Grossmann, J. G., Ebel, C., and Maxwell, A. (2005). Small-angle X-ray scattering reveals the solution structure of the full-length DNA gyrase a subunit. *Structure*, 13:287–296.
- Costenaro, L., Grossmann, J. G., Ebel, C., and Maxwell, A. (2007). Modular structure of the full-length DNA gyrase B subunit revealed by small-angle X-ray scattering. *Structure*, 15:329–339.
- Couturier, M., Bex, F., Bergquist, P. L., and Maas, W. K. (1988). Identification and classification of bacterial plasmids. *Microbiology and Molecular Biology Reviews*, 52:375–395.
- Cullum, J. and Broda, P. (1979). Rate of segregation due to plasmid incompatibility. *Genetics Research*, 33:61–79.
- Das, N., Valjavec-Gratian, M., Basuray, A. N., Fekete, R. A., Papp, P. P., Paulsson, J., and Chattoraj, D. K. (2005). Multiple homeostatic mechanisms in the control of P1 plasmid replication. *Proceedings of the National Academy of Sciences*, 102:2856–2861.
- Davis, M. A. and Austin, S. J. (1988). Recognition of the P1 plasmid centromere analog involves binding of the ParB protein and is modified by a specific host factor. *The EMBO Journal*, 7:1881–1888.

- Di Martino, P., Fursy, R., Bret, L., Sundararaju, B., and Phillips, R. S. (2003). Indole can act as an extracellular signal to regulate biofilm formation of *Escherichia coli* and other indole-producing bacteria. *Canadian Journal of Microbiology*, 49:443–449.
- Di Martino, P., Merieau, A., Phillips, R., Orange, N., and Hulen, C. (2002). Isolation of an *Escherichia coli* strain mutant unable to form biofilm on polystyrene and to adhere to human pneumocyte cells: involvement of tryptophanase. *Canadian Journal of Microbiology*, 48:132–137.
- DiGate, R. J. and Marians, K. J. (1988). Identification of a potent decatenating enzyme from *Escherichia coli*. *The Journal of Biological Chemistry*, 263:13366–13373.
- Donoghue, D. J. and Sharp, P. A. (1978). Replication of Colicin E1 plasmid DNA *in vivo* requires no plasmid-encoded proteins. *Journal of Bacteriology*, 133:1287–1294.
- Drlica, K. and Snyder, M. (1978). Superhelical *Escherichia coli* DNA: relaxation by coumermycin. *Journal of Molecular Biology*, 146:145–154.
- Drlica, K. and Zhao, X. (1997). DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiology and Molecular Biology Reviews*, 61:377–392.
- Dubnau, D. (1991). Genetic competence in *Bacillus subtilis*. *Microbiology and Molecular Biology Reviews*, 55:395–424.
- Eberhard, W. G. (1990). Evolution in bacterial plasmids and levels of selection. *Quarterly Review of Biology*, 65:3–22.
- Ebersbach, G. and Gerdes, K. (2004). Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Molecular Microbiology*, 52:385–398.
- Ebersbach, G. and Gerdes, K. (2005). Plasmid segregation mechanisms. *Annual Review of Genetics*, 39:453–479.
- Edgar, R., Chatteraj, D. K., and Yarmolinsky, M. (2001). Pairing of P1 plasmid partition sites by ParB. *Molecular Microbiology*, 42:1363–1370.
- Eguchi, Y. and Tomizawa, J.-I. (1990). Complex formed by complementary RNA stem-loops and its stabilization by a protein: function of ColE1 Rom protein. *Cell*, 60:199–209.
- Ehrenberg, M. n. (1996). Hypothesis: hypersensitive plasmid copy number control for ColE1. *Biophysical Journal*, 70:135–145.
- Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, 297:1183–1186.
- Engberg, B. and Nordström, K. (1975). Replication of R-factor R1 in *Escherichia coli* K-12 at different growth rates. *Journal of Bacteriology*, 123:179–186.
- Espeli, O. and Marians, K. J. (2004). Untangling intracellular DNA topology. *Molecular Microbiology*, 52:925–931.

- Foster, T. J. (1983). Plasmid-determined resistance to antimicrobial drugs and toxic metal ions in bacteria. *Microbiological Reviews*, 47:361–409.
- Frenkel, L. and Bremer, H. (1986). Increased amplification of plasmids pBR322 and pBR327 by low concentrations of chloramphenicol. *DNA*, 5:539–544.
- Friedman, S. A. and Austin, S. J. (1988). The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon. *Plasmid*, 12:103–112.
- Frolova, L. Y., Meldrays, Y. A., Kochkina, L. L., Giller, S. A., Eremeyev, A. V., Grayevskaya, N. A., and Kisselev, L. L. (1977). DNA-polymerase inhibitors: rifamycin derivatives. *Nucleic Acids Research*, 4:523–538.
- Galibert, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Gurjal, M., Hernandez-lucas, I., Kahn, D., Kahn, M. L., Pohl, T. M., and Portetelle, D. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, 293:668–672.
- Garbe, T. R., Kobayashi, M., and Yukawa, H. (2000). Indole-inducible proteins in bacteria suggest membrane and oxidant toxicity. *Archives of Microbiology*, 173:78–82.
- Gellert, M., Mizuuchi, K., O’Dea, M. H., Itoh, T., and Tomizawa, J.-I. (1977). Nalidixic acid resistance: a second genetic character involved in DNA gyrase activity. *Proceedings of the National Academy of Sciences*, 74:4772–4776.
- Gellert, M., Mizuuchi, K., O’Dea, M. H., and Nash, H. A. (1976). DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proceedings of the National Academy of Sciences*, 73:3872–3876.
- Gerdes, K., Helin, K., Christensen, O. W., and Løbner Olesen, A. (1988). Translational control and differential RNA decay are key elements regulating postsegregational expression of the killer protein encoded by the *parB* locus of plasmid R1. *Journal of Molecular Biology*, 203:119–129.
- Gerdes, K., Howard, M., and Szardenings, F. (2010). Pushing and pulling in prokaryotic DNA segregation. *Cell*, 141:927–942.
- Gerdes, K. and Molin, S. (1986). Partitioning of plasmid R1: structural and functional analysis of the *parA* locus. *Journal of Molecular Biology*, 190:269–279.
- Gerdes, K., Møller-Jensen, J., and Jensen, R. B. (2000). Plasmid and chromosome partitioning: surprises from phylogeny. *Molecular Microbiology*, 37:455–466.
- Gillen, J. R., Willis, D. K., and Clark, A. J. (1981). Genetic analysis of the RecE pathway of genetic recombination in *Escherichia coli* K-12. *Journal of Bacteriology*, 145:521–532.
- Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81:2340–2361.

- Goodman, S. D. and Scocca, J. J. (1988). Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *Proceedings of the National Academy of Sciences*, 85:6982–6986.
- Goss, P. J. E. and Peccoud, J. (1998). Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets. *Proceedings of the National Academy of Sciences*, 95:6750–6755.
- Gould, J. M. and Cramer, W. A. (1977). Studies on the depolarization of the *Escherichia coli* cell membrane by colicin E1. *Journal of Biological Chemistry*, 252:5491–5497.
- Grohmann, E., Muth, G., and Espinosa, M. (2003). Conjugative plasmid transfer in gram-positive bacteria. *Microbiology and Molecular Biology Reviews*, 67:277–301.
- Guhathakurta, A., Viney, I., and Summers, D. K. (1996). Accessory proteins impose site selectivity during ColE1 dimer resolution. *Molecular Microbiology*, 20:613–620.
- Guo, X., Flores, M., Mavingui, P., Fuentes, S. I., Da, G., and Palacios, R. (2003). Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Research*, 13:1810–1817.
- Harrison, P. W., Lower, R. P. J., Kim, N. K. D., and Young, J. P. W. (2010). Introducing the bacterial ‘chromid’: not a chromosome, not a plasmid. *Trends in Microbiology*, 18:141–148.
- Hayes, F. (2003). Toxins-antitoxins: plasmid maintenance, programmed cell death and cell cycle arrest. *Science*, 301:1496–1499.
- Hayes, W. (1953). Observations on a transmissible agent determining differentiation in *Bacterium coli*. *Journal of General Microbiology*, 8:72–88.
- Heinemann, J. A. and Sprague Jr, G. F. (1989). Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature*, 340:205–209.
- Highlander, S. K. and Novick, R. P. (1987). Plasmid repopulation kinetics in *Staphylococcus aureus*. *Plasmid*, 17:210–221.
- Hinnebusch, J. and Barbour, A. G. (1991). Linear plasmids of *Borrelia burgdorferi* have a telomeric structure and sequence similar to those of a eukaryotic virus. *Journal of Bacteriology*, 173:7233–7239.
- Hinnebusch, J. and Tilly, K. (1993). Linear plasmids and chromosomes in bacteria. *Molecular Microbiology*, 10:917–922.
- Hippel, P. H. V., Bear, D. G., McSwiggen, J. A., and Morgan, W. D. (1984). Protein-nucleic acid interactions in transcription: a molecular analysis. *Annual Reviews of Biochemistry*, 53:389–446.

- Hirakawa, H., Inazumi, Y., Masaki, T., Hirata, T., and Yamaguchi, A. (2005). Indole induces the expression of multidrug exporter genes in *Escherichia coli*. *Molecular Microbiology*, 55:1113–1126.
- Hirano, M., Mori, H., Onogi, T., Yamazoe, M., Niki, H., Ogura, T., and Hiraga, S. (1998). Autoregulation of the partition genes of the mini-F plasmid and the intracellular localization of their products in *Escherichia coli*. *Molecular & General Genetics*, 257:392–403.
- Hirochika, H., Nakamura, K., and Sakaguchi, K. (1984). A linear DNA plasmid from *Streptomyces rochei* with an inverted terminal repetition of 614 base pairs. *The EMBO Journal*, 3:761–766.
- Hodgman, T. C., Griffiths, H., and Summers, D. K. (1998). Nucleoprotein architecture and ColE1 dimer resolution: a hypothesis. *Molecular Microbiology*, 29:545–558.
- Holmes, D. S. and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry*, 114:193–197.
- Hooper, D. C., Wolfson, J. S., McHugh, G. L., Swartz, M. D., Tung, C., and Swartz, M. N. (1984). Elimination of plasmid pMG110 from *Escherichia coli* by novobiocin and other inhibitors of DNA gyrase. *Antimicrobial Agents*, 25:586–590.
- Horowitz, D. S. and Wang, J. C. (1987). Mapping the active site tyrosine of *Escherichia coli* DNA gyrase. *The Journal of Biological Chemistry*, 262:5339–5344.
- Hoshino, K., Kitamura, A., Morrissey, I., Sato, K., Kato, J.-I., and Ikeda, H. (1994). Comparison of inhibition of *Escherichia coli* topoisomerase IV by quinolones with DNA gyrase inhibition. *Antimicrobial Agents*, 38:2623–2627.
- Hubschwerlen, C., Pflieger, P., Specklin, J.-L., Gubernator, K., Gmunder, H., Angehrn, P., and Kompis, I. (1992). Pyrimido[1,6-a]benzimidazoles: a new class of DNA gyrase inhibitors. *Journal of Medicinal Chemistry*, 35:1385–1392.
- Hughes, V. M. and Datta, N. (1983). Conjugative plasmids in bacteria of the ‘pre-antibiotic’ era. *Nature*, 302:725–726.
- Ikeda, H., Moriya, K., and Matsumoto, T. (1981). *in vitro* study of illegitimate recombination: involvement of DNA gyrase. *Cold Spring Harbor Symposia on Quantitative Biology*, 45:399.
- Itoh, T. and Tomizawa, J.-I. (1980). Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proceedings of the National Academy of Sciences*, 77:2450–2454.
- Jensen, L. B., Garcia-Migura, L., Valenzuela, A. J. S., Løhr, M., Hasman, H., and Aarestrup, F. M. (2010). A classification system for plasmids from enterococci and other Gram-positive bacteria. *Journal of Microbiological Methods*, 80:25–43.

- Jensen, R. B., Dam, M., and Gerdes, K. (1994). Partitioning of plasmid R1: the *parA* operon is autoregulated by *ParR* and its transcription is highly stimulated by a downstream activating element. *Journal of Molecular Biology*, 236:1299–1309.
- Jensen, R. B., Lurz, R., and Gerdes, K. (1998). Mechanism of DNA segregation in prokaryotes: replicon pairing by *parC* of plasmid R1. *Proceedings of the National Academy of Sciences*, 95:8550–8555.
- Johnsborg, O., Eldholm, V., and Håvarstein, L. S. (2007). Natural genetic transformation: prevalence, mechanisms and function. *Research in Microbiology*, 158:767–778.
- Johnson, R. C., Ball, C. A., Pfeffer, D., and Simon, M. I. (1988). Isolation of the gene encoding the *Hin* recombinational enhancer binding protein. *Proceedings of the National Academy of Sciences*, 85:3484–3488.
- Kaguni, J. M. and Kornberg, A. (1984). Replication initiated at the origin (*oriC*) of the *E. coli* chromosome reconstituted with purified enzymes. *Cell*, 38:183–190.
- Kantor, G. J. and Deering, R. A. (1968). Effect of nalidixic acid and hydroxyurea on division ability of *Escherichia coli* *fil*⁺ and *lon*[−] strains. *Journal of Bacteriology*, 95:520–530.
- Kato, J.-I., Nishimura, Y., Imamura, R., Niki, H., Hiraga, S., and Suzuki, H. (1990). New topoisomerase essential for chromosome segregation in *E. coli*. *Cell*, 63:393–404.
- Kato, J.-I., Suzuki, H., and Ikeda, H. (1992). Purification and characterization of DNA topoisomerase IV in *Escherichia coli*. *The Journal of Biological Chemistry*, 267:25676–25684.
- Keasling, J. D. and Palsson, B. O. (1989). ColE1 plasmid replication: A simple kinetic description from a structured model. *Journal of Theoretical Biology*, 141:447–461.
- Khodursky, A. B., Zechiedrich, E. L., and Cozzarelli, N. R. (1995). Topoisomerase IV is a target of quinolones in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 92:11801–11805.
- Konisky, J. (1982). Colicins and other bacteriocins with established modes of action. *Annual Review of Microbiology*, 36:125–144.
- Koppes, L. J. H., Meyer, M., Oonk, H. B., de Jong, M. A., and Nanninga, N. (1980). Correlation between size and age at different events in the cell division cycle of *Escherichia coli*. *Journal of Bacteriology*, 143:1241–1252.
- Kreuzer, K. N. and Cozzarelli, N. R. (1979). *Escherichia coli* mutants thermosensitive for deoxyribonucleic acid gyrase subunit A: effects on deoxyribonucleic acid replication, transcription, and bacteriophage growth. *Journal of Bacteriology*, 140:424–435.
- Kubitschek, H. E. (1990). Cell volume increase in *Escherichia coli* after shifts to richer media. *Journal of Bacteriology*, 172(1):94–101.

- Kuempel, P. L., Henson, J. M., Dircks, L., Tecklenburg, M., and Lim, D. F. (1991). *dif*, a *recA*-independent recombination site in the terminus region of the chromosome of *Escherichia coli*. *The New Biologist*, 3:799–811.
- Lanka, E. and Wilkins, B. M. (1995). DNA processing reactions in bacterial conjugation. *Annual Review of Biochemistry*, 64:141–169.
- Lawrence, J. G. and Ochman, H. (1997). Amelioration of bacterial genomes: rates of change and exchange. *Journal of Molecular Evolution*, 44:383–397.
- Lederberg, J. (1952). Cell genetics and hereditary symbiosis. *Physiological Reviews*, 32:403–430.
- Lederberg, J., Cavalli, L. L., and Lederberg, E. M. (1952). Sex compatibility in *Escherichia coli*. *Genetics*, 37:720–730.
- Lederberg, J. and Tatum, E. L. (1946). Gene recombination in *Escherichia coli*. *Nature*, 158:558.
- Lee, H. H., Molla, M. N., Cantor, C. R., and Collins, J. J. (2010). Bacterial charity work leads to population-wide resistance. *Nature*, 467:82–85.
- Lehninger, A. L. (1970). *Biochemistry: The Molecular Basis of Cell Structure and Function*. Worth, New York.
- Lewis, R. J., Singh, O. M., Smith, C. V., Skarzynski, T., Maxwell, A., Wonacott, A. J., and Wigley, D. B. (1996). The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *The EMBO Journal*, 15:1412–1420.
- Lin-Chao, S. and Bremer, H. (1986). Effect of the bacterial growth rate on replication control of plasmid pBR322 in *Escherichia coli*. *Molecular & General Genetics*, 203:143–149.
- Lin-Chao, S. and Bremer, H. (1987). Activities of the RNAI and RNAII promoters of plasmid pBR322. *Journal of Bacteriology*, 169:1217–1222.
- Lin-Chao, S., Wen-Tsuan, C., and Ten-Tsao, W. (1992). High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. *Molecular Microbiology*, 6:3385–3393.
- Livermore, D. M. and Brown, D. F. J. (2001). Detection of beta-lactamase-mediated resistance. *The Journal of Antimicrobial Chemotherapy*, 48:59–64.
- Makarova, K. S., Mironov, A. A., and Gelfand, M. S. (2001). Conservation of the binding site for the arginine repressor in all bacterial lineages. *Genome Biology*, 2:research0013.1–0013.8.
- Marmur, J., Rownd, R., Falkow, S., Baron, L. S., Schildkraut, C., and Doty, P. (1961). The nature of intergeneric episomal infection. *Proceedings of the National Academy of Sciences*, 47:972–979.

- Martínez-Robles, M. L., Witz, G., Hernández, P., Schwartzman, J. B., Stasiak, A., and Krimer, D. B. (2009). Interplay of DNA supercoiling and catenation during the segregation of sister duplexes. *Nucleic Acids Research*, 37:5126–5137.
- Masukata, H. and Tomizawa, J.-I. (1986). Control of primer formation for ColE1 plasmid replication: conformational change of the primer transcript. *Cell*, 44:125–136.
- Maxwell, A. (1997). DNA gyrase as a drug target. *Trends in Microbiology*, 5:102–109.
- McCulloch, R., Coggins, L. W., Colloms, S. D., and Sherratt, D. J. (1994). Xer-mediated site-specific recombination at *cer* generates Holliday junctions *in vivo*. *The EMBO Journal*, 13:1844–1855.
- McDowell, D. G. and Mann, N. H. (1991). Characterization and sequence analysis of a small plasmid from *Bacillus thuringiensis* var. *kurstaki* strain HD1-DIPEL. *Plasmid*, 120:113–120.
- Médigue, C., Rouxel, T., Vigier, P., Hénaut, A., and Danchin, A. (1991). Evidence for horizontal gene transfer in *Escherichia coli* speciation. *Journal of Molecular Biology*, 222:851–856.
- Meinhardt, F., Schaffrath, R., and Larsen, M. (1997). Microbial linear plasmids. *Applied Microbiology and Biotechnology*, 47:329–336.
- Metropolis, N. (1987). The beginning of the Monte Carlo method. *Los Alamos Science*, Special Issue:125–130.
- Meynell, E., Meynell, G. G., and Datta, N. (1968). Phylogenetic relationships of drug-resistance factors and other transmissible bacterial plasmids. *Molecular and Microbiology Reviews*, 32:55–83.
- Minh, P. N. L., Devroede, N., Massant, J., Maes, D., and Charlier, D. (2009). Insights into the architecture and stoichiometry of *Escherichia coli* PepA-DNA complexes involved in transcriptional control and site-specific DNA recombination by atomic force microscopy. *Nucleic Acids Research*, 37:1463–1476.
- Mizuuchi, K., O’Dea, M. H., and Gellert, M. (1978). DNA gyrase: subunit structure and ATPase activity of the purified enzyme. *Proceedings of the National Academy of Sciences*, 75:5960–5963.
- Møller Jensen, J., Jensen, R. B., Löwe, J., and Gerdes, K. (2002). Prokaryotic DNA segregation by an actin-like filament. *The EMBO Journal*, 21:3119–3127.
- Mori, H., Mori, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A., and Hiraga, S. (1989). Purification and characterization of SopA and SopB proteins essential for F plasmid partitioning. *The Journal of Biological Chemistry*, 264:15535–15541.
- Nöllmann, M., Crisona, N. J., and Arimondo, P. B. (2007). Thirty years of *Escherichia coli* DNA gyrase: from *in vivo* function to single-molecule mechanism. *Biochimie*, 89:490–499.

- Nordström, K. (1985). Control of plasmid replication: theoretical considerations and practical solutions. *Basic Life Sciences*, 30:189–214.
- Nordström, K. and Gerdes, K. (2003). Clustering versus random segregation of plasmids lacking a partitioning function: a plasmid paradox? *Plasmid*, 50:95–101.
- Nordström, K., Molin, S., and Aagaard-Hansen, H. (1980). Partitioning of plasmid R1 in *Escherichia coli* II: Incompatibility properties of the partitioning system. *Plasmid*, 4:332–349.
- Nordström, K., Molin, S., and Light, J. (1984). Control of replication of bacterial plasmids: genetics, molecular biology, and physiology of the plasmid R1 system. *Plasmid*, 12:71–90.
- Nordström, K. and Wagner, E. G. H. (1994). Kinetic aspects of control of plasmid replication by antisense RNA. *Trends in Biochemical Sciences*, 19:294–300.
- Norrande, J., Kempe, T., and Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene*, 26:101–106.
- Oblak, M., Grdadolnik, S. G., Kotnik, M., Jerala, R., Filipic, M., and Solmajer, T. (2005). In silico fragment-based discovery of indolin-2-one analogues as potent DNA gyrase inhibitors. *Bioorganic & Medicinal Chemistry Letters*, 15:5207–5210.
- Oblak, M., Grdadolnik, S. G., Kotnik, M., Poterszman, A., Atkinson, R. A., Nierengarten, H., Desplancq, D., Moras, D., and Solmajer, T. (2006). Biophysical characterization of an indolinone inhibitor in the ATP-binding site of DNA gyrase. *Biochemical and Biophysical Research Communications*, 349:1206–1213.
- Oblak, M., Kotnik, M., and Solmajer, T. (2007). Discovery and development of ATPase inhibitors of DNA gyrase as antibacterial agents. *Current Medicinal Chemistry*, 14:2033–2047.
- O’Brien, R. L., Olenick, J. G., and Hahn, F. E. (1966). Reactions of quinine, chloroquine, and quinacrine with DNA and their effects on the DNA and RNA polymerase reactions. *Proceedings of the National Academy of Sciences*, 55:1511–1517.
- O’Gorman, R. B., Dunaway, M., and Matthews, K. S. (1980). DNA binding characteristics of lactose repressor and the trypsin-resistant core repressor. *The Journal of Biological Chemistry*, 255:10100–10106.
- Ogura, T. and Hiraga, S. (1983). Partition mechanism of F plasmid: two plasmid gene-encoded products and a cis-acting region are involved in partition. *Cell*, 32:351–360.
- Oram, M., Marko, J. F., and Halford, S. E. (1997). Communications between distant sites on supercoiled DNA from non-exponential kinetics for DNA synapsis by resolvase. *Journal of Molecular Biology*, 270:396–412.

- Orr, E. and Staudenbauer, W. L. (1981). An *Escherichia coli* mutant thermosensitive in the B subunit of DNA gyrase: effect on the structure and replication of the colicin E1 plasmid *in vitro*. *Molecular & General Genetics*, 181:52–56.
- Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nature Genetics*, 31:69–73.
- Patient, M. E. and Summers, D. K. (1993). ColE1 multimer formation triggers inhibition of *Escherichia coli* cell division. *Molecular Microbiology*, 9:1089–1095.
- Paulsson, J. and Ehrenberg, M. n. (1998). Trade-off between segregational stability and metabolic burden: a mathematical model of plasmid ColE1 replication control. *Journal of Molecular Biology*, 279:73–88.
- Paulsson, J. and Ehrenberg, M. n. (2001). Noise in a minimal regulatory network: plasmid copy number control. *Quarterly Reviews of Biophysics*, 34:1–59.
- Paulsson, J., Nordström, K., and Ehrenberg, M. n. (1998). Requirements for rapid plasmid ColE1 copy number adjustments: a mathematical model of inhibition modes and RNA turnover rates. *Plasmid*, 39:215–234.
- Peng, H. and Marians, K. J. (1993). Decatenation activity of topoisomerase IV during *oriC* and pBR322 DNA replication *in vitro*. *Proceedings of the National Academy of Sciences*, 90:8571–8575.
- Peng, H. and Marians, K. J. (1995). The Interaction of *Escherichia coli* Topoisomerase IV with DNA. *Journal of Biological Chemistry*, 270:25286–25290.
- Pogliano, J., Ho, T. Q., Zhong, Z., and Helinski, D. R. (2001). Multicopy plasmids are clustered and localized in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 98:4486–4491.
- Pritchard, R. H., Barth, P. T., and Collins, J. (1969). Control of DNA synthesis in bacteria. *Symposium of the Society for General Microbiology*, 19:263–297.
- Reanney, D. (1976). Extrachromosomal elements as possible agents of adaptation and development. *Microbiology and Molecular Biology Reviews*, 40:552–590.
- Reece, R. J. and Maxwell, A. (1991). The C-terminal domain of the *Escherichia coli* DNA gyrase A subunit is a DNA-binding protein. *Nucleic Acids Research*, 19:1399–1405.
- Reijns, M., Lu, Y., Leach, S., and Colloms, S. D. (2005). Mutagenesis of PepA suggests a new model for the Xer/*cer* synaptic complex. *Molecular Microbiology*, 57:927–941.
- Rowe, D. C. D. and Summers, D. K. (1999). The quiescent-cell expression system for protein synthesis in *Escherichia coli*. *Applied and Environmental Microbiology*, 65:2710–2715.
- Roychowdhury, P. and Basak, B. S. (1975). The crystal structure of indole. *Acta Crystallographica*, 31:1559–1563.

- Sabik, J. F., Suit, J. L., and Luria, S. E. (1983). *cea-kil* operon of the ColE1 plasmid. *Journal of Bacteriology*, 153:1479–1485.
- Salje, J. (2010). Plasmid segregation: how to survive as an extra piece of DNA. *Critical Reviews in Biochemistry and Molecular Biology*, 45:296–317.
- Sambrook, J. and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbour Laboratory Press, 3rd edition.
- Saye, D. J., Ogunseitan, O., Sayler, G. S., and Millerl, R. V. (1987). Potential for transduction of plasmids in a natural freshwater environment: effect of plasmid donor concentration and a natural microbial community on transduction in *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*, 53:987–995.
- Schaaper, R. M., Danforth, B. N., and Glickman, B. W. (1986). Mechanisms of spontaneous mutagenesis: an analysis of the spectrum of spontaneous mutation in the *Escherichia coli lacI* gene. *Journal of Molecular Biology*, 189:273–284.
- Schaechter, M., Williamson, J. P., Hood, J. R., and Koch, A. L. (1962). Growth, cell and nuclear divisions in some bacteria. *Journal of General Microbiology*, 29:421–434.
- Schneider, R., Travers, A., Kutateladze, T., and Muskhelishvili, G. (1999). A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Molecular Microbiology*, 34:953–964.
- Schneider, R., Travers, A., and Muskhelishvili, G. (2000). The expression of the *Escherichia coli* *fis* gene is strongly dependent on the superhelical density of DNA. *Molecular Microbiology*, 38:167–175.
- Seo, J. H. and Bailey, J. E. (1985). Effects of recombinant plasmid content on growth properties and cloned gene product formation in *Escherichia coli*. *Biotechnology and Bioengineering*, 27:1668–1674.
- Sharpe, M. E., Chatwin, H. M., Macpherson, C., Withers, H. L., and Summers, D. K. (1999). Analysis of the ColE1 stability determinant *Rcd*. *Microbiology*, 145:2135–2144.
- Shehata, T. E. and Marr, A. G. (1970). Synchronous growth of enteric bacteria. *Journal of Bacteriology*, 103:789–792.
- Sherratt, D. J., Blakely, G., Burke, M., Colloms, S. D., Leslie, N., McCulloch, R., May, G., and Roberts, J. (1993). *Site-Specific Recombination and the Partition of Bacterial Chromosomes*, pages 25–41.
- Smith, G. (1988). Homologous recombination in procaryotes. *Microbiology and Molecular Biology Reviews*, 34:2258–2268.
- Snyder, M. and Drlica, K. (1979). DNA gyrase on the bacterial chromosome: DNA cleavage induced by oxolinic acid. *Journal of Molecular Biology*, 131:287–302.

- Solar, G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M., and Díaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews*, 62:434–464.
- Som, T. and Tomizawa, J.-I. (1983). Regulatory regions of ColE1 that are involved in determination of plasmid copy number. *Proceedings of the National Academy of Sciences*, 80:3232–3236.
- Sompayrac, L. and Maaloe, O. (1973). Autorepressor model for control of DNA replication. *Nature: New Biology*, 241:133–135.
- Sonnenwirth, A. C. (1980). *The Enteric Bacteria and Bacterioides*, pages 645–672. Harper & Row Publishers Inc., 3rd edition.
- Stanisich, V. A. (1988). Identification and analysis of plasmids at the genetic level. In Grinsted, J. and Bennett, P. M., editors, *Methods in Microbiology*, volume 21 of *Methods in Microbiology*, pages 11–47. Academic Press.
- Stark, W. M., Boocock, M. R., and Sherratt, D. J. (1992). Catalysis by site-specific recombinases. *Trends in Genetics*, 8:432–439.
- Staudenbauer, W. L. (1976). Replication of small plasmids in extracts of *Escherichia coli*: requirement for both DNA polymerases I and III. *Molecular & General Genetics*, 149:151–158.
- Steck, T. R. and Drlica, K. (1984). Bacterial chromosome segregation: evidence for DNA gyrase involvement in decatenation. *Cell*, 36:1081–1088.
- Steiner, W., Liu, G., Donachie, W. D., and Kuempel, P. (1999). The cytoplasmic domain of FtsK protein is required for resolution of chromosome dimers. *Molecular Microbiology*, 31:579–583.
- Stirling, C. J., Colloms, S. D., Collins, J. F., Szatmari, G., and Sherratt, D. J. (1989). *xerB*, an *Escherichia coli* gene required for plasmid ColE1 site-specific recombination, is identical to *pepA*, encoding aminopeptidase A, a protein with substantial similarity to bovine lens leucine aminopeptidase. *The EMBO Journal*, 8:1623–1627.
- Stirling, C. J., Szatmari, G., Stewart, G., Smith, M. C. M., and Sherratt, D. J. (1988). The arginine repressor is essential for plasmid-stabilizing site-specific recombination at the ColE1 *cer* locus. *The EMBO Journal*, 7:4389–4395.
- Sträter, N., Sherratt, D. J., and Colloms, S. D. (1999). X-ray structure of aminopeptidase A from *Escherichia coli* and a model for the nucleoprotein complex in Xer site-specific recombination. *The EMBO Journal*, 18:4513–4522.
- Stueber, D. and Bujard, H. (1982). Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes. *The EMBO Journal*, 1:1399–1404.

- Subramanya, H. S., Arciszewska, L. K., Baker, R. A., Bird, L. E., Sherratt, D. J., and Wigley, D. B. (1997). Crystal structure of the site-specific recombinase, Xer. *The EMBO Journal*, 16:5178–5187.
- Sugino, A., Higgins, N. P., Brown, P. O., Peebles, C. L., and Cozzarelli, N. R. (1978). Energy coupling in DNA gyrase and the mechanism of action of novobiocin. *Proceedings of the National Academy of Sciences*, 75:4838–4842.
- Sugino, A., Peebles, C. L., Kreuzer, K. N., and Cozzarelli, N. R. (1977). Mechanism of action of nalidixic acid: purification of *Escherichia coli* *nalA* gene product and its relationship to DNA gyrase and a novel nicking-closing enzyme. *Proceedings of the National Academy of Sciences*, 74:4767–4771.
- Summers, D. K. (1996). *The Biology of Plasmids*.
- Summers, D. K., Beton, C. W. H., and Withers, H. L. (1993). Multicopy plasmid instability: the dimer catastrophe hypothesis. *Molecular Microbiology*, 8:1031–1038.
- Summers, D. K. and Sherratt, D. J. (1984). Multimerization of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. *Cell*, 36:1097–1103.
- Summers, D. K. and Sherratt, D. J. (1988). Resolution of ColE1 dimers requires a DNA sequence implicated in the three-dimensional organization of the *cer* site. *The EMBO Journal*, 7:851–858.
- Tam, J. E. and Kline, B. C. (1989). The F plasmid *ccd* autorepressor is a complex of CcdA and CcdB proteins. *Molecular & General Genetics*, 219:26–32.
- Tomizawa, J.-I. (1984). Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript. *Cell*, 38:861–870.
- Tomizawa, J.-I. (1986). Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. *Cell*, 47:89–97.
- Tomizawa, J.-I. (1990). Control of ColE1 plasmid replication Intermediates in the binding of RNAI and RNAII. *Journal of Molecular Biology*, 212:683–694.
- Tomizawa, J.-I., Itoh, T., Selzer, G., and Som, T. (1981). Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proceedings of the National Academy of Sciences*, 78:1421–1425.
- Tomizawa, J.-I. and Som, T. (1984). Control of ColE1 plasmid replication: enhancement of binding of RNA I to the primer transcript by the Rom protein. *Cell*, 38:871–878.
- Toussaint, A. and Merlin, C. (2002). Mobile elements as a combination of functional modules. *Plasmid*, 47:26–35.
- Twigg, A. J. and Sherratt, D. J. (1980). Trans-complementable copy-number mutants of plasmid ColE1. *Nature*, 283:216–218.

- Uhlin, B. E. and Nordström, K. (1985). Preferential inhibition of plasmid replication *in vivo* by altered DNA gyrase activity in *Escherichia coli*. *Journal of Bacteriology*, 162:855–857.
- Van Duyne, G. D. (2001). A structural view of cre-*loxP* site-specific recombination. *Annual Review of Biophysics and Biomolecular Structure*, 30:87–104.
- Van Duyne, G. D., Ghosh, G., Maas, W. K., and Sigler, P. B. (1996). Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *Journal of Molecular Biology*, 256:377–391.
- Van Melder, L. and Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genetics*, 5:e1000437.
- Varsaki, A., Lucas, M., Afendra, A. S., Drainas, C., and de La Cruz, F. (2003). Genetic and biochemical characterization of MbeA, the relaxase involved in plasmid ColE1 conjugative mobilization. *Molecular Microbiology*, 48:481–493.
- Varsaki, A., Moncalián, G., Garcillán-Barcia, M. D. P., Drainas, C., and de La Cruz, F. (2009). Analysis of ColE1 MbeC unveils an extended ribbon-helix-helix family of nicking accessory protein. *Journal of Bacteriology*, 191:1446–1455.
- Vieira, J. and Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*, 19:259–268.
- Vogelstein, B. and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy of Sciences*, 76:615–619.
- Walker, K. A., Mallik, P., Pratt, T. S., and Osuna, R. (2004). The *Escherichia coli* Fis promoter is regulated by changes in the levels of its transcription initiation nucleotide CTP. *The Journal of Biological Chemistry*, 279:50818–50828.
- Wang, D., Ding, X., and Rather, P. N. (2001). Indole can act as an extracellular signal in *Escherichia coli*. *Journal of Bacteriology*, 183:4210–4216.
- Wang, J. C. (2002). Cellular roles of DNA topoisomerases: a molecular perspective. *Nature Reviews: Molecular Cell Biology*, 3:430–440.
- Wang, X., Reyes-lamothe, R., and Sherratt, D. J. (2008). Modulation of *Escherichia coli* sister chromosome cohesion by topoisomerase IV. *Genes & Development*, 22:2426–2433.
- Wang, Z., Jin, L., Yuan, Z., Wegrzyn, G., and Wegrzyn, A. (2009). Classification of plasmid vectors using replication origin, selection marker and promoter as criteria. *Plasmid*, 61:47–51.
- Watanabe, T. (1963). Infective heredity of multiple drug resistance in bacteria. *Bacteriological Reviews*, 27:87–115.

- Weinstein-Fischer, D. and Altuvia, S. (2007). Differential regulation of *Escherichia coli* topoisomerase I by Fis. *Molecular Microbiology*, 63:1131–1144.
- Willetts, N. and Wilkins, B. (1984). Processing of plasmid DNA during bacterial conjugation. *Microbiological Reviews*, 48:24–41.
- Winkler, W. C. and Breaker, R. R. (2005). Regulation of bacterial gene expression by riboswitches. *Annual Review of Microbiology*, 59:487–517.
- Wolfson, J. S., Hooper, D. C., Swartz, M. N., and McHugh, G. L. (1982). Antagonism of the B subunit of DNA gyrase eliminates plasmids pBR322 and pMG110 from *Escherichia coli*. *Journal of Bacteriology*, 152:338–344.
- Wood, W. A., Gunsalus, I. C., and Umbreit, W. W. (1947). Function of pyridoxal phosphate: resolution and purification of the tryptophanase enzyme of *Escherichia coli*. *Journal of Biological Chemistry*, 170:313–321.
- Yao, S., Helinski, D. R., and Toukdarian, A. (2007). Localization of the naturally occurring plasmid ColE1 at the cell pole. *Journal of Bacteriology*, 189:1946–1953.
- Zavitz, K. H. and Marians, K. J. (1991). Dissecting the functional role of PriA protein-catalysed primosome assembly in *Escherichia coli* DNA replication. *Molecular Microbiology*, 5:2869–2873.
- Zechiedrich, E. L. and Cozzarelli, N. R. (1995). Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in *Escherichia coli*. *Genes & Development*, 9:2859–2869.
- Zechiedrich, E. L., Khodursky, A. B., Bachellier, S., Schneider, R., Chen, D., Lilley, D. M. J., and Cozzarelli, N. R. (2000). Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli*. *Journal of Biological Chemistry*, 275:8103–8113.
- Zinder, N. D. and Lederberg, J. (1952). Genetic exchange in Salmonella. *Journal of Bacteriology*, 64:679–699.