DNA Charge Neutralisation by Linear Polymers I: Irreversible Binding

E Maltsev,* JAD Wattis,† and HM Byrne‡
Centre for Mathematical Medicine, Division of Applied Mathematics, School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK

We develop a deterministic mathematical model to describe the way in which polymers bind to DNA by considering the dynamics of the gap distribution that forms when polymers bind to a DNA plasmid. In so doing, we generalise existing theory to account for overlaps and binding cooperativity whereby the polymer binding rate depends on the size of the overlap The proposed mean-field models are then solved using a combination of numerical and asymptotic methods. We find that overlaps lead to higher coverage and hence higher charge neutralisations, results which are more in line with recent experimental observations. Our work has applications to gene therapy where polymers are used to neutralise the negative charges of the DNA phosphate backbone, allowing condensation prior to delivery into the nucleus of an abnormal cell.

PACS numbers: 82.20.-w, 82.30.-b,82.39.-k, 05.90.+m,89.75.Fb

I. INTRODUCTION

A. Importance of DNA charge neutralisation in gene therapy

One approach to gene therapy that is being used to treat a range of inherited and acquired diseases involves the introduction of DNA into the nucleus of abnormal cells to restore their function to normal [9, 11, 19, 23]. Delivery into such cells requires the DNA to be compacted either by polymers [5, 49], or transferred within another organism, such as a virus [13, 27]. Liposomes, cationic lipids [10, 39], or cationic polymers such as dendrimers [5, 24] or polyethylenimines [37], polyamidoamines [38, 42] are all examples of non-viral vectors.

Repulsive forces from negatively charged phosphate groups which prevent DNA from forming compact structures have to be neutralised by the vector to achieve condensation. The work presented in this paper focuses on modelling charge neutralisation of DNA in such non-viral gene delivery systems. It involves generalising and solving models of random sequential absorption [7, 44] using deterministic mean-field approaches, deriving expressions for the distribution of gap sizes and the overall charge neutralisation.

B. Experimental observations and the Counterion Condensation Theory

The main body of experimental work involves identifying polymers and ions that cause DNA condensation. While all studies agree that the charge of the polymers must exceed one in order for condensation to occur they are often focused on different types of polymers. Most studies consider salt solutions and polymers of relatively small valency but there is some experimental work involving polymers with much higher charges [5, 21, 48]. Experiments usually focus on the charge neutralisation required to condense the DNA and the morphology of the condensate.

Different polymers have been found to produce different shapes of condensate; there may be rods, toroids and spheres. The ability to control the shape, size and charge of the DNA-polymer complexes is important in gene therapy since these factors influence the complexes' suitability for transfection. Condensation studies performed by Roberts and coworkers [32, 41] using Atomic Force Microscopy (AFM) allow visualisation of DNA movement and conformational changes. Both toroidal and extended linear structures are observed when the DNA-polymer complex condenses. In [41], the authors conclude that rings are formed from the bending of the linear structures; that the two structures exist in a dynamic equilibrium, the balance of which can be influenced by the type of polymer used.

Linear polyamines, spermidine and spermine have been reported to employ a common molecular mechanism of DNA-binding [14]. Phosphates were found to be the primary binding sites of these polyamines. It was suggested that electrostatic shielding of the phosphates causes closer helix-helix surface contacts, facilitating the condensation of DNA. DNA may also be condensed by polyamidoamine dendrimers [5]. When present in solution in excess, the dendrimers produced complexes with sufficiently large net positive charge to ensure efficient cellular uptake. In [5] the authors conclude that the most transfection-efficient complexes are not those which are the most compact but those which have highest positive charge.

Overall, the experimental evidence suggests that DNA condenses when approximately 90% of its phosphate charges are neutralized [50].

*Electronic address: Eugene.Maltsev@maths.nottingham.ac.uk
†Corresponding author; Electronic address: Jonathan.Wattis@nottingham.ac.uk
‡Electronic address: Helen.Byrne@nottingham.ac.uk

Mathematical modelling of DNA condensation has resulted in two theories which admit a description of experimental results with varying levels of accuracy and parameter fitting. The term "counterion condensation" refers to mobile positive ions being attached to negative ions which are permanently fixed to DNA [31]. Solving the Poisson-Boltzmann equation confirms the results of counterion condensation theory (see [46] for example). The equations for binding isotherms derived from approximate analytical solution of Poisson-Boltzmann equation are given in [46]. Counterion condensation theory can be used to calculate the amount of negative DNA charge neutralised at equilibrium when non-specific binding of small polymers to DNA backbone occurs [50].

The second theory is the excluded-site binding model, which can be used to estimate the time-variation of the charge as well as the equilibrium state and is suitable for larger polymers [38, 48]. It is the second theory which is generalised in this paper.

C. Random sequential absorption and the excluded site binding model

The original model describing the binding of relatively long polymers to a number of contiguous sites of DNA was developed by McGhee & von Hippel [34]. Later, Epstein extended this work on the excluded-site binding model: in [15] an exact solution is derived for irreversible ligand binding to lattices of finite length and for studying the amount of ligand bound to a DNA molecule as a function of time; in [16] the results of Monte-Carlo simulations of reversible binding are reported; and in [17] a recurrence relation for the binding capacity is derived, however, this cannot be solved explicitly. More recently, Munro *et al.* [35] tabulated data on the capacities for various lengths of polymer and DNA in the case of irreversible binding.

Alongside this mathematical modelling of the biochemical process of polymer adsorption, a similar problem has been studied by sections of the theoretical physics community. The continuous formulation of the problem of *random sequential absorption* (RSA) is equivalent to polymers of unit length landing on a molecule of infinite length and is known as the *parking problem*; it is studied in [44]. Its discrete analogue is derived by Bonnier *et al.* [7], who calculate the proportion of the infinite linear lattice covered by the landing of polymers which cover a discrete number of sites. This quantity is equivalent in our example system, to the neutralisation of DNA charge, which we will denote by θ.

It is argued in [47] that the excluded-site binding model of [34] has a number of deficiencies. One is that the number of sites occupied by polymer and the binding rate do not vary with salt concentration. However, in principle such effects could be included by allowing the binding rate to depend on pH and the concentration of the ionic species. Bossmann & Schulman [8] discuss the difficulties of interpreting experimental data on the distribution of gap sizes arising in DNA adsorption studies, and point out the different effects which may dominate in studies on small vs long strands of DNA.

Extensions of standard RSA models include cooperative sequential adsorption (CSA) where polymers preferentially attach adjacent to existing adsorbed polymers, as described by Evans [18]. Whilst CSA leads to the higher charge neutralisations required for DNA condensation, it is hard to see the mechanism leading to cooperativity and associated preferential attachment. There are connections between models of island nucleation and growth and CSA, as noted by Evans [18] and Barma [2]. Barma [2] analyses the dynamics of deposition and evaporation in terms of spin-chain models; these effectively allow diffusion of adsorbed particles along the substrate. Bartelt & Privman [3] and Nielaba [36] consider multilayer adsorption initially of unit length polymers, later generalising further to consider longer polymers, mixtures, and surface diffusion following adsorption; connections with models of surface roughening and KPZ theory are noted. Such generalisations will be analysed using our methods in forthcoming papers[29, 30]. The topic of large-time asymptotics of RSA is studied by Ben-Naim & Krapivsky [4] who find exact solutions for some special cases. Krapivsky & Ben-Naim [25] consider reversible adsorption, a problem which we will analyse in a future work [29].

A variation of the model that approximates large ligands attached to polynucleotide as a one-dimensional fluid of hard rods [51] is used in [43]. However, this extension to counterion condensation theory still fails to accurately model the experimental data. It is suggested that the charge spacing decreases as more polymers are bound to polynucleotides. Only if binding rates different from the calculated ones are used, can the experimental data be made to fit the theoretical binding isotherms.

DNA condensation into multi-molecular rod- or toroid-like complexes is consistent with the idea that there are attractive forces between parts of a DNA molecule (or DNA molecules) which dominate when their charge is neutralised. In some systems, the fraction of neutralised charge on the DNA has exceeded one, the DNA having attracted so much positively charged polymer that the DNA-polymer complex acquires a net positive charge. This situation is known as charge inversion. The physics of charge inversion in chemical and biological systems is reviewed in [22].

The attractive forces between circular polyions of the same charge are studied in [26, 33]. Overcharging occurs because of the highly favourable gain in electrostatic free energy due to strong positional correlations between condensed counterions. The result is the appearance of purely electrostatic attraction between the like-charged macromolecules.

DNA condensation has been shown to occur when a large proportion of its negative charge is neutralised by positively charged polymers [38, 41, 49, 50]. Since many

sites become unavailable for binding, describing the interactions of a condensing polymer with the DNA from knowledge of the rate of binding of the polymer to unoccupied DNA is complicated. In particular, the overall rate of polymer binding to DNA is much larger at the start of the process when the DNA is empty, than at later stages, when the DNA is almost fully occupied. Our irreversible binding model, based on [12], describes these features. We are not claiming to describe fully all the processes occurring as polymers adsorb to DNA, there are many effects which we ignore: for example the curvature of the DNA due to its helical form, the relative spacing of charges on the polymer may not be the same as that on the DNA, that polymers may bind at each end but not in the middle. However, our model is an extension to existing RSA-based models in that we allow partial adherence of polymers, so that overlapped binding is permitted. One of the aims of this paper is to investigate if the effects of this extra model generality is sufficient to explain the higher charge neutralisations and charge inversion that is observed in experiments and not realisable in the standard (non-overlapped-binding) RSA model.

D. Gap distribution model

In this work, we model the DNA as a one-dimensional strand, with uniformly spaced binding sites. The model is used to analyse how the distribution of gap sizes evolves when polymers attach to the DNA. The distribution of gap sizes allows us to calculate the fraction of DNA sites that are occupied by charged polymers and the resulting charge neutralisation. We derive the governing equations by considering the general evolution of $N_p(t)$ the number of gaps of length p at time t. This may be stated as follows

$$\frac{dN_p}{dt} = -\begin{pmatrix}\text{rate at which gaps of}\\ \text{length } p \text{ disappear due}\\ \text{to polymer binding}\end{pmatrix} + \begin{pmatrix}\text{rate at which gaps of}\\ \text{length } p \text{ are created due}\\ \text{to polymer binding}\end{pmatrix}. \quad (1)$$

The initial irreversible binding model is derived and solved using numerical and asymptotic techniques in Section II. The model is then extended to allow for partially overlapped polymers (Section III), including the case when the binding rate depends on the gap-size (Section IV). Following [17] recursive relations are constructed for the steady state gap distributions that are realised in the large-time limit (Section II D). Using these the equilibrium charge neutralisation and final distribution of gap lengths can also be calculated. The paper ends in Section V, with a discussion of our results and possible directions for future research.

II. BINDING WITHOUT OVERLAPS

A. Kinetics

We denote by x the length of the polymer, and by p the length of the gap in which the incoming polymer will bind. Both x and p are positive integers (the unit being one base-pair). When such a polymer lands in such a gap on the DNA two smaller gaps are produced, one of length y the second of length $p-x-y$, where $0 \le y \le p-x$. Thus the possibilities are as follows

$$(p) \to (y) + (p-x-y), \qquad 0 \le y \le p-x, \qquad (2)$$

where the terms in brackets represent gaps of the corresponding size, as illustrated in Figure 1. There are $p-x+1$ ways in which a polymer of length x can position itself in a gap of length p, provided that $p \ge x$. We assume that each case occurs with equal probability.

We denote the total number of gaps between polymers on the DNA by M_0 so that

$$M_0(t) = \sum_{p=0}^{P_0} N_p(t). \qquad (3)$$

The number of polymers bound to the DNA is $M_0 - 1$, and the concentration of bound polymers is $B(t) = A_0(M_0(t)-1)$, where A_0 is the molar concentration of the DNA. Hence the molar concentration of free polymers $L(t)$ can be expressed as

$$L(t) = L_0 - B(t) = L_0 - A_0(M_0(t)-1), \qquad (4)$$

where $L_0 = L(t=0)$ is the molar concentration of polymers in the solution before any binding occurs.

Using (2) in (1) we have the differential equations

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = -K_f(p-x+1)N_p \quad (P_0-x+1 \le p \le P_0), \qquad (5a)$$

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = -K_f(p-x+1)N_p + \sum_{g=p+x}^{P_0} 2K_f N_g \quad (x \le p \le P_0-x), \qquad (5b)$$

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = \sum_{g=p+x}^{P_0} 2K_f N_g \quad (0 \le p \le x-1), \qquad (5c)$$

where P_0 is the length of the DNA plasmid, and $K_f = k_f L(t)$ denotes the rate at which the polymer lands, k_f being a rate constant.

The sink term on the RHS of equations (5a) and (5b) describes the loss of gaps of length p which are filled by incoming polymer. The source term that appears in equations (5b) and (5c) describes the formation of gaps

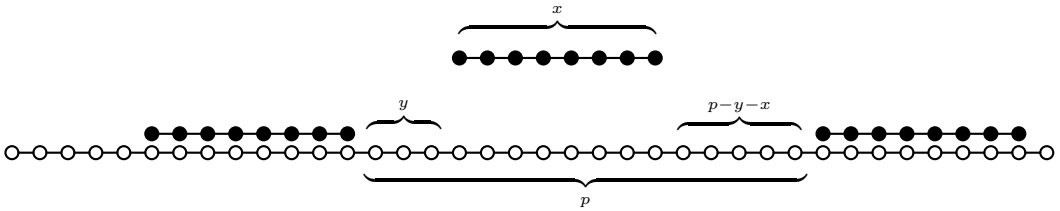


FIG. 1: Illustration of a polymer of length x attempting to land in a gap of length p on a DNA plasmid which has some sites already occupied.

of length p when polymer lands in a gap of length $g > p$. The factor of two is due to the fact that two gaps are formed when a landing event occurs. Since gaps of length p for which $P_0 - x + 1 \leq p \leq P_0$ can only be destroyed, equation (5a) contains just a sink term. By contrast, since gaps of length $p < x$ are too small to admit a polymer, no sink term is included in equation (5c).

At $t = 0$, we assume that no polymer has been bound and so simulations are performed with the initial condition $N_p(0) = \delta_{p,P_0}$ (using the Kronecker δ notation).

B. Charge neutralisation

As polymers adhere, they neutralise the negative charge of the DNA. Two physical quantities derived from the gap size distribution can be used to calculate the extent of charge neutralisation. They are the total number of gaps, M_0, as defined by (3) and the total length of gaps, M_1, defined by

$$M_1(t) = \sum_{p=1}^{P_0} pN_p(t). \tag{6}$$

The charge neutralisation θ is defined to be the proportion of charges on the DNA neutralised by the polymer. This can be calculated in two ways

$$\theta(t) = \frac{x\,(M_0(t) - 1)}{P_0} = \frac{P_0 - M_1(t)}{P_0}, \tag{7}$$

since $M_0 - 1$ is the number of polymer molecules attached to the DNA plasmid and $P_0 - M_1$ is the total number of sites occupied by the polymers. We thus have the identity

$$xM_0(t) + M_1(t) = P_0 + x, \tag{8}$$

which is valid for all t.

C. Numerical solution

Equations (5) were solved using a semi-implicit extrapolation method [40] with adaptive step-size control written and compiled using Fortran 90. The charge neutralisation θ was then calculated using equation (7).

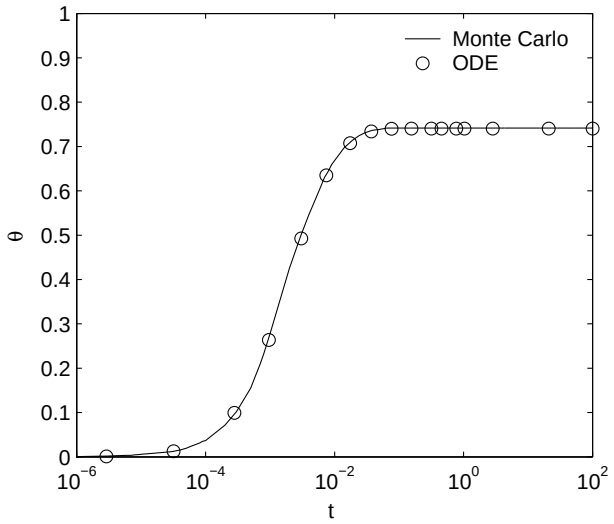


FIG. 2: Evolution of charge neutralisation θ calculated from equation (5) and compared to Monte Carlo simulation results. Parameter values: $L_0 = 8$, $A = 1$, $P_0 = 20$ sites, $x = 5$ sites, $k_f = 10s^{-1}$.

The evolution of the charge neutralisation for a typical simulation is plotted in Figure 2. The accuracy of the numerical solution is confirmed by comparing it to results from a Monte-Carlo simulation with the same parameters [28]. From Figure 2 we note in both cases a rapid approach to an equilibrium charge neutralisation.

D. Irreversible equilibrium asymptotics

A recursive relation, developed in [17], allows us to calculate the equilibrium binding capacity, $R(x, P_0)$. This is the average number of polymers of length x that irreversibly bind to DNA of length P_0. $R(x, P_0)$ corresponds to an equilibrium value of $M_0(t) - 1$ from Section II (equation (3)). Clearly if the DNA is shorter than the polymer ($P_0 < x$) then no polymers can land and so $R(x, P_0) = 0$.

If $x \leq P_0 < 2x$ then only one polymer can be accommodated so that $R(x, P_0) = 1$. For longer DNA chains ($P_0 \geq 2x$) the number of polymers landing will depend upon precisely where earlier polymers landed. When a polymer lands on a lattice of length P_0, two shorter, sublattices are generated. The total capacity of the long lattice is one more than the sum of the capacities of the two shorter lattices. By conditioning on the landing site, the recursive relation

$$R(x, P_0) = 1 + \frac{2\sum_{q=x}^{P_0-x} R(x, q)}{P_0 - x + 1} \qquad (2x \leq P_0), \quad (9a)$$

$$R(x, P_0) = 1 \qquad (x \leq P_0 < 2x), \quad (9b)$$

$$R(x, P_0) = 0 \qquad (P_0 < x), \quad (9c)$$

for the expected capacity can be derived. In equations (9) we view P_0 as the recurrence parameter. We now derive a recurrence relation for $R(x, P_0)$. From (9a) we have that

$$R(x, P_0 - 1) = 1 + 2\frac{-R(x, P_0 - x) + \sum_{p=x}^{P_0-x} R(x, p)}{P_0 - x}, \quad (10)$$

which, upon rearranging, yields

$$2\sum_{p=x}^{P_0-x} R(x, p) = (P_0 - x)\left(R(x, P_0 - 1) - 1\right) + 2R(x, P_0 - x). \quad (11)$$

By substituting from (11) into (9) and noting that $R(x, P_0)$ can be calculated explicitly for $P_0 \leq 3x - 1$, we deduce

$$R(x, P_0) = 1 + \frac{(P_0 - x)(R(x, P_0 - 1) - 1) + 2R(x, P_0 - x)}{P_0 - x + 1} \qquad (3x \leq P_0), \quad (12a)$$

$$R(x, P_0) = 1 + 2\left(\frac{P_0 - 2x + 1}{P_0 - x + 1}\right) \qquad (2x \leq P_0 < 3x), \quad (12b)$$

$$R(x, P_0) = 1 \qquad (x \leq P_0 < 2x), \quad (12c)$$

$$R(x, P_0) = 0 \qquad (P_0 < x). \quad (12d)$$

Equation (12a) is an xth order recurrence relation in P_0, with P_0-dependent coefficients, for which an explicit solution is not available.

We are interested in the behaviour of solutions when the polymers occupy many base pairs and the DNA is many times longer than the polymer, that is, $1 \ll x \ll P_0$. There are many such scalings, and so for the purposes of giving illustrative calculations we choose one particular limit, namely that of $P_0 = \mathcal{O}(x^2) \gg 1$. Accordingly we define a small parameter, ϵ, by $\epsilon = 1/x \ll 1$. We assume that $P_0 \sim O(\epsilon^{-2})$ and write $y = \epsilon^2 P_0$, where $y \sim O(1)$. This scaling is typical of those used in experiments on DNA-condensation by cationic polymers where the DNA plasmid is typically of length 10^4 base pairs and polymers are of length 10^2 units.

If $\theta \sim O(1)$ then, using equation (7), $R = P_0\theta/x \sim O(1/\epsilon)$, that is, a plasmid of length $O(1/\epsilon^2)$ can accept $O(1/\epsilon)$ polymers of length $O(1/\epsilon)$. We therefore rescale the irreversible binding capacity as $R(x, P_0) = r(y)/\epsilon$ with $r \sim O(1)$. We substitute x, P_0, and R in equation (12a), to yield

$$\left(\frac{y}{\epsilon^2} - \frac{1}{\epsilon} + 1\right) r(y) = \epsilon + \left(\frac{y}{\epsilon^2} - \frac{1}{\epsilon}\right) r(y - \epsilon^2) + 2r(y - \epsilon). \quad (13)$$

We construct an asymptotic expansion for $r(y)$ approximating $r(y-\epsilon)$ by $r(y) - \epsilon\frac{d}{dy}r(y)$ since $y \sim O(1)$. Making this approximation and retaining only terms of leading order and first order, yields the following ODE for $r(y)$

$$(y + \epsilon)\frac{dr(y)}{dy} = r(y) + \epsilon, \quad (14)$$

with solution

$$r(y) = c(y + \epsilon) - \epsilon, \quad (15)$$

for some constant c. Rewriting (15) in terms of the original variables yields

$$\theta(x, P_0) = \theta(x, \infty) - \frac{x}{P_0}\left(1 - \theta(x, \infty)\right), \quad (16)$$

where the constant $\theta(x, \infty)$ is determined by matching equation (16) to the largest solution of equation (12) that is easily obtained analytically.

From equation (12b), the largest value of P_0 for which we can determine $R(x, P_0)$ without using the recurrence relation occurs when $P_0 = 3x - 1$. Substituting $R(x, 3x - 1) = 2$ in equation (16) gives

$$\theta(x, 3x - 1) = \theta(x, \infty) - \frac{x}{3x - 1}\left(1 - \theta(x, \infty)\right) = \frac{2x}{3x - 1}, \quad (17)$$

which, on rearrangement, yields

$$\theta(x, \infty) = \frac{3x}{4x - 1}. \quad (18)$$

Substitution of $\theta(x, \infty)$ in equation (16) yields the following expression for the charge neutralisation of DNA molecules having at least $3x$ sites

$$\theta(x, P_0) = \frac{3x}{4x - 1}\left(1 - \frac{x - 1}{3P_0}\right). \quad (19)$$

In Figure 3 the asymptotic solution (19) is shown to be in very close agreement with the exact recurrence relations over a range of values for DNA length P_0.

Equation (19) implies that when polymers bind irreversibly without overlaps the highest charge neutralisation occurs when $x = 2$ and $\theta(2, \infty) = 0.86$. Such neutralisation is insufficient to condense DNA. Applying equation (19) to the continuous parking problem results in $\lim_{x\to\infty, P_0\to\infty} \theta(x, P_0) = 0.750$, which is close to the solution obtained in [44], where $\theta = 0.748$.

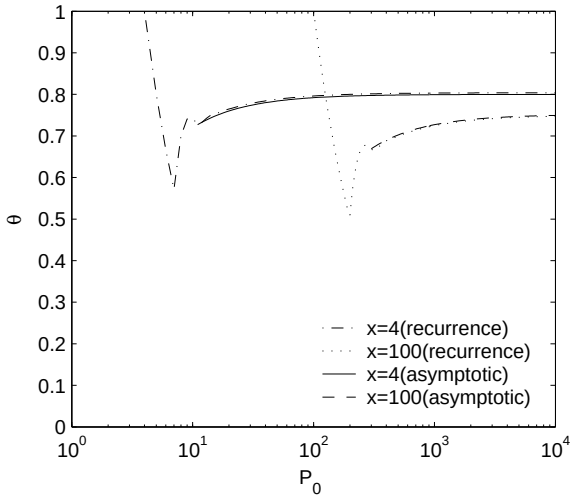


FIG. 3: Exact charge neutralisation θ when non-overlapped polymers that occupy 4 and 100 sites bind irreversibly to DNA of various lengths P_0. There is excellent agreement between the asymptotic results and those obtained from the recurrence relation.

From [20], the kinetics of the neutralisation of charge for a lattice of infinite length are

$$\theta(x,t) = \int_0^{x(1-\exp(-k_f t))} \exp\left(-2\int_0^u \frac{1-(1-v/x)^{x-1}}{v}\mathrm{d}v\right)\mathrm{d}u, \quad (20)$$

where k_f is the binding rate.

In Figure 4 the asymptotic solution (19) is plotted and shown to be in good agreement with exact solutions compiled in [20] of the equilibrium charge neutralisation that were obtained by evaluating equation (20) in the limit $t=\infty$.

E. Equilibrium gap length distribution

By generalising the argument presented in the previous section it is possible to derive recurrence relations for the equilibrium distribution of gaps, that is $\lim_{t\to\infty} N_p(t)$, which we shall now denote by $N_p^{\mathrm{eq}}(x,P_0)$, explicitly stating the dependence of N_p^{eq} on x and P_0. Irreversible binding without overlaps results in all gaps larger than $x-1$ being filled by polymers so that $N_p^{\mathrm{eq}}(x,P_0)=0$ for $p\geq x$. As stated in Section II D, an empty DNA molecule of length P_0 has P_0-x+1 landing positions for a polymer of length x. Each point of initial attachment results in two sub-lattices, a sub-lattice and a gap, or two gaps. Here, by 'gap' we mean a gap of size $p<x$ and by 'sub-lattice' we mean a gap of size $p\geq x$ which will be further subdivided by the binding of more polymers.

The total number of gaps is calculated recursively by considering the distribution of gap sizes on smaller sub-lattices.

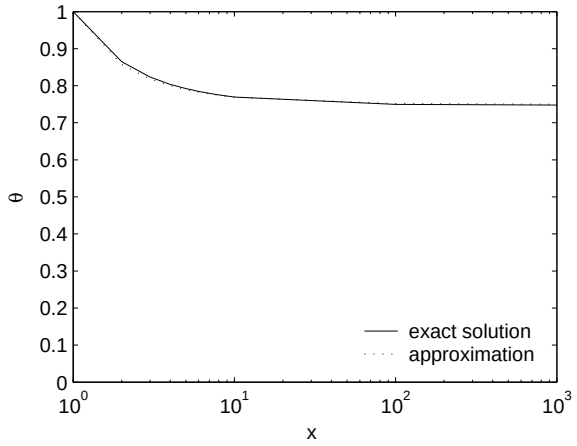


FIG. 4: Diagram showing how the charge neutralisation changes as the length of the polymer varies when the polymer binds irreversibly and with no overlaps to a strip of DNA of infinite length.

$$\begin{pmatrix}\text{Number of gaps}\\ \text{of length } p \text{ on}\\ \text{plasmid of length } P_0\end{pmatrix} = 2\frac{\begin{pmatrix}\text{Possibility of gap}\\ \text{generation on}\\ \text{the first landing}\end{pmatrix} + \begin{pmatrix}\text{Sum of all gaps of size } p\\ \text{from sub-lattices of DNA}\\ \text{formed by the landing}\end{pmatrix}}{\begin{pmatrix}\text{Total number of}\\ \text{landing positions}\end{pmatrix}}. \quad (21)$$

The resulting recursive relation for the distribution of gaps of length $(0\leq p\leq x-1)$ is

$$N_p^{\mathrm{eq}}(x,P_0) = \frac{2}{(P_0-x+1)}\left(1+\sum_{q=x+p}^{P_0-x} N_p^{\mathrm{eq}}(x,q)\right), \quad (p+2x\leq P_0), \quad (22a)$$

$$N_p^{\mathrm{eq}}(x,P_0) = \frac{2}{P_0-x+1}, \qquad (p+x\leq P_0<p+2x), \quad (22b)$$

$$N_p^{\mathrm{eq}}(x,P_0) = 0, \qquad (P_0<p+x). \quad (22c)$$

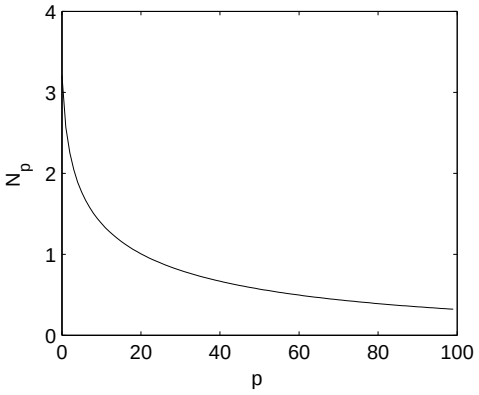
Using an argument similar to that used to derive equation (12), equation (22) can be rewritten as an xth order recurrence relation:

$$N_p^{\mathrm{eq}}(x,P_0) = \frac{(P_0-x)N_p^{\mathrm{eq}}(x,P_0-1)+2N_p^{\mathrm{eq}}(x,P_0-x)}{P_0-x+1}, \quad (p+x<P_0), \quad (23a)$$

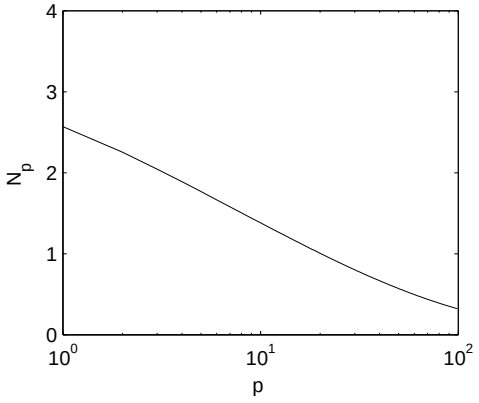
$$N_p^{\mathrm{eq}}(x,P_0) = \frac{2}{P_0-x+1}, \qquad (P_0=p+x), \quad (23b)$$

$$N_p^{\mathrm{eq}}(x,P_0) = 0, \qquad (P_0<p+x). \quad (23c)$$

A typical case is presented in Figure 5, where the equilibrium gap distribution $N_p^{\rm eq}(x, P_0)$ is plotted for fixed values of x and P_0. The log-scale plot in Figure 5(b) suggests that the distribution is approximately of the form $N_p = a - b \log p$ (where a and b are constants).



(a) linear scale



(b) log scale

FIG. 5: Series of plots showing the equilibrium gap distribution when binding is irreversible, with no overlaps. Parameter values: $P_0 = 10000$ and $x = 100$ ($\theta = 0.7473$).

III. OVERLAPPED BINDING

Experimental results suggest that at equilibrium a net resultant positive charge can arise when positively charged polymers land on negatively charged DNA [1]. One explanation of this phenomenon is that some polymers are only partially bound to the DNA surface. This could occur when polymers attach to gaps on the DNA which are shorter in length than the polymers (seeFigure 6).

In this section we consider polymer-binding with overlaps, assuming as before that the binding rate is independent of the length of the binding region.

A. Kinetics

As before, when a polymer of length x lands in a gap of width p, two smaller gaps are produced, one of length y and the other of length $p - x - y$. The possibilities are as follows

$$(p) \to (y) + (p - x - y) \qquad 1 - x \le y \le p - 1. \quad (24)$$

Negative values of y and $p - x - y$ correspond to regions where there is an overlap of length $-y$ or $-(p - x - y)$. Polymers cannot land in "negative" gaps: they are only formed when polymers land.

As illustrated in Figure 6, a polymer can bind in a gap of length $p > 0$ in $p + x - 1$ ways. If the polymers land on the DNA plasmid at the rate K_f then gaps of size p are removed at a rate $-K_f(p + x - 1)N_p$, which is proportional to the the number of possible landing positions and the number of such gaps, N_p. As in Section II, the assumption that the polymer lands on the DNA plasmid at the rate K_f means that the rate at which gaps of size p are created is given by $2K_f \sum_{g=p+1}^{P_0} N_g$. The resulting equations are

$$\frac{{\rm d}N_p}{{\rm d}t} = -K_f(p + x - 1)N_p \qquad (p = P_0), \quad (25a)$$

$$\frac{{\rm d}N_p}{{\rm d}t} = -K_f(p + x - 1)N_p + \sum_{g=p+1}^{P_0} 2K_f N_g \qquad (1 \le p \le P_0 - 1), \quad (25b)$$

$$\frac{{\rm d}N_p}{{\rm d}t} = \sum_{g=1}^{P_0} 2K_f N_g \qquad (1 - x \le p \le 0). \quad (25c)$$

Although equations (25) appear to be identical to (5), their domains of applicability in p-space differ, and the factors of $p - x + 1$ have been replaced by $p + x - 1$. In both cases gaps of length P_0 cannot be created as binding leads only to the formation of smaller gaps. Gaps of size $(1 \le p \le P_0 - 1)$ may be created and destroyed. Once a gap of zero, or negative size is created, no further binding event can remove it; therefore such gaps can only be created. Since we allow overlaps to occur, the lower limit of the sum appearing in equations (25) is $g = 1$ rather than $g = p + 1$. This is because the smallest gap into which a polymer may land has size 1.

As in the case of nonoverlapped binding (8), we have $M_1(t) + xM_0(t) = P_0 + x$, and

$$\theta = \frac{x(M_0 - 1)}{P_0} = 1 - \frac{M_1}{P_0}. \quad (26)$$

However, M_0 and M_1 are now defined slightly differently:

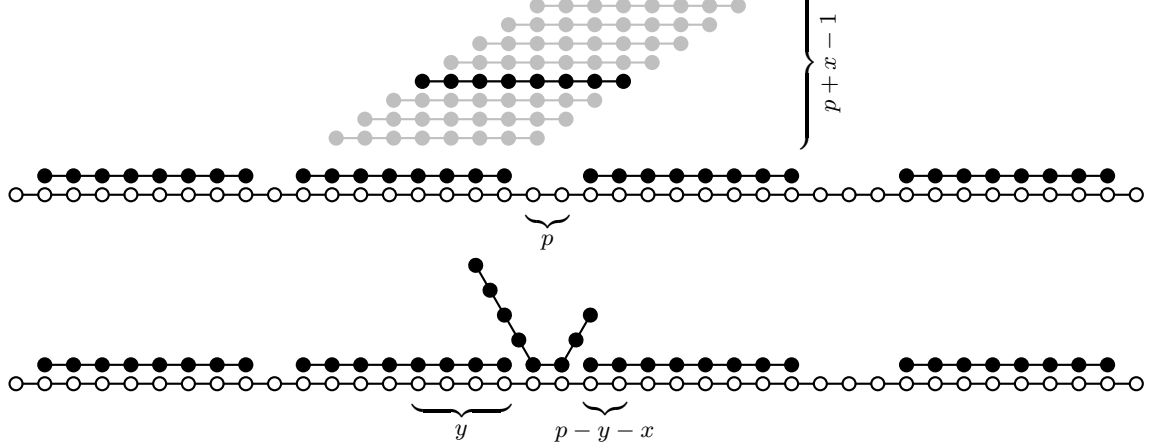


FIG. 6: There are $p + x - 1$ possible landing positions (shown in upper panel, in gray) for a polymer of length x binding to DNA in a gap of width $p < x$.

in place of (3) and (6) we have

$$M_0(t) = \sum_{p=1-x}^{P_0} N_p(t), \qquad M_1(t) = \sum_{p=1-x}^{P_0} pN_p(t), \quad (27)$$

thus $M_0(t) \geq 0$, but $M_1(t) < 0$ is allowable, leading to the possibility that $\theta > 1$ in (26).

B. Numerical solution

As in Section II C, equations (25) were solved using a semi-implicit extrapolation method with adaptive step-size control. The charge neutralisation corresponding to a numerical solution of equations (25) is plotted in Figure 7 for the binding of a 5-site polymer to a strip of DNA of length 200. We observe much greater charge neutralisation when overlapping occurs.

The equilibrium gap distributions corresponding to the simulations presented in Figure 7 are plotted in Figure 8 and show that overlapped irreversible binding results in a uniform distribution of overlaps. This surprising result is due to the unrealistic assumption that the adherence rate of polymers does not depend on the number of bonds that form (that is, the size of the gap in which the polymer lands). A more realistic scenario in which the rate of adherence depends on gap size is analysed in the next section.

C. Equilibrium plasmid capacity

It is possible to construct a recursive relation for the equilibrium binding capacity of a strip of DNA when polymers bind irreversibly with overlaps (i.e. the equilibrium solution of the model presented in Section III A). The derivation is similar to that presented in Section II D, except that binding can now occur in any gap on the

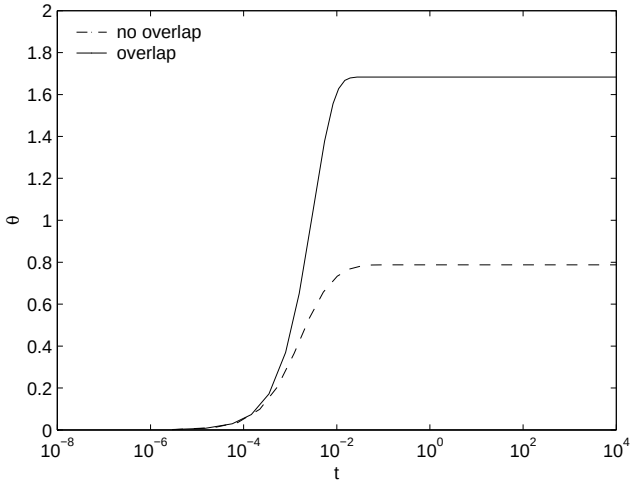


FIG. 7: Plot showing how for irreversible binding the charge neutralisation changes when binding with overlaps occurs. Parameter values: $L_0 = 10^{-6}M$, $A = 5 \times 10^{-9}M$, $P_0 = 200$ sites, $x = 5$ sites, $k_f = 10^8 M^{-1}s^{-1}$.

DNA of length greater than zero. It follows that there are $P_0 + x - 1$ landing positions for polymer of length x landing on a stretch of DNA of length P_0. Therefore the average number of polymers of length x that bind irreversibly (with overlaps) to a DNA of length P_0 is

$$R(x, P_0) = 1 + \frac{2\sum_{p=1}^{P_0-1} R(x,p)}{P_0 + x - 1}. \quad (28)$$

We remark that the minimum gap length that can accommodate another polymer has decreased from x in equation (9) (lower limit of the sum) to 1. The largest gap created by the landing polymers has also increased from $P_0 - x$ to $P_0 - 1$ (upper limit of the sum). Since DNA of length $P_0 = 1$ always accepts just one polymer, we

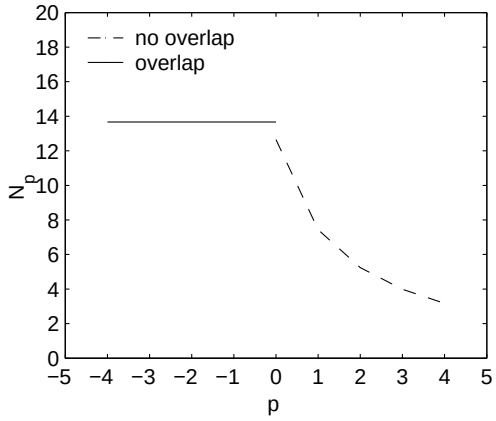


FIG. 8: Plot showing how for irreversible binding the equilibrium gap distribution changes when binding with overlaps occurs. Parameter values: as per Figure 7.

impose the boundary condition

$$R(x,1) = 1, \tag{29}$$

and we may use (29) to simplify equation (28) to obtain

$$R(x,P_0) = 1 + \frac{2(P_0-1)}{x+1}. \tag{30}$$

It follows from (30) that the fraction of neutralised charge is

$$\theta(x,P_0) = \frac{x}{P_0}\left(1 + \frac{2(P_0-1)}{x+1}\right) = 1 + \frac{(x-1)(P_0+x)}{P_0(x+1)}, \tag{31}$$

so that charge inversion occurs for all $x > 1$ and

$$\theta \sim 2 + \frac{x}{P_0}, \tag{32}$$

as $P_0 \to \infty$ and $x \to \infty$, with x/P_0 fixed. For $x = O(1)$ and $P_0 \to \infty$, we have

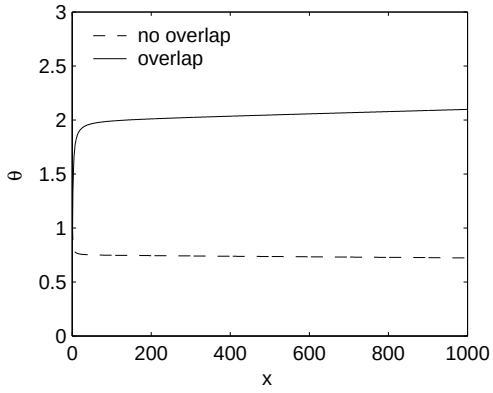
$$\theta \sim 2 - \frac{2}{x+1}. \tag{33}$$

It is possible to combine equations (32) and (33) by considering the asymptotic limit $P_0 \gg 1$ with $x \sim \sqrt{P_0}$, since then equation (31) yields

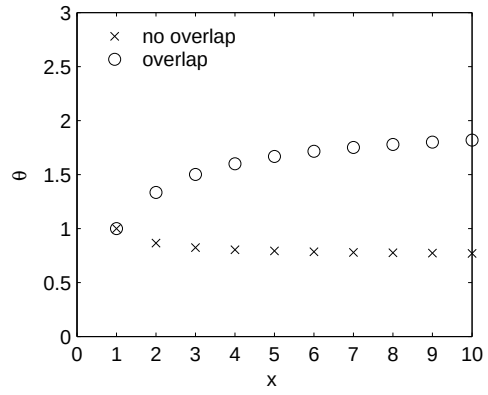
$$\theta \sim 2 - \frac{2}{x} + \frac{x}{P_0}. \tag{34}$$

The scaling $P_0 = \mathcal{O}(x^2) \gg 1$ was chosen in Section II D as *an* example of $P_0 \gg x \gg 1$ due to its applicability to polymer adsorption onto a DNA substrate. We see from (34) that the scaling $P_0 \sim x^2 \gg 1$ is a useful example in that it simultaneously illustrates both the effects (32) which occurs for $P_0 \sim x \gg 1$ and (33) which holds when $P_0 \gg x \sim 1$.

Figure 9 shows how θ varies with x for DNA of length 10^4 sites. As expected from (31), charge inversion is evident for all values of $x > 1$ and θ increases with x.



(a)



(b)

FIG. 9: Equilibrium charge neutralisation θ, as defined by equation (31), when polymers of length x bind irreversibly to DNA of length $P_0 = 10^4$. In (a) the charge neutralisation θ is plotted for the range of polymer lengths $1 \le x \le 1000$, showing a region of rapid variation for smaller x. In (b) we focus on the range $1 \le x \le 10$, showing that the two cases coincide at $x = 1$.

Figure 10 displays a series of curves showing how, for a given polymer of fixed length, the equilibrium charge neutralisation varies when the length of the DNA is altered. We note that $\theta \to 1 + (x-1)/(1+x)$ as $P_0 \to \infty$, with $x = O(1)$, as expected from equation (33).

D. Equilibrium gap length distribution

To find the equilibrium gap distribution when overlaps occur we use a similar technique to that described above for θ. A polymer of length x can land on an empty DNA plasmid of length P_0 in $P_0 - x + 1$ ways, as in the case of non-overlapped binding. In addition, there are $x - 1$ landing positions on each side of the DNA plasmid which

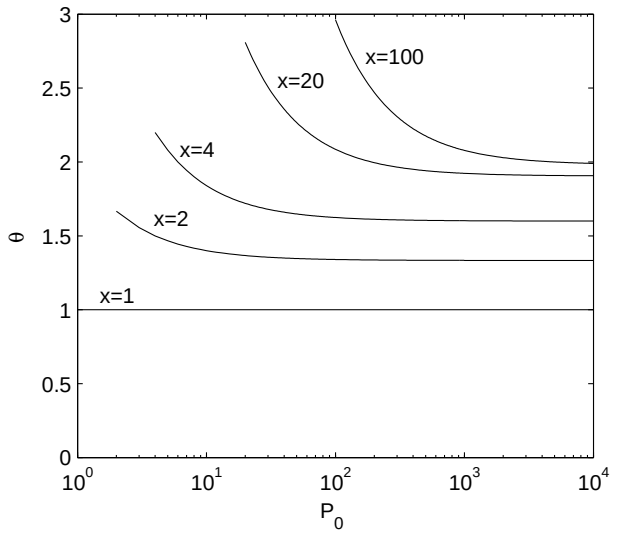


FIG. 10: Series of curves showing how, for polymers of fixed length x, the equilibrium charge neutralisation θ varies with P_0, the length of the DNA plasmid.

result in an overlapping. Thus, following arguments similar to those preceding equation (22), we obtain the recursive relation

$$N_p^{\rm eq}(x,P_0)=\frac{2}{(P_0+x-1)}\left(1+\sum_{q=1}^{P_0-1}N_p^{\rm eq}(x,q)\right)\quad(1<P_0), \tag{35}$$

for the gap distribution $(1-x\le p\le 0)$; with $N_p^{\rm eq}(x,1)=2/x$, since only one polymer binds to DNA of length $P_0=1$ and it can do so in x different ways, leading to gaps of size p and $1-p-x$ $(1-x\le p\le 0)$, each of which are equally likely. The recurrence relation and the initial condition are independent of p. Therefore when overlapped binding occurs at a constant rate, the equilibrium gap distribution is uniform.

Equation (35) can be simplified by noting that

$$N_p^{\rm eq}(x,P_0-1)=\frac{2}{(P_0+x-2)}\times\left(1+\sum_{q=1}^{P_0-1}N_p^{\rm eq}(x,q)-N_p^{\rm eq}(x,P_0-1)\right); \tag{36}$$

combining (35) and (36) it follows that

$$N_p^{\rm eq}(x,P_0)=\left(\frac{P_0+x}{P_0+x-1}\right)N_p^{\rm eq}(x,P_0-1)\quad(1<P_0), \tag{37a}$$

$$N_p^{\rm eq}(x,1)=\frac{2}{x}. \tag{37b}$$

Equation (37) can be solved to give

$$N_p^{\rm eq}(x,P_0)=\left(\frac{P_0+x}{1+x}\right)N_p^{\rm eq}(x,1)=\frac{2\,(P_0+x)}{x\,(x+1)}, \tag{38}$$

which we note is independent of gap size p. It is also possible to calculate the charge neutralisation. Suppose that the average gap length is L. Then $M_1=LM_0$ (note $L<0$, $M_0>0$, $M_1<0$). The fraction of neutralised charge is then

$$\theta=\frac{(P_0-L)}{P_0}\frac{x}{(x+L)}=\frac{1-L/P_0}{1+L/x}. \tag{39}$$

Now we recall that $xM_0+M_1=P_0+x$ (see equation (8)). Since there is an equal number of gaps of every length, the average gap length L satisfies

$$L=\frac{1}{x}\sum_{i=1-x}^{0}i=\frac{1-x}{2}, \tag{40}$$

as expected for a uniform distribution on $1-x\le p\le 0$. Substitution of L in equation (39) confirms equation (31).

IV. IRREVERSIBLE BINDING WITH OVERLAP-DEPENDENT RATES

A. Kinetics

It is reasonable to assume that a polymer is less likely to bind to the DNA if the location where it binds is partially occupied by other polymers. We account for this effect by allowing the binding rate K_f to vary with p, the number of bonds formed, in the following way

$$K_f^{(p)}=\frac{K_f}{\lambda^{x-p}}\qquad(p\le x). \tag{41}$$

In (41) λ is the binding cooperativity constant. If polymer is positioned so that all x bonds form then $K_f^{(x)}=K_f$. The binding rate is reduced if the gap between polymers already attached to the DNA is not large enough to accept the full length of the landing polymer: every time the gap size is reduced by one site, it is decreased by a factor of λ. Hence the lowest non-zero adhesion rate corresponds to a polymer landing and binding to the DNA in a gap of unit length, which occurs at a rate $K_f^{(x)}/\lambda^{x-1}$. Varying λ from unity to infinity interpolates between the previously described models of overlapped binding ($\lambda=1$) and binding without overlaps ($\lambda=\infty$). Typically we expect λ to be close to unity. This is because the probability of forming each individual bond is high, but sufficiently far away from unity that the probability of forming all x bonds is distinct from unity. Treating each bond-formation event as independent, which we acknowledge is an approximation, the probability of forming p bonds is λ^p. Other choices for $K_f^{(p)}$ could be made, for example $K_f^{(p)}=K_f(p/x)^\nu$ with $\nu>0$.

1. *Destruction of gaps by polymer-binding*

We now assume that the rate of binding depends on the number of sites at which the polymer binds. Figures

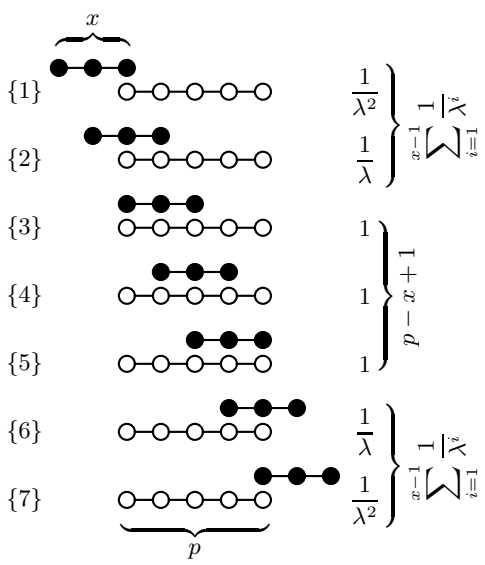


FIG. 11: Overlap-length cooperativity of overlapped landing ($x = 3$, $p = 5$).

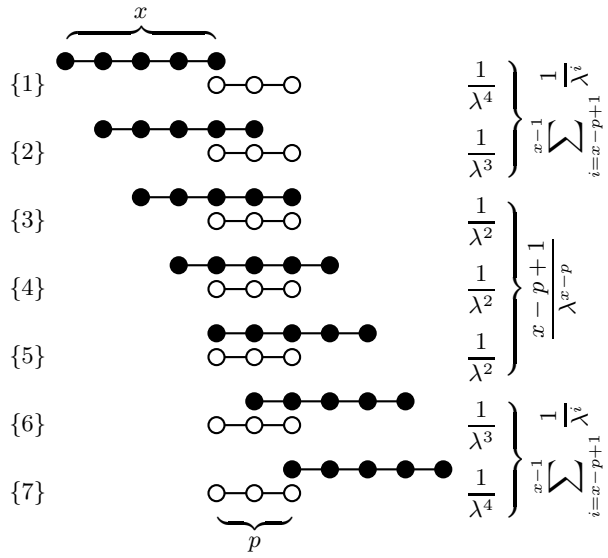


FIG. 12: Overlap-length cooperativity of overlapped landing ($x = 5$, $p = 3$).

11 and 12 illustrate all the contributions to the sink terms of the binding kinetics equations. When the gap is longer than the polymer (i.e. $p > x$) there are $p + x - 1$ sites at which a polymer can bind partially if we allow overlaps. This contrasts with $p - x + 1$ sites at which the polymer may bind when overlaps are not allowed. The binding rate depends on the number of sites to which the polymer attaches. Individual binding rates corresponding to each position and expressions combining them are displayed on the RHS of Figure 11. The rate at which gaps of length p are removed is obtained by summing all the possible binding rates and in this case it results in

$$\begin{aligned}\begin{pmatrix}\text{rate at which gaps of}\\ \text{length } p \text{ are removed}\end{pmatrix} &= K_f\left(p - x + 1 + 2\sum_{i=1}^{x-1}\frac{1}{\lambda^i}\right)N_p \\ &= K_f\left(p - x + 1 + \frac{2(1-\lambda^{1-x})}{\lambda - 1}\right)N_p. \quad (42)\end{aligned}$$

When the gap is smaller than the polymer $p < x$ (Figure 12) then any polymer that attaches will bind partially. As a result, the rate at which gaps of length $p < x$ form depends on the number of bound sites; there are still $p + x - 1$ binding configurations.

When the polymer fully covers the gap, the number of overlapped sites is $x - p$ and hence the binding rate is

$$K_f^{(p)} = \frac{K_f}{\lambda^{x-p}}.$$

Since there are $x - p + 1$ such positions, the contribution to the rate of destruction of gaps from polymers which fully cover the gap is:

$$K_f \lambda^{p-x}(x - p + 1)N_p. \quad (43)$$

Polymers which partially cover the gap contribute as follows to gap destruction

$$2K_f\left(\sum_{i=x-p+1}^{x-1}\frac{1}{\lambda^i}\right)N_p = K_f\left(\frac{2(\lambda^{p-1}-1)}{\lambda^{x-1}(\lambda-1)}\right)N_p. \quad (44)$$

Combining (43) and (44) we deduce that the rate at which gaps of length $1 \le p \le x$ are destroyed is given by

$$K_f\left(\frac{2\lambda^{1-x}(\lambda^{p-1}-1)}{\lambda-1} + \lambda^{p-x}(x-p+1)\right)N_p. \quad (45)$$

2. *Creation of gaps and overlaps by polymer-binding*

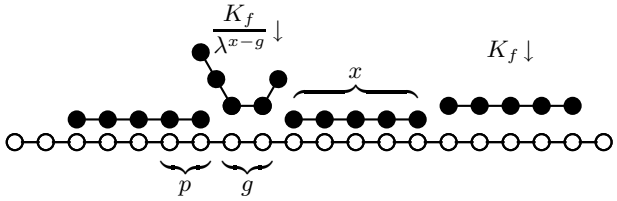


FIG. 13: An overlap of the length p created inside of the short gap of the length g.

A gap of length $1 - x \le p \le 0$ is created when a polymer binds partially to a gap of length $1 \le g \le p + x$ (illustrated in Figure 13) by g sites only. It follows from (41) that the binding rate is $K_f^{(x)}/\lambda^{x-g}$ and the contribution to the creation of gaps is

$$2K_f\sum_{g=1}^{p+x}\frac{N_g}{\lambda^{x-g}}. \quad (46)$$

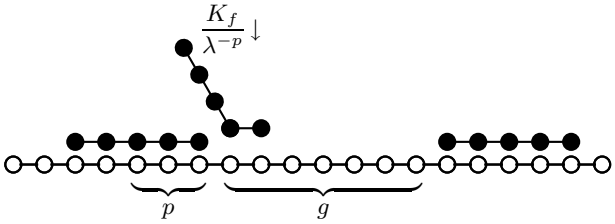


FIG. 14: An overlap of length p is created when a polymer partially binds inside a gap of length g.

When larger gaps of length $p+x+1 \leq g \leq P_0$ are destroyed, and the only sites where the polymer does not bind are due to the overlaps, then the binding rate is $K_f^{(x)}/\lambda^{-p}$ (see Figure 14) and the contribution to the creation of gaps is

$$2K_f \sum_{g=p+x+1}^{P_0} \frac{N_g}{\lambda^{-p}}. \tag{47}$$

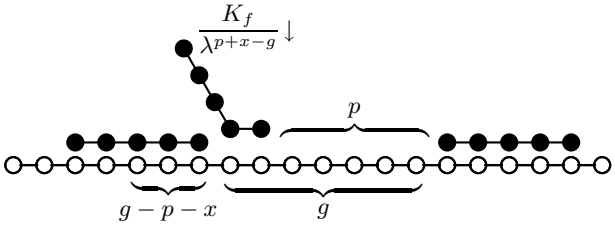


FIG. 15: A gap of length p is created when a polymer partially binds inside a gap of the length g.

If a gap of length $1 \leq p \leq P_0 - x - 1$ is created when a polymer binds to a gap of length $p+1 \leq g \leq p+x$, then the polymer is attached by $g-p$ sites only (see Figure 15). It follows from (41) that the binding rate is $K_f^{(x)}/\lambda^{p+x-g}$ and the contribution to the creation of gaps is

$$2K_f \sum_{g=p+1}^{p+x} \frac{N_g}{\lambda^{p+x-g}}. \tag{48}$$

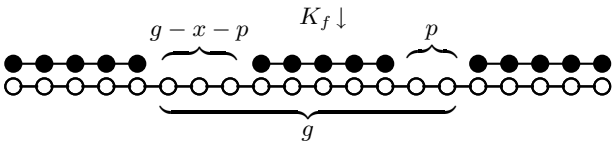


FIG. 16: A gap of length p is created when a polymer lands inside a gap of length g.

When larger gaps of length $p+x+1 \leq g \leq P_0$ are destroyed, leaving a gap of size p, with no overlaps being formed, the binding rate is $K_f^{(x)}$ (see Figure 16) and the contribution to the creation of gaps is

$$2K_f \sum_{g=p+x+1}^{P_0} N_g. \tag{49}$$

Larger gaps of length $P_0 - x \leq p \leq P_0 - 1$ can only be created by partially bound polymers which attach with $g-p$ sites. It follows from (41) that the binding rate is $K_f^{(x)}/\lambda^{p+x-g}$ and the contribution to the creation of gaps is the same as (48) with upper limit of the sum set to P_0 as the length of the DNA limits the size of the gap being destroyed

$$2K_f \sum_{g=p+1}^{P_0} \frac{N_g}{\lambda^{p+x-g}}. \tag{50}$$

Using the above results we deduce that when equations (5) or (25) are adjusted to allow for overlap-dependent binding rates, the following system of differential equations is obtained:

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = -K_f N_p \left(p - x + 1 + \frac{2(1-\lambda^{1-x})}{\lambda - 1}\right) \quad (p = P_0), \tag{51a}$$

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = -K_f N_p \left(p - x + 1 + \frac{2(1-\lambda^{1-x})}{\lambda - 1}\right) + 2K_f \sum_{g=p+1}^{P_0} \frac{N_g}{\lambda^{p+x-g}} \quad (P_0 - x \leq p \leq P_0 - 1), \tag{51b}$$

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = -K_f N_p \left(p - x + 1 + \frac{2(1-\lambda^{1-x})}{\lambda - 1}\right) + 2K_f \sum_{g=p+x+1}^{P_0} N_g + 2K_f \sum_{g=p+1}^{p+x} \frac{N_g}{\lambda^{p+x-g}} \quad (x+1 \leq p \leq P_0 - x - 1), \tag{51c}$$

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = -K_f \left(\frac{2\lambda^{1-x}(\lambda^{p-1}-1)}{\lambda - 1} + \lambda^{p-x}(x-p+1)\right) N_p + 2K_f \sum_{g=p+x+1}^{P_0} N_g + 2K_f \sum_{g=p+1}^{p+x} \frac{N_g}{\lambda^{p+x-g}} \quad (1 \leq p \leq x), \tag{51d}$$

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = 2K_f \sum_{g=p+x+1}^{P_0} \frac{N_g}{\lambda^{-p}} + 2K_f \sum_{g=1}^{p+x} \frac{N_g}{\lambda^{x-g}} \quad (1 - x \leq p \leq 0). \tag{51e}$$

A semi-implicit extrapolation method was used to solve the system of equations (51). Figure 17 shows how the charge neutralisation evolves for different values of the binding cooperativity constant, λ, when a polymer of length $x = 5$ binds to a DNA plasmid of length $P_0 = 200$. The uppermost curve in Figure 17 corresponds to overlapped binding at a rate which is independent of the size of the landing site ($\lambda = 1$) as in Section III, and has a shape similar to the plot of binding without overlaps ($\lambda \to \infty$). The main differences between the two

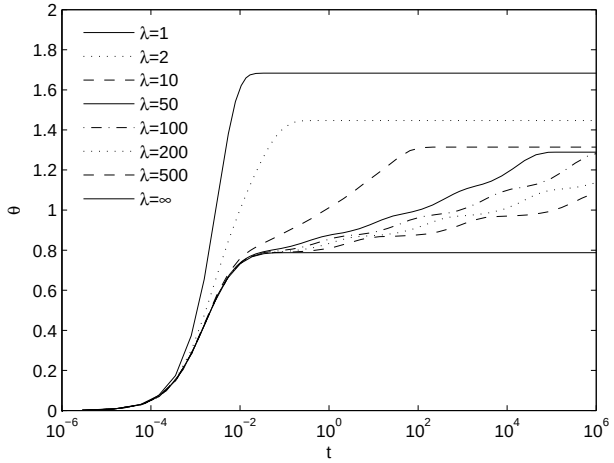


FIG. 17: Series of curves showing how the dynamics of the charge neutralisation change for irreversible binding with different binding cooperativity constants λ. Binding with no overlaps ($\lambda = \infty$) corresponds to the lowest curve and $\lambda = 1$ where overlapped binding is not penalised in any way corresponds to the top curve. Parameter values: $L_0 = 10^{-6}M$, $A = 5\times10^{-9}M$, $x = 5$ sites, $P_0 = 200$ sites, $k_f = 10^8 M^{-1}s^{-1}$.

curves are that when $\lambda \to \infty$ the equilibrium value of θ is smaller. Plots for intermediate values of λ appear to have steps. These may be explained by gradual filling of the smaller gaps by binding with increasingly large overlaps. The clearest example is for the case $\lambda = 100$ in Figure 17 where 5 steps are clearly visible. The first step corresponds to non-overlapped binding equilibrium (at $t = 0.1s$), the second when gaps of length $x - 1 = 4$ sites are filled. Since the binding rate decreases by a factor of λ as the gap size decreases by one site, the second plateau occurs when $t = 1s$, the third when $t = 100s$, and so on.

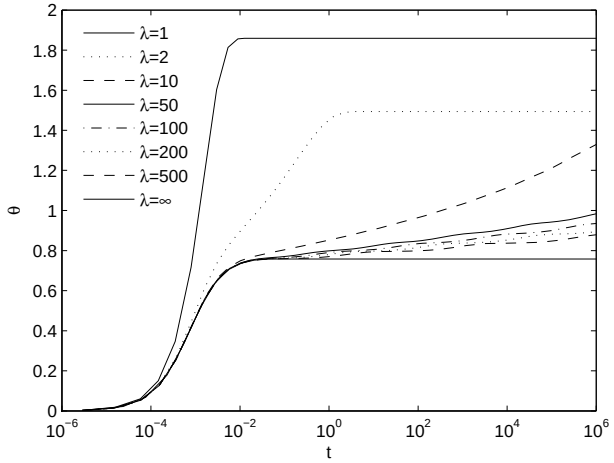


FIG. 18: Irreversible binding with different binding cooperativity constants. Parameter values: as per Figure 17, except $x = 10$ sites.

The only difference between the parameters used to construct Figures 17 and 18 is polymer length, ($x = 5$ in Figure 17 and $x = 10$ in Figure 18). We note that convergence to equilibrium is slower when x increases (this is particularly evident when $\lambda = 2$).

B. Irreversible binding capacity

When the binding rate varies with the gap size p according to (41), the binding capacity, that is the total number of polymers of length x which can be expected to adhere to a plasmid of length P_0, is

$$R(x,P_0) = 1 + \frac{2\sum_{p=1}^{P_0-x} R(x,p) + 2\sum_{p=P_0-x+1}^{P_0-1} \lambda^{P_0-x-p} R(x,p)}{P_0 - x + 1 + 2\sum_{i=1}^{x-1} \lambda^{-i}} \qquad (x \le P_0), \quad (52a)$$

$$R(x,P_0) = 1 + \frac{2\sum_{p=1}^{P_0-1} \lambda^{P_0-x-p} R(x,p)}{(x-P_0-1)\lambda^{P_0-x} + 2\sum_{i=x-P_0}^{x-1} \lambda^{-i}} \qquad (2 \le P_0 < x), \quad (52b)$$

$$R(x,1) = 1. \quad (52c)$$

We recall that non-overlapped binding is equivalent to size-dependent overlapped binding with $\lambda = \infty$ and note that equation (52a) reduces to equation (9a) in this case. Similarly equation (52a) reduces to equation (28) when $\lambda = 1$. Equations (52) may be derived by considering all possible landing configurations of the polymers on the DNA plasmid. Figure 19 illustrates what cases may arise when the polymer is shorter than the DNA. The factor by which $k_f^{(x)}$ should be multiplied to determine the effective binding rate is given on the RHS.

There are $P_0 + x - 1$ landing positions (as in the case of constant-rate binding with overlaps) but these configurations occur with different probabilities because polymers are less likely to bind when they cannot adhere with all of their charges.

Figure 20 shows how the charge neutralisation varies with λ when $P_0 = 10000$ and $x = 100$. In the limit $\lambda \to \infty$, θ hardly changes from its value of 1.46 (which it attains at $\lambda = 1.5$).

Note that there is no connection between the result derived in Section II D where $\theta < 1$ and $\lambda = \infty$ and the asymptotic limit $\lambda \to \infty$ which gives $\theta \approx 1.5$. Even for very large values of λ gaps of size $p < x$ are filled over very long timescales, the plasmid will eventually be completely covered in the limit $t \to \infty$, and have $\theta \ge 1$. However if $\lambda = \infty$, gaps of size $p < x$ are never filled and so $\theta \le 1 \ \ \forall \ \ t$.

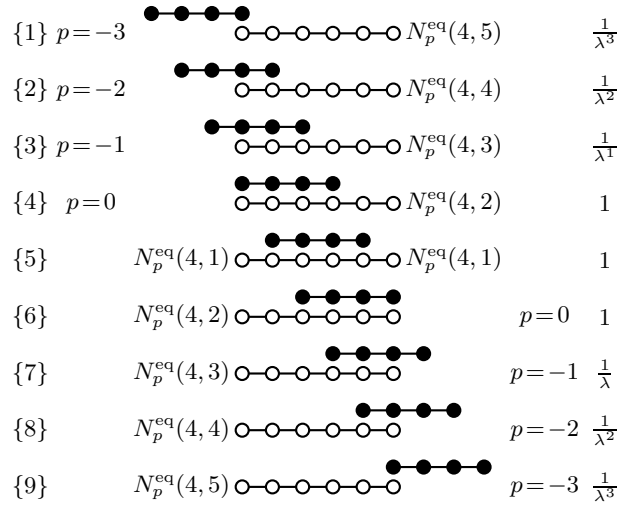


FIG. 19: Gap-length cooperativity of overlapped landing ($x = 4, P_0 = 6$).

Figure 21 shows the equilibrium charge neutralisation θ as a function of the polymer length x and cooperativity constant λ. As expected, increasing λ and/or x leads to an increase in θ.

C. Equilibrium gap distribution for variable-rate overlapped binding

Recursive relations for the equilibrium gap distribution ($N_p^{\rm eq}(x, P_0)$) when there is variable-rate binding with overlaps are derived in this section. Figure 19 illustrates the derivation of the steady-state gap distribution when polymers of length $x = 4$ land on a stretch of DNA of length $P_0 = 6$.

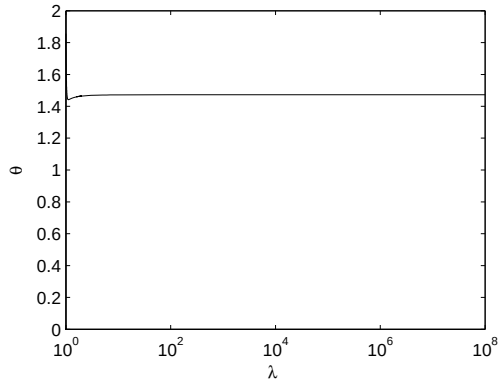
Let $N_{-3}^{\rm eq}(4, 6)$ be the number of gaps of length -3 when polymers of length 4 land on a stretch of DNA which is 6 sites long. The relative rate at which each possible landing event occurs is shown on the right hand side of Figure 19. In this case the sum of the weights of all outcomes is

$$\omega = (1/\lambda^3 + 1/\lambda^2 + 1/\lambda + 3 + 1/\lambda + 1/\lambda^2 + 1/\lambda^3),$$

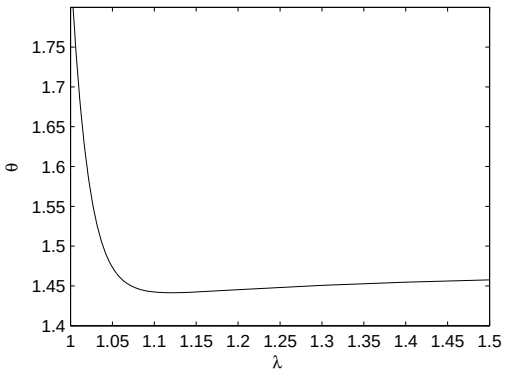
which generalises to $(P_0 - x + 1 + 2\sum_{i=1}^{x-1} \lambda^{-i})$ for polymers of length x landing on gaps of length P_0.

The probability of a polymer landing in configuration {1} (see Figure 19) is the rate at which this happens ($1/\lambda^3$) normalised by the sum of all possible weights (ω), giving

$$\begin{pmatrix}\text{probability of}\\ \text{polymer landing in}\\ \text{configuration } \{1\}\end{pmatrix} = \frac{1/\lambda^3}{\omega} = \frac{1/\lambda^3}{1/\lambda^3 + 1/\lambda^2 + 1/\lambda + 3 + 1/\lambda + 1/\lambda^2 + 1/\lambda^3}.$$



(a) log scale



(b) small λ

FIG. 20: Charge neutralisation for the case of overlapped binding. Parameter values: $P_0 = 10000$, $x = 100$.

This landing position creates one gap of length -3 but there is also the possibility of gaps of this size being created when polymers land on the remaining 5-site gap. Therefore the total contribution to overlaps of length 3 from configuration {1} is

$$\frac{1/\lambda^3\left(1 + N_{-3}^{\rm eq}(4,5)\right)}{\frac{1}{\lambda^3} + \frac{1}{\lambda^2} + \frac{1}{\lambda} + 3 + \frac{1}{\lambda} + \frac{1}{\lambda^2} + \frac{1}{\lambda^3}}.$$

Similar formulae hold for for configurations {2}, {3}, {4}. Configuration {5} is a special case since 2 gaps are created. Hence the total contribution from {5} is

$$\frac{2N_{-3}^{\rm eq}(4,1)}{\frac{1}{\lambda^3} + \frac{1}{\lambda^2} + \frac{1}{\lambda} + 3 + \frac{1}{\lambda} + \frac{1}{\lambda^2} + \frac{1}{\lambda^3}}.$$

Arrangements {6}, {7}, {8}, {9} are identical to {4}, {3}, {2}, {1}. Combining the above results we deduce

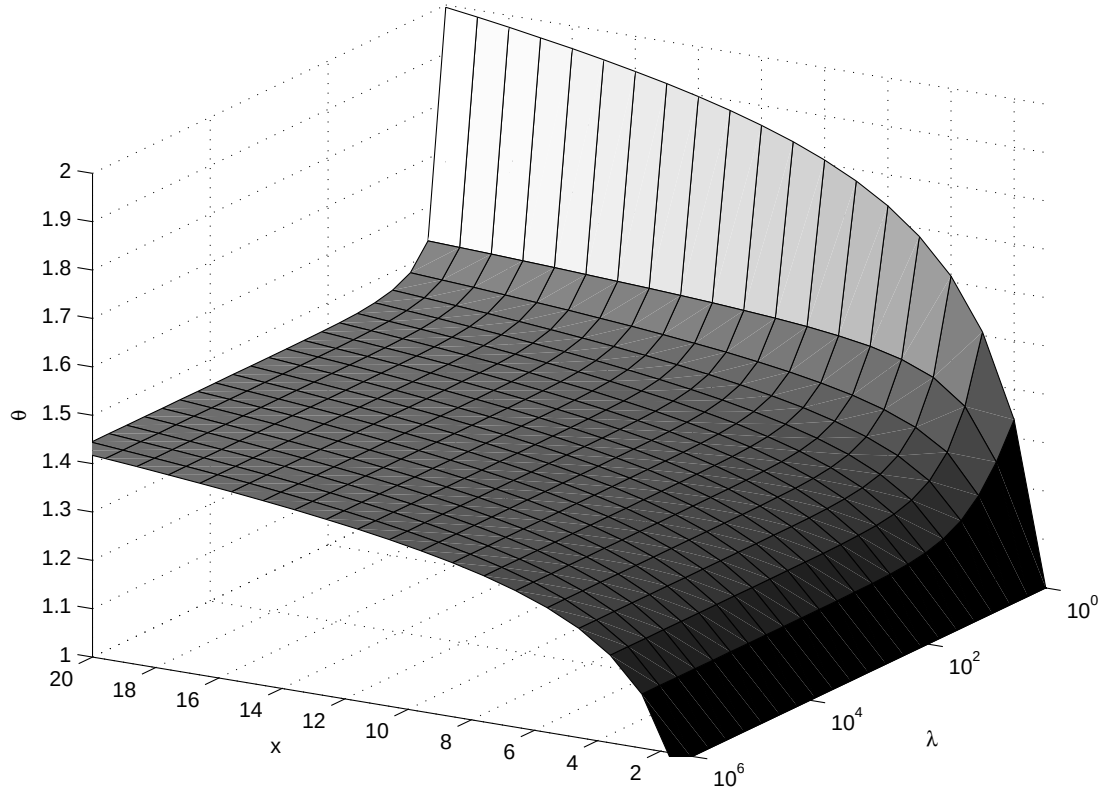


FIG. 21: Surface plot showing how, for a fixed value of P_0 ($P_0 = 2000$), the equilibrium charge neutralisation, θ changes as the polymer length x and binding cooperativity λ vary.

that

$$N_{-3}^{\mathrm{eq}}(4,6) = \frac{2}{\frac{1}{\lambda^3}+\frac{1}{\lambda^2}+\frac{1}{\lambda}+3+\frac{1}{\lambda}+\frac{1}{\lambda^2}+\frac{1}{\lambda^3}} \times \left(\frac{1}{\lambda^3}\left(1+N_{-3}^{\mathrm{eq}}(4,5)\right)+\frac{N_{-3}^{\mathrm{eq}}(4,4)}{\lambda^2} +\frac{N_{-3}^{\mathrm{eq}}(4,3)}{\lambda}+N_{-3}^{\mathrm{eq}}(4,2)+N_{-3}^{\mathrm{eq}}(4,1)\right). \quad (53)$$

Similar reasoning results in recursive relations for steady-state gap distributions for any polymer and DNA length. The number of gaps $N_p^{\mathrm{eq}}(x,P_0)$ of size p is determined by constructing a recurrence relation in which the length of the DNA molecule varies. We do this by first considering a DNA molecule of unit length, ($\widehat{P}_0 = 1$) and increasing $\widehat{P}_0$ until $\widehat{P}_0 = P_0$. It is then possible to show that

$$N_p^{\mathrm{eq}}(x,P_0) = 2\frac{\lambda^p+\sum_{\widehat{P}_0=1}^{P_0-x} N_p^{\mathrm{eq}}(x,\widehat{P}_0)+\sum_{\widehat{P}_0=P_0-x+1}^{P_0-1}\frac{N_p^{\mathrm{eq}}(x,\widehat{P}_0)}{\lambda^{\widehat{P}_0-P_0+x}}}{P_0-x+1+2\sum_{\widehat{P}_0=1}^{x-1}\frac{1}{\lambda^{\widehat{P}_0}}} \quad (x \le P_0), \quad (54a)$$

$$N_p^{\mathrm{eq}}(x,P_0) = 2\frac{\lambda^p+\sum_{\widehat{P}_0=1}^{P_0-1}\frac{N_p^{\mathrm{eq}}(x,\widehat{P}_0)}{\lambda^{\widehat{P}_0-P_0+x}}}{\frac{x-P_0+1}{\lambda^{x-P_0}}+2\sum_{\widehat{P}_0=1}^{P_0-1}\frac{1}{\lambda^{\widehat{P}_0-P_0+x}}} \quad (p+x \le P_0 < x), \quad (54b)$$

$$N_p^{\mathrm{eq}}(x,P_0) = 2\frac{\lambda^{P_0-x}+\sum_{\widehat{P}_0=1}^{P_0-1}\frac{N_p^{\mathrm{eq}}(x,\widehat{P}_0)}{\lambda^{\widehat{P}_0-P_0+x}}}{\frac{x-P_0+1}{\lambda^{x-P_0}}+2\sum_{\widehat{P}_0=1}^{P_0-1}\frac{1}{\lambda^{\widehat{P}_0-P_0+x}}} \quad (2 \le P_0 < p+x), \quad (54c)$$

$$N_p^{\mathrm{eq}}(x,1) = \frac{2}{x}. \quad (54d)$$

Equations (54) allow us to investigate how gap distributions vary with λ. Plots of the gap distribution $N_p^{\rm eq}(x, P_0)$ for DNA of length $P_0 = 2000$, and polymer of length $x = 50$ are shown in Figure 22 for different values of λ. We note that when $\lambda = 1$ a uniform distribution of gap lengths is observed; higher values of λ result in a greater proportion of smaller gaps, and increasing λ above 2 has little additional effect on the gap distribution.

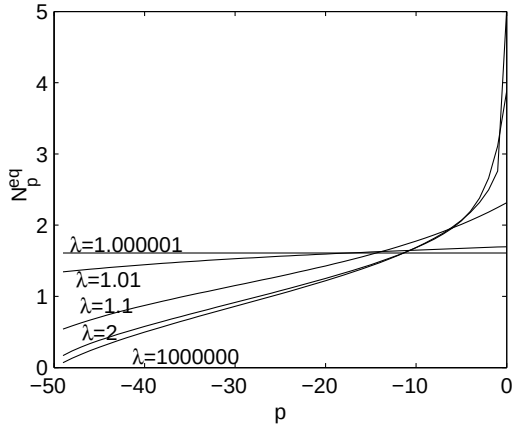


FIG. 22: Series of curves showing how the equilibrium gap distribution changes as the binding cooperativity λ varies. Parameter values: $P_0 = 2000$, $x = 50$.

V. CONCLUSIONS

We have adapted an existing model of binding without overlaps [12] to construct new models of overlapped binding with constant and variable binding rates. We have shown that the simpler mechanism of overlapped binding gives rise to higher coverage in RSA than that allowed by standard RSA models. We believe that in systems such as polymer adsorption onto DNA, overlapped binding is a more likely mechanism than cooperative sequential adsorption (CSA), which is the other mechanism for acheiving the high charge neutralisations required for DNA condensation. In particular we point to charge inversion as the clinching argument, which is easily explained by overlapped binding but impossible both with standard RSA and even with CSA. Overlapped binding has some similarity with multilayer binding models analysed by, for example, Bartelt & Privman [3] and Nielaba [36]. Ideally, of course, a mean field model encompassing both overlapped binding and multilayer adsorption would be useful.

Neither overlapped nor non-overlapped irreversible binding scenarios lead to an 'equilibrium' configuration, rather the dynamics simply leads to a final jammed state, which is degenerate. However, we have characterised the most likely final state using a mean-field approach. The method used in [17] to derive recurrence relations for the equilibrium charge neutralisation has been generalised to determine equilibrium gap length distributions for non-overlapped binding. Similar methods were used to study overlapped binding and to determine the associated charge neutralisation and equilibrium gap length distribution. New exact expressions for DNA charge neutralisation were derived for binding with overlaps and new asymptotic approximations for the case without overlaps were also obtained.

The numerical simulations of irreversible binding without overlaps presented here agree with previous work [15, 16] in that they yield charge neutralisations that are insufficient to condense DNA. Our asymptotic formula for charge neutralisation indicates that it is impossible to achieve the 90% charge neutralisation required to condense the DNA [6] for any polymers when binding is irreversible and there are no overlaps. It is possible to achieve 100% coverage with monomers but the counterion condensation model and experimental studies [31] suggest that such a combination would not condense. This is due to the fact that monomers being polymers of unit length are would behave in a similar manner to monovalent ions, which are known not to cause DNA condensation. 100% coverage with monomers is unlikely to be observed in realistic non-overlapping systems due to monomers having non-zero unbinding and motion rates (giving rise to higher transational entropy).

Since DNA condensation by polymers is observed, there is something lacking in the traditional excluded site binding model: we believe that this is the phenomenon of overlapped binding. We have developed a model that describes the dynamics of the gap distributions that occur when polymers overlap and have shown that this allows higher charge neutralisations to occur than a model which forbids overlapped binding.

Furthermore, overlapped binding can explain charge inversion, where adhered polymers more than neutralise the DNA's negative phosphate charges, and form a complex with net positive charge. This phenomenon, where the charge neutralisation ratio exceeds unity, which has been experimentally observed [22] and cannot be explained by nonoverlapped binding, is entirely consistent with our overlapped binding models.

Simulations involving irreversible binding with overlaps and a binding rate that is independent of the number of bound sites lead to much higher coverage of the DNA molecule than the non-overlapping binding case.

In Section III C we constructed recurrence relations for the distribution of gaps and the charge neutralisation for the case of irreversible binding with overlaps. We found that at equilibrium the gap sizes for overlapped binding were uniformly distributed. The binding rate was then modified to account for the number of sites to which the polymer binds. As a result, the steady-state charge neutralisation reduced to more plausible levels.

In future work, the model of overlapped binding should be calibrated further by comparing experimental results with the solutions from our models to estimate the co-

operativity parameter λ. Alternatively the model could be generalised further by imposing a limit on the smallest gap size in which a polymer can land. A more refined model would determine the rate's dependence on electrostatic DNA-polymer interactions and so be consistent with electrostatic/thermodynamic models of [45]. In forthcoming papers we explain how to generalise the theory and results outlined in this paper to the case of reversible polymer-binding and polymer motion along the DNA plasmid [29] and generalise all these results to the case of polymer mixtures in which the polymer lengths are nonuniform [30].

Acknowledgements

We thank Snjezana Stolnik-Trenkic and Clive Roberts for helpful discussions and the BBSRC for financial support (EM). We are grateful to the referees for bringing to our attention several relevant references.

[1] C-H Ahn, SY Chae, YH Bae & SW Kim. Biodegradable poly(ethilenimine) for plasmid DNA delivery. *J. Control. Release*, **80**, 273–282, (2002)

[2] M Barma. Deposition-evaporation dynamics: jamming, conservation laws and dynamical diversity. In Nonequilibrium Statistical Mechanics in One Dimension, ed V Privman, CUP, (1997), (ISBN 052101834X).

[3] MC Bartelt & V Privman. Kinetics of irreversible monolayer and multilayer adsorption. *Int J Modern Physics B*, **5**, 2883–2907, (1991).

[4] E Ben-Naim & PL Krapivsky. On irreversible deposition of disordered substrates. *J Phys A; Math Gen*, **27**, 3575–3577, (1994).

[5] AU Bielinska, C Chen, J Johnson, JR Baker Jr. DNA complexing with polyamidoamine dendrimers: implications for transfection. *Bioconjugate Chemistry*, **10**, 843–850, (1999).

[6] VA Bloomfield. DNA condensation by multivalent cations. *Biopolymers*, **44**, 269–282, (1998).

[7] B Bonnier, D Boyer & P Viot. Pair correlation function in random sequential absorption process. *J Phys A: Math and Gen*, **27**, 3671–3682, (1994).

[8] SH Bossmann & LS Schulman. Luminescence quenching as a probe of particle distribution. In Nonequilibrium Statistical Mechanics in One Dimension, ed V Privman, CUP, (1997), (ISBN 052101834X).

[9] G Brooks. Gene Therapy. The Use of DNA as a Drug. Pharmaceutical Press, UK, (2001).

[10] R Bruinsma & J Mashl. Long-range electrostatic interaction in DNA-cationic lipid complexes. *Europhysics Letters*, **41**, 165–170, (1998).

[11] SH Cheng & RK Scheule. Airway delivery of cationic lipid: DNA complexes for cystic fybrosis. *Advanced Drug Delivery Reviews*, **30**, 173–184, (1998).

[12] ER Cohen & H Reiss. Kinetics of reactant isolation. I. one-dimensional problems. *J Chem Phys*, **38**, 680–691, (1962).

[13] ST Crooke. Editorial: Delivery of oligonucleotides and polynucleotides. *Journal of Drug Targeting*, **3**, 185–190, (1995).

[14] H Deng, VA Bloomfield, JM Benevides, & GJ Thomas Jr. Structural basis of polyamine-DNA recognition: spermidine and spermine interactions with genomic B-DNAs of different GC content probed by Raman spectroscopy. *Nucleic Acids Research*, **28**, 3379–3385, (2000).

[15] IR Epstein. Kinetics of large-ligand binding to one-dimensional lattices: theory of irreversible binding. *Biopolymers*, **18**, 765–788, (1979).

[16] IR Epstein. Kinetics of nucleic acid-large ligand interactions: exact Monte-Carlo treatment and limiting cases of reversible binding. *Biopolymers*, **18**, 2037–2050, (1979).

[17] IR Epstein. Cooperative and non-cooperative binding of large ligands to a finite one-dimensional lattice, a model for ligand - large oligonucleotide interactions. *Biophys Chem*, **8**, 327–339, (1978).

[18] JW Evans. Random and cooperative sequential adsorption: exactly solvable problems on 1D lattices, continuum limits, and 2D extensions. In Nonequilibrium Statistical Mechanics in One Dimension, ed V Privman, CUP, (1997), (ISBN 052101834X).

[19] S Ferrari, DM Geddes & EWFW Alton. Barriers to and new approaches for gene therapy and gene delivery in cystic fibrosis. *Advanced Drug Delivery Reviews*, **54**, 1373–1393, (2002).

[20] JJ González, PC Hemmer, & JS Hoye. Cooperative effects in random sequential polymer reactions. *Chemical Physics*, **3**, 228–238, (1974).

[21] I Gőssl, L Shu, AD Schlűter & JP Rabe. Molecular structure of single DNA complexes with positively dendronized polymers. *J Am Chem Soc*, **124**, 6850–6865, (2002),

[22] AY Grosberg, TT Nguyen & BI Shklovskii. Colloquium: the physics of charge inversion in chemical and biological systems. *Rev Mod Phys*, **74**, 329–345, (2002).

[23] RC Hoeben. Gene therapy for haemophilia. *Gene Therapy and Molecular Biology*, **1**, 293–300, (1998).

[24] VA Kabanov, VG Sergeyev, OA Pyshkina, AA Zinchenko, AB Zezin, JGH Joosten, J Brackman & K Yoshikawa. Interpolyelectrolyte complexes formed by DNA and astramol poly(propylene imine) Dendrimers. *Macromolecules*, **33**, 9587–9593, (2000).

[25] PL Krapivsky & E Ben-Naim. Collective properties of adsorption-desorption processes. *J Chem Phys*, **100**, 6778–6782, (1994).

[26] Y Levin. Electrostatic correlations: from plasma to biology. *Reports on Progress in Physics*, **65**, 1577–1632, (2002).

[27] I MacLachlan, P Cullis & RW Graham. Progress towards a synthetic virus for systematic gene delivery. *Curr Opin in Molecular Therapeutics*, **1**, 252–259, (1999).

[28] E Maltsev. Mathematical Modelling of DNA Condensation. PhD Thesis, University of Nottingham, (2005).

[29] E Maltsev, JAD Wattis & HM Byrne. DNA charge neutralisation by linear polymers II: reversible binding.

preprint, (2005).

[30] E Maltsev, JAD Wattis & HM Byrne. DNA charge neutralisation by linear polymers III: mixed polymers. *preprint*, (2005).

[31] GS Manning. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Quarterly Review of Biophysics*, **2**, 179–246, (1979).

[32] AL Martin, MC Davies, BJ Rackstraw, CJ Roberts, S Stolnik, SJB Tendler & PM Williams. Observation of DNA-polymer condensate formation in real time at a molecular level. *FEBS Letters* ,**480**, 106–112, (2000).

[33] S Martins & JF Stilck. Attractive forces between circular polyions of the same charge. *Physica A*, **311**, 23–34, (2002).

[34] JD McGhee & PH von Hippel. Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J Mol Biol*, **86**, 469–489, (1974). ,

[35] PD Munro, CM Jackson & DJ Winzor. On the need to consider kinetic as well as thermodynamic consequences of the parking problem in quantitative studies of nonspecific binding between proteins and linear polymer chains. *Biophys Chem*, **71**, 185–198, (1998).

[36] P Nielaba. Lattice models of irreversible adsorption and diffusion. In Nonequilibrium Statistical Mechanics in One Dimension, ed V Privman, CUP, (1997), (ISBN 052101834X).

[37] M Ogris, P Steinlein, S Carotta, S Brunner & E Wagner. DNA/polyethylenimine transfection particles: influence of ligands, polymer size, and PEGylation on internalization and gene expression. *AAPS PharmSci*, **3**(3) E21, (2001).

[38] D Porschke. Dynamics of DNA condensation. *Biochemistry*, **23**, 4821–4828, (1984).

[39] C Pouton, P Lucas, B Thomas, A Uduehi, D Milroy & S Moss. Polycation-DNA complexes for gene delivery: a comparison of the biopharmaceutical properties of cationic polypeptides and cationic lipids. *J Controlled Release*, **53**, 289–299, (1998).

[40] WH Press, SA Teukolsky, WT Vetterling & BP Flannery. Numerical Recipes in Fortran 90. The Art of Parallel Scientific Computing. Cambridge University Press, Volume 2, (1996).

[41] BJ Rackstraw, AL Martin, S Stolnik, CJ Roberts, MC Garnett, MC Davies & SJB Tendler. Microscopic investigations into PEG-cationic polymer-induced DNA condensation. *Langmuir*, **17**, 3185–3193, (2001).

[42] E Raspaud, I Chaperon, A Leforestier & F Livolant. Spermine-induced aggregation of DNA, nucleosome, and chromatin. *Biophys J*, **77**, 1547–1555, (1999).

[43] J Ray & GS Manning. Theory of delocalized ionic binding to polynucleotides: structural and excluded-volume effects. *Biopolymers*, **32**, 541–549, (1992).

[44] A Rényi. On a one-dimensional problem concerning random space-filling. *Publ Math Inst Hung Acad Sci*, **3**, 109–127, (1958).

[45] I Rouzina & VA Bloomfield. Competitive electrostatic binding of charged ligands to polyelectrolytes: planar and cylindrical geometries. *J Phys Chem*, **100**, 4292–4304, (1996).

[46] I Rouzina & VA Bloomfield. Influence of ligand spatial organization on competitive electrostatic binding to DNA. *J Phys Chem*, **100**, 4305-4313, (1996).

[47] I. Rouzina & VA Bloomfield. Competitive electrostatic binding of charged ligands to polyelectrolytes: practical approach using the non-linear Poisson-Boltzmann equation. *Biophys Chem*, **64**, 139–155, (1997).

[48] U Rungsardthong. Physiochemical evaluation of polymer-DNA complexes for DNA delivery. PhD Thesis, University of Nottingham, (2002).

[49] J Widom & RL Baldwin. Cation-induced toroidal condensation of DNA: studies with $Co^{3+}(NH_3)_6$. *J Mol Biol*, **144**, 431–453, (1980).

[50] RW Wilson & VA Bloomfield. Counterion-induced condensation of Deoxyribonucleic acid, A light-scattering study. *Biochemistry*, **18**, 2192–2196, (1979).

[51] CP Woodbury Jr. Free-sliding ligands: an alternative model of DNA-protein interactions. *Biopolymers*, **20**, 2225-2241, (1981).