INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) L181–L187 PII: S0305-4470(03)56211-4

LETTER TO THE EDITOR

Helicase on DNA: a phase coexistence based mechanism

Somendra M Bhattacharjee[1] **and Flavio Seno**[2]

[1] Institute of Physics, Bhubaneswar 751 005, India
[2] INFM-Dipartimento di Fisica, Università di Padova, Via Marzolo 8, 35131 Padova, Italy

E-mail: somen@iopb.res.in and flavio.seno@pd.infn.it

Received 14 November 2002, in final form 20 January 2003
Published 19 March 2003
Online at stacks.iop.org/JPhysA/36/L181

Abstract

We propose a phase coexistence based mechanism for activity of helicases, ubiquitous enzymes that unwind double stranded DNA. The helicase–DNA complex constitutes a fixed-stretch ensemble that entails the coexistence of domains of zipped and unzipped phases of DNA, separated by a domain wall. The motor action of the helicase leads to a change in the position of the fixed constraint thereby shifting the domain wall on dsDNA. We associate this off-equilibrium domain wall motion with the unzipping activity of the helicase. We show that this proposal gives a clear and consistent explanation of the main observed features of helicases.

PACS numbers: 87.15.Aa, 87.15.He, 87.14.Gg

Nucleic acid helicases are defined as enzymes that translocate directionally through double stranded nucleic acid substrates to catalyse the separation of the complementary strands. They facilitate various biological processes such as DNA replication, recombination and repair, RNA transcription, editing and splicing [1]. There are several structural varieties of helicases such as monomeric (e.g., PcrA), dimeric (e.g., Rep), trimeric (e.g., RecBCD), tetrameric (e.g., RNA polymerase) or closed hexameric (e.g., DnaB), but all use the hydrolysis[3] of ATP to ADP as the preferred source of energy [1–8].

Bulk behaviour in solutions such as average unwinding rates, step size, average number of base-pairs opened per helicase, etc, are known for a few helicases such as the Hepatitis-C virus helicase [6], PcrA [7], DnaB [5, 10] and others. Much attention has recently been devoted [12–14] towards a quantitative characterization of RecBCD enzymes, a multifunctional trimeric protein complex (the products of the recB, recC and recD genes [11]) that participates in the repair of chromosomal DNA through homologous recombination. In bacteria, such as *Escherichia coli*, RecBCD is involved, e.g., in protection against damage by UV or

[3] In general, some nucleoside triphosphate (NTP) is required, e.g. T7 gp4 can use GTP, SV40 large T antigen uses non-ATP nucleotides for unwinding RNA [9].

gamma irradiation, and infection by bacteriophages. In all cases, the full functionality of RecBCD relies on the helicase and the nuclease actions of its subunits. The use of single-molecule, micromanipulation tools allowed for monitoring in detail the translocation [12, 14], unwinding [13] and processivity (rate of dissociation) [13] of individual RecBCD enzyme molecules on dsDNA. Such experiments elucidated several new aspects of helicase behaviour and showed that many properties could be more related to general principles than on specific chemical details. In particular, it has been observed that (a) RecBCD unwinds dsDNA at a uniform rate, over a wide range of ATP concentrations, as it moves on one strand, (b) the nuclease activity does not affect unzipping and (c) the helicase can work in the presence of DNA gaps up to certain lengths. More recently, winding–rewinding for *E. coli* Rep helicase–DNA complex [8] has been observed at a single molecule level.

Despite these varieties of experimental findings no clear mechanism coupling the motor action and the helicase activity is known yet. To fill this gap, in this letter we present a simple, but powerful argument, based on the principle of phase coexistence [15–19], that provides a clear and robust explanation for the gross observed features. There are a few biological operational models built on how a helicase presumably might work [5]. The analysis reported here gives a thermodynamic basis to a model called the 'wedge model' according to which the motion of the helicase 'provides enough force to enable the helicase to destabilize the base pairs at the junction by a process resembling the action of a wedge' [5]. In our proposed mechanism, energy (from ATP) is required for translocation activity or the motor action of the helicase and not directly for base pair breaking and therefore, according to the classification scheme of [3], this corresponds to passive helicases. Additional features required for active helicases are ignored in this first study. Our proposal is supported by computer simulation of an exactly tractable model. To our knowledge, this is the first theoretical study of the dynamics of a DNA–helicase complex.

We study the joint dynamics of the helicase and the DNA in a two-dimensional fork model (Y-model) [16, 17]. The two strands of DNA are represented by two directed and mutually avoiding walks formed by N bases each. In two dimensions, on the square lattice (see figure 1) this implies that the two walks follow the positive direction of the diagonal axis (z); in other words the coordinate along such a direction always increases. The perpendicular direction x measures, in units of the elementary square diagonal, the distance between the two complementary monomers belonging to the two strands. When this distance is equal to 1 they are considered in contact: a binding energy is gained which is uniform ($\epsilon = 1$) for homogeneous model of DNA (homo-DNA) but chosen randomly from two different values (ϵ_1, ϵ_2) for heterogeneous DNA (hetero-DNA). Note that due to the geometrical properties of the lattice the two complementary monomers are labelled by the same z-coordinate, as one would require for base pairing in DNA. In the Y-model unzipping can occur only processively, e.g. bubbles are suppressed along the chain: the only conformations considered have the first $N - m$ monomers bounded, whereas the remaining m are separated in a Y-like conformation. The fact that the Y-model does not allow rejoining of the unzipped portion of the dsDNA is similar to the geometry observed in the experiment of [13] (see also figure 4 of [20]). Also, bubbles are suppressed for DNA at temperatures much below its melting temperature T_m, a temperature below which the two strands are zipped. In the case of homogeneous interaction the exact phase diagram and other static and dynamical quantities can be exactly determined (also in the presence of a stretching force) [16, 17].

The coarse-grained nature of the model [21] needs to be stressed here. Monomers are to be thought of as groups of bases, and ignoring helicity or restricting to two dimensions,

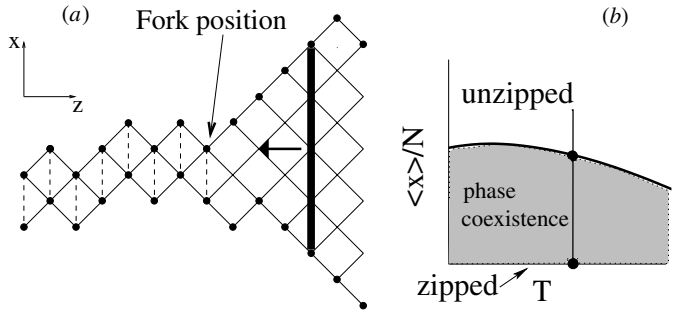

Figure 1. (*a*) A typical configuration of the simulated DNA with the helicase (thick rod) as modelled on a square lattice (thin lines). Bases are represented by dots and paired bases are shown by dotted lines. The position of the fork coincides with the last paired base. The arrow indicates the motion of the helicase. Except for the rigid hard-core constraint, there is no other direct interaction between the DNA and the helicase. (*b*) Schematic phase diagram in the separation (stretch) versus temperature plane. At fixed temperature, a finite endpoint separation leads to a coexistence of the two phases indicated by the filled circles. The interface of the two phases is the domain wall.

are more for simplification of the calculation than artefacts[4]. Such coarse-grained models, or even simpler ones, are used in various DNA related problems [19, 24] and even in analysis of thermal melting of DNA [23]. The spirit behind our approach is that the key element that can influence universal behaviour of helicase translocation is the competition between a Y-fork conformation which can be unzipped by paying energy and the movement of an opening machine.

Several studies of theoretical DNA models [15–19, 22] have established the existence of a sharp unzipping phase transition of a dsDNA at a critical stretching force applied at one end on the two strands. This implies that in the conjugate ensemble of a fixed-separation constraint for the two strands, there is a coexistence of domains of zipped and unzipped phases. For the fixed-force ensemble for a homo-DNA in the large length limit ($N \to \infty$), the critical force g_c for the zipping–unzipping transition, and the end separation under a force $g > g_c$ are given by [16, 17]

$$g_c(T) = \frac{T}{2}\cosh^{-1}(\mathrm{e}^{1/T} - 1) \qquad \text{and} \qquad \frac{\langle x\rangle}{N} = \tanh\frac{g}{2T} \tag{1}$$

where temperature T is measured in units of k_B/ϵ, k_B being the Boltzmann constant. From these equations the phase coexistence curve $\langle x\rangle/N$ versus T can be determined exactly and it is schematically shown in figure 1(*b*). Under a fixed-distance constraint, represented by the vertical line in figure 1(*b*), the DNA chain splits into domains of zipped and unzipped phases. The length of the unzipped strand can be read off from the upper line of the coexistence curve. This fact can be checked independently and we have verified it not only for the Y-model, but also for models that allow bubbles.

We now make our hypothesis. If we think that the helicase, by virtue of its size larger than the separation of the two DNA strands and the excluded volume interaction, acts as a geometrical separator, *the helicase–DNA complex constitutes exactly a fixed-stretch ensemble*. The unwinding activity can then be simply associated with the motion of the domain wall which necessarily forms and follows the motor action of the helicase. In other words, the helicase plays a double role: firstly, its presence thermodynamically implies the existence of a

[4] The qualitative features of the unzipping phase transition of [15–17] are observed in more complex models in [22].

domain wall, and secondly, its translocation induces a motion of the domain wall to reach its equilibrium position. Such an approach puts the primary role on the translocation motion. The thermodynamic force, that drives the domain wall towards its equilibrium position, provides the mechanism for base pair opening.

To verify our hypothesis we have numerically studied the DNA model described above mimicking the effects of the helicase with a rod of length l (see figure 1(a). Dynamics is introduced by a Monte Carlo procedure. For the DNA, one among the $2N$ monomers is randomly chosen and an attempt is made to modify its position with respect to all the others (which remain fixed). The move can in principle increase or decrease by one unit the distance between the strands. The move is accepted according to standard Metropolis rules. A Monte Carlo unit time is defined as $2N$ single-monomer attempted moves. The scaling properties of this DNA dynamics, also in the presence of a stretching force, have already been determined [16]. The helicase moves forward (motor action) by unit steps, along the $-z$ direction, on the DNA, if it is not hindered by the chain configuration (excluded volume interaction). The motion is kept unidirectional as found for RecBCD in [12] (see below also). The motion of the helicase is attempted at every Monte Carlo step. In our simulation we start with a configuration where the helicase is attached to one end of a dsDNA (as in the experiments of [12–14]). The temperature is maintained constant for the DNA to be in the zipped phase (below T_m). The overall dynamics is off-equilibrium. Note that there is no specific interaction but both DNA and helicase dynamics are constrained by the excluded volume.

Throughout the simulation we monitored the position of the last zipped base pair (fork position, see figure 1(a)), the zipping probability $p(s, t)$ that a base pair at site s at time t is zipped, and the average position of the helicase at time t. In most simulations the length l of the helicase was taken equal to 6 (though we studied up to $l = 14$) whereas 1000 thermal averages were necessary. The length N of the strands was varied according to the different experiments ranging up to $N = 1000$. For an analysis of the domain wall, the zipping probability $p(s, t)$ can be fitted by a function:

$$p(s, t) = \tfrac{1}{2}[1 - \tanh\{(s - s_0(t))/w(t)\}] \tag{2}$$

where $s_0(t)$ and $w(t)$ are the position and the width, respectively, of the domain wall.

In figure 2 we plot for two different times the zipping probability $p(s, t)$ for the homogeneous and for heterogeneous cases. In all cases the domain wall behaviour predicted is well fitted by a tanh profile (equation (2)). We point out that such domain walls are not found if the DNA strands are noninteracting, i.e. $\epsilon = 0$ (see inset of figure 2). In figure 3 we show the instantaneous positions of the helicase and the domain wall (computed through equation (2)) as a function of time. Movement starts at $t = 0$ when the helicase is loaded. The two quantities proceed uniformly and cooperatively through the DNA unwinding it. The domain wall evolves towards the equilibrium position whenever the helicase motion is stopped and this position turns out to be not very far away from the instantaneous position. This indicates an adiabatic adjustment of the domain wall to the instantaneous position of the mobile helicase. We stress that had there been no interaction between the DNA and the helicase, the latter, as a phantom motor, would have moved with the assigned speed with the DNA remaining bound (since we are below T_m). It is important to compare this motion with the unzipping dynamics in a fixed-force ensemble which shows a characteristic scaling behaviour [16], namely a nonlinear evolution $m(t) \sim t^{1/3}$, where $m(t)$ is the number of unzipped bases at time t. In contrast, and this is the central point of our work, we find that the combined dynamics involving the excluded volume interaction between the helicase and the DNA (but no external force) leads to the uniform motion of both the helicase and the domain wall. Our results of figure 3 should be compared with figure 2 of [13]. The effective velocity is smaller than the unhindered one

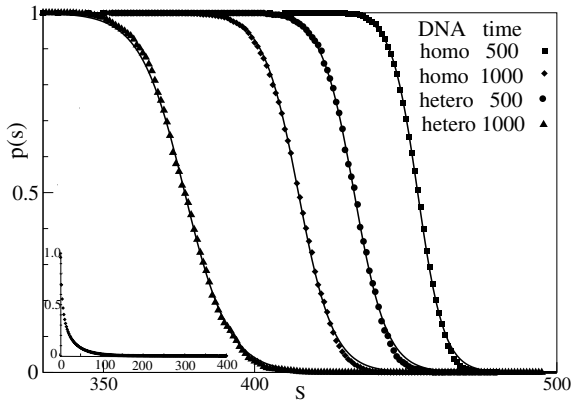

Figure 2. Zipping probabilities at two different times (after thermal averaging) for the homo-DNA ($\epsilon = 1$) and hetero-DNA ($\epsilon_1 = 1, \epsilon_2 = 0.5$). The solid lines are fits to equation (2). The inset shows the equilibrium zipping probabilities (no domain wall) for the case of $\epsilon = 0$ when the ends at $z = 400$ are kept separated.

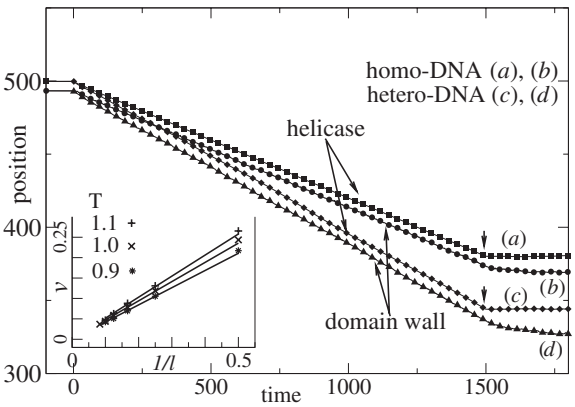

Figure 3. Positions of the helicase and the domain wall as a function of time for (*a*), (*b*) homo- and (*c*), (*d*) hetero-DNA. The size of the helicase is $l = 6$. When the helicase is stopped at time = 1000 (indicated by the arrows), the domain wall evolves towards the equilibrium position. The inset shows the speed of the helicase for different temperatures and lengths, with the solid lines representing $v = 1.28[\exp(-1/T)]/l$.

and depends on the size l of the helicase, the temperature and the sequence. The l and T dependence can be estimated if we think that to move the helicase we have to unzip a base pair at the fork position, and allow the formed kink to reach the position of the helicase. The base pair is broken with a probability $\exp\left(-\frac{1}{T}\right)$ and the kink needs a time of order l to reach the helicase. Therefore, the velocity is proportional to $\left[\exp\left(-\frac{1}{T}\right)\right]/l$ (confirmed by the fit of our data, see the inset of figure 3). A strong dependence on T has also been found in [13]. In our simulation no sequence-dependent nonuniformity in motion was ever discernible so long as the heterogeneity was uncorrelated. The velocity we observed, originating solely from the motor action, is a lower bound because any periodic conformational change of the helicase [12] during its motion (ignored here mainly to illustrate the role of the domain wall) would assist the motion of the domain wall itself.

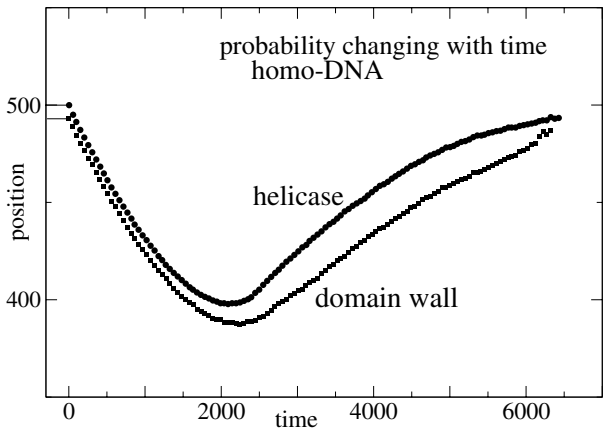

Figure 4. Positions of the helicase ($l = 6$) for homogeneous DNA with a biased probability $P(t)$. Probability decreases linearly from 1 for $t = 0$ to $\frac{1}{2}$ for $t \approx 2500$ after which it remains constant at $1/2$.

Our proposal lends itself to several predictions. We discuss a few here. Since the position $s_0(t)$ of the wall is determined by the location of the stretching constraint put up by the helicase, the unzipped part beyond the helicase should not affect the action of the helicase. We verified this explicitly in our simulation: a part of the unzipped DNA beyond the helicase is chopped off (which mimics RecD activity [12]) at arbitrarily chosen times and there is neither any change in the nature of the wall (e.g. $p(s, t)$) nor in the rate of unzipping. This agrees with the observations in [12, 13] that the nuclease activity does not modify the unwinding action of RecBCD. Next, the width of the wall, as defined by equation (2), gives a length-scale for the helicase activity. The sequence randomness in a heterogeneous chain does not matter if there is no correlation beyond this scale, as found both in simulation and real experiments. We have seen a periodic modulation (not shown) in the helicase motion if the DNA sequence is periodic of two pairing energies with periodicity larger than the width of the domain wall. Another consequence of this scale is that a nick or break smaller than the domain wall width will not be recognized, providing an interpretation of the observations in [14] regarding the size of gap a helicase can negotiate.

We have also simulated cases where the helicase undergoes a biased random walk-type motion as expected for RecQ [25]. We introduced a probability $P(t)$ that the helicase could step along the $-z$ direction on the DNA (when this is sterically acceptable) but also a probability $1 - P(t)$ that it could move in the opposite direction (i.e., away from the domain wall). When $P(t)$ is kept fixed at a value P_0 in the interval $1 \geqslant P_0 \geqslant \frac{1}{2}$ ($P_0 = 1$ being the case studied in the first part of the letter) we find again a uniform motion with velocity related to P_0. At $P_0 \equiv \frac{1}{2}$ (random walk), the situation changes: unzipping does not proceed and the helicase dissociates from the double strand. This behaviour is illustrated in figure 4 where we plot the position of the helicase as a function of time: $P(t)$ decreases linearly from 1 down to $\frac{1}{2}$ and then remains constant. As expected the unwinding proceeds nonuniformly until the random walk regime is reached. At that time the helicase is discarded and the DNA zips again. It suggests, though a bit speculative, that a probability affecting the forward motion of the helicase could be a phenomenological characterization of the gradual wobbliness of the helicase on the track.

In conclusion, we have shown that associating the helicase activity with the domain wall motion in a fixed-stretch ensemble accounts for several observed features, such as

e.g., the uniformity of unzipping, no sequence-dependent nonuniformity and the insensitivity to nuclease action, without any requirement of extra specific bond-cutting chemical processes. A domain wall also gives a quantitative meaning to the Y-fork in the terminology of DNA replication. The underlying thermodynamic basis gives a robustness to the mechanism that could be at work for hexameric helicases also.

Acknowledgments

FS was supported by MIUR-COFIN01. FS thanks the Institute of Physics, Bhubaneswar for kind hospitality.

References

[1] von Hippel P H and Delagoutte E 2001 *Cell* **104** 177–90
[2] Matson S W, Bean D W and George J W 1994 *Bioessays* **16** 13–22
[3] Lohman T M and Bjornson K P 1996 *Ann. Rev. Biochem.* **65** 169–214
[4] West S C 1996 *Cell* **86** 177–80
[5] Patel S S and Picha K M 2000 *Ann. Rev. Biochem.* **69** 651
[6] Porter D J T *et al* 1998 *J. Biol. Chem.* **273** 18906
[7] Soultanas P *et al* 2000 *EMBO J.* **19** 3799
[8] Ha Taekjip *et al* 2002 *Nature* **419** 638
[9] Scheffner M, Knippers R and Stahl H 1991 *Eur. J. Biochem.* **195** 49
[10] Kim S *et al* 1996 *Cell* **84** 643
[11] Amundsen S K, Taylor A F, Chaudhury A M and Smith G R 1986 *Proc. Natl Acad. Sci. USA* **83** 5558–62
[12] Dohoney K M and Gelles J 2001 *Nature* **409** 370–4
[13] Bianco P R, Brewer L R, Corzett M, Balhorn R, Yeh Y and Kowalczykowski S C 2001 *Nature* **409** 374–8
[14] Bianco P R and Kowalczykowski S C 2000 *Nature* **405** 368–72
[15] Bhattacharjee S M 2000 *J. Phys. A: Math. Gen.* **33** L423
Bhattacharjee S M 2000 *J. Phys. A: Math. Gen.* **33** 9003(E)
[16] Marenduzzo D, Bhattacharjee S M, Maritan A, Orlandini E and Seno F 2002 *Phys. Rev. Lett.* **88** 028102
[17] Marenduzzo D, Trovato A and Maritan A 2001 *Phys. Rev.* E **64** 031901
[18] Bhattacharjee S M and Marenduzzo D 2002 *J. Phys. A: Math. Gen.* **35** L141
[19] Sebastian K L 2000 *Phys. Rev.* E **62** 1128
Zhou H 2000 *Preprint* cond-mat/0007015
Lubensky D and Nelson D R 2002 *Phys. Rev.* E **65** 031917
Cocco S, Monasson R and Marko J F 2001 *Proc. Natl Acad. Sci. USA* **98** 8608
[20] Henn A, Medalia O, Shi S P, Franceschi F and Sagi I 2001 *Proc. Natl Acad. Sci. USA* **98** 5007–12
[21] Poland D and Scheraga H A 1966 *J. Chem. Phys.* **45** 1464
[22] Orlandini E, Bhattacharjee S M, Marenduzzo D, Maritan A and Seno F 2001 *J. Phys. A: Math. Gen.* **34** L751
[23] Daune M 1999 *Molecular Biophysics: Structures in Motion* (Oxford: Oxford University Press)
[24] Causo M S, Coluzzi B and Grassberger P 2000 *Phys. Rev.* E **62** 3958
Carlon E, Orlandini E and Stella A L 2002 *Phys. Rev. Lett.* **88** 198101
[25] Harmon F G and Kowalczykowski S C 2001 *J. Biol. Chem.* **276** 232