

Designs of Autonomous Unidirectional Walking DNA Devices*

Peng Yin¹, Andrew J. Turberfield², and John H. Reif¹

¹ Department of Computer Science, Duke University
Box 90129, Durham, NC 27708-0129, USA.
{py, reif}@cs.duke.edu

² University of Oxford, Department of Physics, Clarendon Laboratory
Parks Road, Oxford OX 1 3PU, UK.
a.turberfield@physics.ox.ac.uk

Abstract. Imagine a host of nanoscale DNA robots move autonomously over a microscale DNA nanostructure, each following a programmable route and serving as a nanoparticle and/or an information carrier. The accomplishment of this goal has many applications in nanorobotics, nano-fabrication, nano-electronics, nano-diagnostics/therapeutics, and nano-computing. Recent success in constructing large scale DNA nanostructures in a programmable way provides the structural basis to meet the above challenge. The missing link is a DNA walker that can autonomously move along a route programmably embedded in the underlying nanostructure – existing synthetic DNA mechanical devices only exhibit localized non-extensible motions such as bi-directional rotation, open/close, and contraction/extension, mediated by external environmental changes. We describe in this paper two designs of autonomous DNA walking devices in which a walker moves along a linear track unidirectionally. The track of each device consists of a periodic linear array of anchorage sites. A walker sequentially steps over the anchorages in an autonomous unidirectional way. Each walking device makes use of alternating actions of restriction enzymes and ligase to achieve unidirectional translational motion.

1 Introduction

A major challenge in nanotechnology is to precisely transport a nanoscale object from one location on a nanostructure to another location following a programmable path. DNA has been explored as an excellent building material for the construction of both large scale nanostructures and individual nanomechanical devices [11]. The successful constructions of two dimensional DNA lattices and one dimensional DNA arrays made from DX molecules [16], TX molecules [5], rhombus molecules [8], and 4x4 molecules [17] provide the structural base for realization of the above goal. However, most existing DNA nanomechanical devices only exhibit localized non-extensible motions such as open/close [7, 13, 14, 21], extension/contraction [1, 4, 6], and reversible rotation motion [9, 18]. Furthermore, these motions are not autonomously executed but rather mediated by external environmental changes such as the addition and removal

* Extended abstract. For full version, see [19].

of DNA fuel strands [1, 4, 6, 13, 14, 18, 21] or the change of ionic compositions of the solution [7, 9]. Autonomous unidirectional DNA devices executing linear translational motions are hence desirable.

There are already some exciting progress in this direction. Turberfield and colleagues have proposed to use DNA fuels to design autonomous free running DNA machines [15]. Reif has described theoretical designs of autonomous DNA walking and rolling devices that demonstrate random bidirectional translational motion along DNA tracks [10]. On the experimental side, Mao's group has recently constructed an autonomous DNA motor powered by a DNA enzyme [3]; Seeman's group has constructed a DNA walking device mediated by DNA fuel strands [12].

In the rest of the paper, we present two designs of autonomous DNA walking devices. Each device consists of a track and a walker. The track of each device contains a periodic linear array of anchorage sites. A walker sequentially steps over the anchorages in an autonomous unidirectional fashion. Each walking device makes use of alternating actions of restriction enzymes and ligase to achieve unidirectional translational motion. The action of ligase consumes ATP as energy source. The walking devices described here make the following improvements over the walking device presented in [10]. Firstly, they demonstrate unidirectional motion rather than random bidirectional motion. Secondly, the moving part (walker) in each walking device is a physical entity with a flexible body size rather than a symbolic entity, and thus the walker can serve not only as an information carrier but also as a nanoparticle carrier. These walking device designs are also different from the walking device construction by Seeman's group [12] in that they are autonomous. A limitation of our first device is that it has a low probability of falling off the track. Our second device has zero probability of falling off the track, but it is a more complicated (hence less practical) construction and assumes a restriction enzyme property that has not yet been fully-substantiated. For each walking device, we first present its structure and operation, and then describe its implementation using conceptual enzymes followed by one or more concrete examples using commercially available enzymes. The design using conceptual enzymes illustrates the general principle of the design and reveals the essential information encoding of the device that dictates its operation, while the examples using real enzymes attempt to validate the practicality of the design principles and to illustrate some technical complications in mapping the conceptual design to real enzymes.

=====

2 Definitions

A basic structural unit used in the construction of the walking devices is a *dangler*. A dangler is a duplex DNA fragment with single strand extensions at both ends: one end is the *fixed end* that is usually attached to another structural unit (*e.g.* the backbone of the track or the body of the walker); the other end is the *sticky end*. The flexible single strand DNA at the fixed end allows the otherwise stiff dangler to move rather freely around the fixed end. This property is crucial to the operation of the devices. The fixed end only serves to structurally join a dangler to another component of the device in a flexible fashion (*e.g.* the linkage of an anchorage to the backbone of the track/the linkage of a

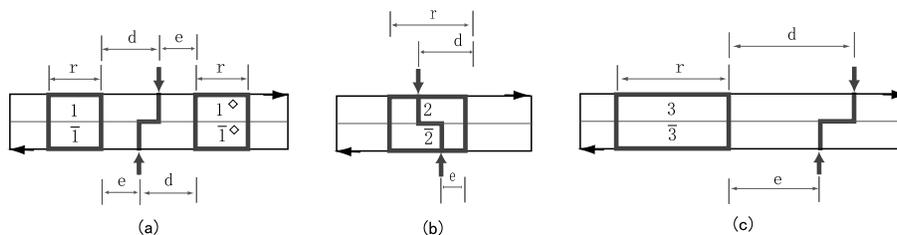


Fig. 1. Conceptual endonucleases used in the construction of the walking devices. The sequences constituting the recognition site of the endonuclease in (a), (b), and (c) are labeled with 1, $\bar{1}$, 1^\diamond , and $\bar{1}^\diamond$; 2 and $\bar{2}$; and 3 and $\bar{3}$, respectively. Symbols r , d , and e are signed length parameters in number of bases. Recognition sites and cleavage sites are indicated with bold boxes and pairs of bold arrows, respectively. N indicates the position of a base whose value does not affect recognition by an endonuclease

foot to the body a walker); the sticky end, in contrast, usually encodes information and participates actively in dictating the motion of the walker.

Two basic operational events driving the unidirectional motion of the devices are *ligations* and *cleavages*. Two neighboring dangles with complementary sticky ends can associate with each other via the hybridization of their sticky ends. Subsequent to this hybridization, the nicks at either end of the hybridized section can be sealed by a *ligase* and the two duplex fragments are hence joined covalently in a process known as *ligation*. When the context is clear, the whole process of hybridization and subsequent ligation (joining of two DNA strands) is referred to as ligation, for simplicity. In *cleavage*, an approximately reverse process to ligation, a duplex DNA fragment is cut into two separate duplex parts (with each usually possessing a complementary sticky end) by enzymes known as *restriction endonucleases*. Following cleavage, the two duplex DNA fragments (each with a sticky end) can dissociate in a process known as *melting*. When the context is clear, the whole process of cleavage and subsequent melting is referred to as cleavage. Figure 1 illustrates the conceptual restriction enzymes used in the construction of our devices. Cleavage uses no energy input from external environment while ligation consumes one molecule of ATP as energy source.

=====

3 Device I

Design overview. Device I consists of two parts: the *track* and the *walker*. The walker is the moving part of the device while the track is the immobile part along which the walker moves. Figure 2 (a) gives a schematic drawing of the structure of device I. The track contains a linear array of anchorages, *A* and *B*. Each anchorage is a duplex DNA fragment with a sticky end on the top, and rigidly attached to the backbone of the track. The walker stands on top of the track. The walker consists of two parts, the body and the feet (a *front foot C* and a *hind foot D*). The body is a duplex DNA segment and each foot is a DNA dangler tethered to the body via a flexible single strand DNA joint.

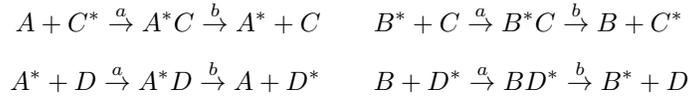
The flexible joint allows a foot of the walker to rove to and only to the two anchorages immediately neighboring the current anchorage on which it stands. The sticky end of a foot is complementary to the sticky end of the anchorage on which it stands and hence the foot can hybridize and be ligated with the anchorage. The ligation product between a foot and an anchorage will be cut by an endonuclease such that both the foot and the anchorage change their sticky ends. As a result, the foot will possess a sticky end that is complementary to the sticky end of the anchorage immediately ahead of the anchorage X on which the foot has been standing, but *not* complementary to the sticky end of the anchorage immediately behind X . Consequently, the foot can only hybridize and be ligated with the anchorage immediately ahead of X , but not with the one immediately behind it. This guarantees the forward motion of the walker. See Figure 2 (b) for illustration.

A foot or an anchorage X can exist in two forms, X and X^* , where $X = A, B, C$, and D . X^* is derived from X by altering its sticky end. X and X^* are required to satisfy certain properties that will be described later. At any moment during the motion, the track in front of the front foot C and behind the hind foot D consists of alternating dangles A and B^* while the track between them consists of

alternating A^* and B . Assume w.l.o.g. that at the start of the motion, both feet C and D are ligated with anchorages of type A , forming A^*C and A^*D respectively. Thus the initial configuration of the walker and track complex can be written as,

$$(AB^*)_i[A^*D]B(A^*B)_j[A^*C]B^*(AB^*)_k$$

where $[A^*C]$ (resp. $[A^*D]$) is the complex between anchorage A^* and the front foot C (resp. hind foot D). To make the walker move unidirectionally down the track, we implement the following reactions between a foot and an anchorage,



In phase a of each reaction, a foot is ligated with an anchorage; in phase b , the foot and the anchorage are cut separate by a restriction enzyme, each now possessing a new sticky end. Applying the reactions to the walker-track complex, we have the following motion of the walker along the track,

$$\begin{aligned} & (AB^*)_i[A^*D]B(A^*B)_j[A^*C]B^*(AB^*)_k \\ & \rightarrow (AB^*)_iA[B^*D](A^*B)_jA^*[B^*C](AB^*)_k \\ & \rightarrow (AB^*)_{i+1}[A^*D]B(A^*B)_j[A^*C]B^*(AB^*)_{k-1} \end{aligned}$$

The above is a full induction cycle of the motion of the walker, and hence the walker can (in principle) move forward along the track infinitely. We further require that phase a of each reaction is *not* reversible, thus the whole reaction is irreversible. Consequently, the walker can only move unidirectionally along the track.

Note that the hind foot D can be viewed as the dual to the front foot C : during the motion, front foot C changes the configuration of the track from (AB^*) to A^*B ; hind foot D moves on the modified track and restores it to its original configuration AB^* .

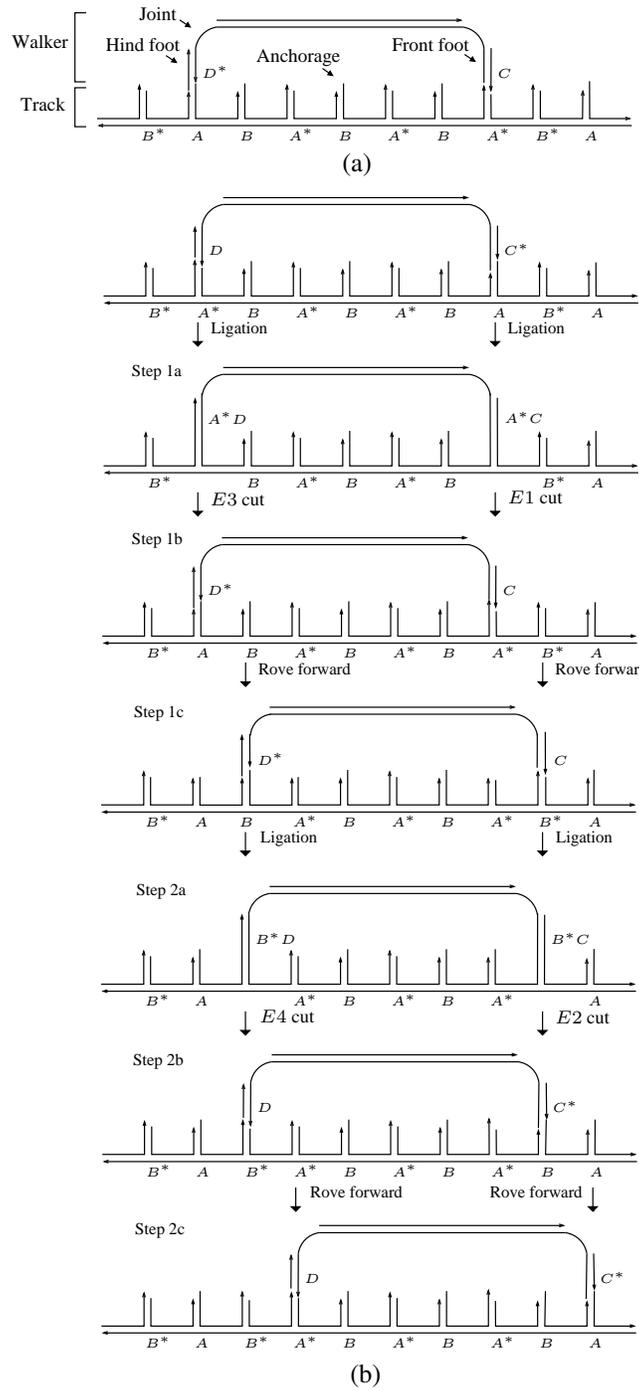


Fig. 2. The structural design and step by step operation of device I. (a) Structural design of the device. (b) Step by step operation of the device. The walker moves unidirectionally to the right

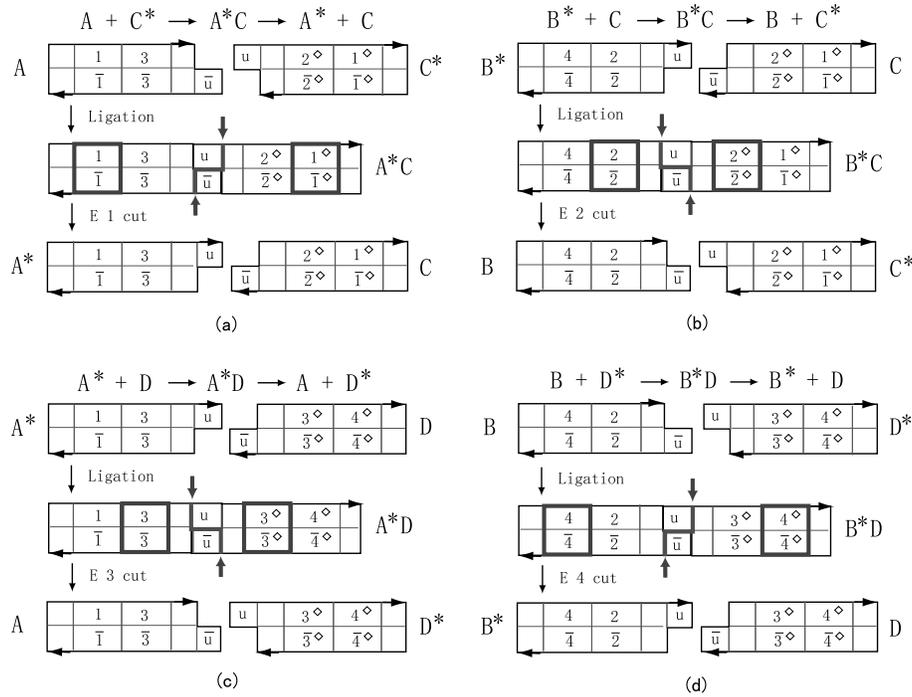


Fig. 3. Implementation of device I using four conceptual restriction enzymes. Endonuclease recognition sites and cleavage sites are indicated with bold boxes and pairs of bold arrows, respectively

Implementation with conceptual endonucleases. To implement the designed reactions, we use four conceptual enzymes $E1$, $E2$, $E3$, and $E4$. The cutting patterns of these enzymes are similar to the one depicted in Figure 1 (a). Here we require that $d_1 - e_1 = d_4 - e_4 = e_2 - d_2 = e_3 - d_3$, where d_i and e_i are the length parameters for endonuclease Ei . Figure 3 describes the detailed step by step reactions that dictate the motion of the walker. Since only the region near the end of an anchorage or a foot is relevant for the reactions, we only depict the end regions in Figure 3.

Figure 3 (a) depicts reaction $A + C^* \rightarrow A^*C \rightarrow A^* + C$. In this reaction, the sticky end \bar{u} of anchorage A is first ligated with the sticky end u (complementary to \bar{u}) of foot C^* , generating ligation product A^*C . This corresponds to the reaction of the front foot in Step 1a in Figure 2 (b): $A + C^* \rightarrow A^*C$. A^*C contains a recognition site for endonuclease $E1$ and is cut by $E1$ into A^* and C (Step 1b in Figure 2 (b): $A^*C \rightarrow A^* + C$). Note that now front foot C possesses a new sticky end \bar{u} . Recall that the anchorage immediately ahead of the anchorage A^* , on which front foot C is standing, is anchorage B^* , which possesses a sticky end u (complementary to \bar{u}). Thus C can rove forward and hybridize with B^* (Step 1c in Figure 2 (b)). This leads to the reaction in Figure 3 (b): $B^* + C \rightarrow B^*C \rightarrow B + C^*$. First, the hybridization product

Table 1. Implementation of device I with endonucleases Ahd I, Fnu4H I, ScrF I, and Xcm I. Ligation sites and cleavage sites are denoted with — and ^, respectively. The bases that determine recognition sites in action are in upper case

Reactions	Enzymes	DNA Sequences
$A + C^* \rightarrow A^*C$	Ligase	5'...gaccn-gcgtc... 3' 3'...ctgggn-cgcag... 5'
$A^*C \rightarrow A^* + C$	Ahd I	5'...GACcc n^gcGTC... 3' 3'...CTGgg^ncgCAG... 5'
$B^* + C \rightarrow B^*C$	Ligase	5'...ccanngcn-gcgtc... 3' 3'...ggtnncg-ncgcag... 5'
$B^*C \rightarrow B + C^*$	Fnu4H I	5'...ccannGC^n GCgtc... 3' 3'...ggtnnCG n^CGcag... 5'
$A^* + D \rightarrow A^*D$	Ligase	5'...gaccn-ggnntgg... 3' 3'...ctggg-nccnacc... 5'
$A^*D \rightarrow A + D^*$	ScrF I	5'...gacCC^n GGnntgg... 3' 3'...ctgGG n^CCnacc... 5'
$B + D^* \rightarrow B^*D$	Ligase	5'...ccanngc-nggnntgg... 3' 3'...ggtnncgn-ccnacc... 5'
$B^*D \rightarrow B^* + D$	Xcm I	5'...CCAnngc n^ggnTGG... 3' 3'...GGTnncg^ncnnACC... 5'

between B^* and C is ligated to form B^*C (Step 2a in Figure 2 (b): $B^* + C \rightarrow B^*C$). This ligation product is subsequently cut into B and C^* by endonuclease $E2$ (Step 2b in Figure 2 (b): $B^*C \rightarrow B + C^*$). Now front foot C^* possesses sticky end u , and hence it will rove forward and hybridize with anchorage A down the track (Step 2c in Figure 2 (b)). This completes a full induction cycle for the front foot.

Note that reaction $A + C^* \rightarrow A^*C$ is irreversible: no restriction enzyme is present to cut A^*C into A and C^* . This effectively establishes the irreversibility of the motion of foot C . However, we note that after A^*C is cut into A^* and C , the two can be religated into A^*C (which is subsequently cut back into A^* and C). This represents an idling step in the motion of the walker. Similar analysis applies to the reaction $B^* + C \rightarrow B^*C \rightarrow B + C^*$.

The motion of hind foot D is similar to motion of front foot C and we omit its description for brevity. —

Molecular implementation using real enzymes. We give two implementations with real enzymes. The first one is a direct mapping of the implementation using the conceptual enzymes in Figure 3. The real enzymes used are shown in Figure 4 (a). Here, real endonucleases *AhdI*, *Fnu4HI*, *ScrFI*, and *XcmI* correspond to conceptual endonucleases $E1$, $E2$, $E3$, and $E4$, respectively. The reactions are shown in Table 1.

The second implementation reduces the number of endonucleases to three by using a non-palindromic endonuclease (*Aci I*) and its slightly more involved construction is shown in Table 2. The real enzymes used are shown in Figure 4 (b). Note that *Aci I* and *Hha I* correspond to the conceptual enzyme depicted in Figure 1 (b).

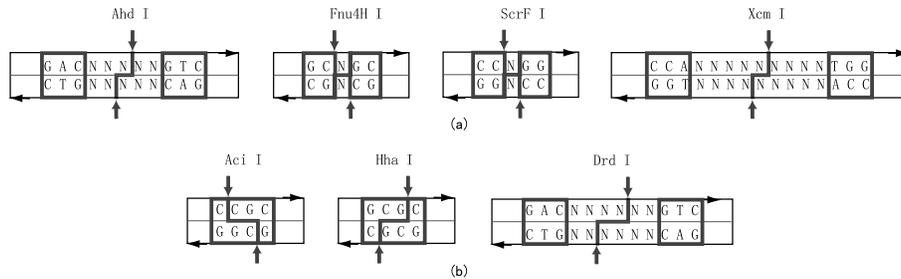


Fig. 4. Real enzymes used in the construction of device I. Endonuclease recognition sites and cleavage sites are indicated with bold boxes and pairs of bold arrows, respectively. N indicates the position of a base whose value does not affect recognition by an endonuclease

Table 2. Implementation of device I with endonucleases Aci I, Hha I, and Drd I. Ligation sites and cleavage sites are denoted with – and ^, respectively. The bases that determine recognition sites in action are in upper case

Reactions	Enzymes	DNA Sequences
$A + C^* \rightarrow A^*C$	Ligase	5'...gacnccg-c... 3' 3'...ctgng-gcg... 5'
$A^*C \rightarrow A^* + C$	Aci I	5'...gacn C [^] CG C... 3' 3'...ctgn G GC [^] G... 5'
$B^* + C \rightarrow B^*C$	Ligase	5'...c-cgc... 3' 3'...cgc-g... 5'
$B^*C \rightarrow B + C^*$	Hha I	5'... G CG [^] C... 3' 3'... C [^] GC G... 5'
$A^* + D \rightarrow A^*D$	Ligase	5'...gacnc-cgngtc... 3' 3'...ctgnggc-cncag... 5'
$A^*D \rightarrow A + D^*$	Drd I	5'...G A Cnc cg [^] gn G T C ... 3' 3'...CT G ng [^] gc cn C A G ... 5'
$B + D^* \rightarrow B^*D$	Ligase	5'...gcg-gngtc... 3' 3'...c-gccncag... 5'
$B^*D \rightarrow B^* + D$	Aci I	5'... G [^] CG Gngtc... 3' 3'... C GC [^] Cncag... 5'

Processivity of device I. A key technical issue in the construction of device I is to ensure that the walker is constrained to stay on or near the track. An isolated foot *C* or *D* would easily fall off the track and diffuse away. However, we can reduce the falling-off probability by constructing a multi-footed walker. Instead of possessing only two feet as in Figure 2, the walker has an array of alternate *C* and *D* feet. The feet are attached to a common backbone. Hence the walker is held to the track by multiple bonds - even if none are ligated (so all bonds are weak 1- or 2-base hydrogen bonds) then the probability of detachment is small. This is precisely what is needed - feet are held in the right place with the right amount of freedom to move - it introduces the

constraint that no foot can move more than two anchorages forward until all feet have moved at least one anchorage. =====

4 Design II

Overview. A potential problem of device I is that it may fall off the track. Though a walker with more feet has lower probability of falling off, we can not completely eliminate such risk. In contrast, the device we describe next is guaranteed to stay on the track, though it has a more complicated (hence less practical) construction and assumes a restriction enzyme property that has not yet been fully-substantiated. In device II, a two-footed walker steps over the anchorages along a track unidirectionally. The design of device II is based on the following principle: the lifting of one foot off the track is conditional on the attachment (ligation) of the other foot to the track. This attachment principle can ensure that at any moment, at least one foot of the walker is attached to the track. We describe the structure and step by step operation of device II below.

The track and the walker are depicted in Figure 5 (a). The track contains a linear array of anchorages. Each anchorage is a duplex DNA fragment with single strand DNA overhangs at both ends and its midpoint is tethered to the backbone of the track via single strand DNA. Thus the anchorage can be viewed as a two-ended dangler. In addition, between every two neighboring anchorages is tethered another dangler, referred to as a *switch*. As we shall see below, the alternating arrangement of anchorages and switches are used to construct a signaling mechanism which ensures the unidirectional and non-falling-off-track motion of the walker. The anchorages and switches are denoted as T_i and S_i respectively, where $i = 1, 2, 3, \dots, n$. A switch S_i can only be ligated with its immediate anchorage neighbors T_{i-1} and T_i . The upper ends of T are of type C^* , and the lower end of T_i is of type A^* (B^*) for odd (even) i .

The walker consists of two dangles connected with a single strand DNA. The two dangles serve as the feet of the walker and are denoted as F_1 and F_2 . The ends of both F_1 and F_2 are of type C . The walker stands on top of the upper ends of the anchorages and walks down the track unidirectionally, with the switch/anchorage complex of the road serving both as attaching points and as a signal transducing device to dictate the lifting and attaching of its feet in an alternating fashion such that it never falls off the track. In particular, at any point, if one foot is attached to anchorage T_i , the other foot can only be attached to T_i 's immediate neighbors, T_{i-1} and T_{i+1} .

The ends of the feet of the walker, of the anchorages, and of the switches have the following properties:

1. The complementary end pairs are: (A, A^*) , (A, B^*) , (B^*, A^*) , (B, B^*) , and (C, C^*) . Only two dangles with these complementary ends can be ligated.
2. The formation of CC^* ligation product at the upper end of the anchorage introduces a recognition site on the anchorage for endonuclease $E3$, which has a similar cleavage pattern as the one depicted in Figure 1 (c). And this results in a cleavage at the other end of the anchorage such that the anchorage is cut from the switch it is currently ligated with (if there is one). Similarly, the formation of A^*A (resp. B^*B) at the lower end of the anchorage will produce a recognition site on the anchorage for endonucle-

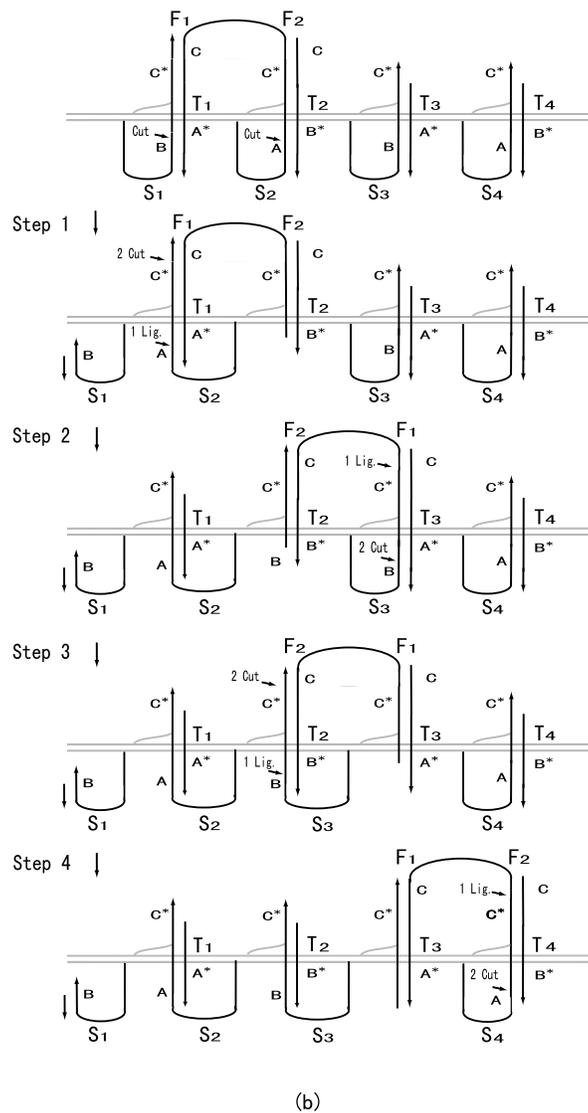
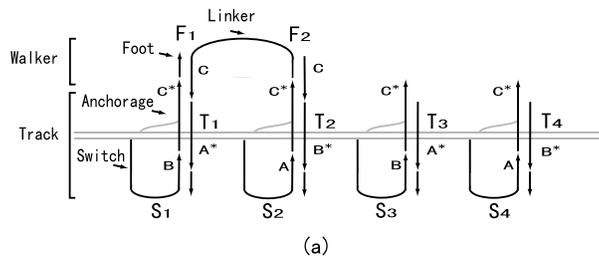


Fig. 5. The structural design and step by step operation of device II. (a) Structural design of the device. (b) Step by step operation of the device. The track is depicted in light gray

ase $E1$ (resp. $E2$) and this will result in the cleavage of CC^* at the upper end of the anchorage if there is a foot end C ligated with C^* .

We will next see how these properties guarantee the desired motion of the walker as we go through a step by step description of the walker's motion.

Step by step motion. Now we describe the four steps of the walker's motion that completes a full inductual cycle. Initially, the walker and track complex is assembled in such a way that the feet F_1 and F_2 of the walker are ligated with anchorages T_1 and T_2 , respectively; each switch S_i is ligated to the lower end of T_i , forming BA^* for odd i and AB^* for even i . Note that BA^* and AB^* are different.

Step 0. Upon introduction of enzymes into the system, switches S_1 and S_2 are cut from anchorages T_1 and T_2 respectively, since the CC^* sequences at the upper ends of T_1 and T_2 constitute endonuclease $E3$ recognition sites and hence result in cleavages at the lower ends of T_1 and T_2 . Now S_2 (with end A) can explore its neighboring space and be ligated with either T_1 (with end A^*) or T_2 (with end B^*). Ligation between S_2 and T_2 is a just an idling step: the ligation product will be subsequently cut again. In contrast, ligation of S_2 and T_1 brings the system to Step 1.

Step 1. The ligation of S_2 (with end A) and T_1 (with end A^*) introduces a recognition site for $E1$, and thus results in the cleavage of F_1 from the upper end of T_1 . Note that the ligation product between the lower end of T_1 and S_2 contains recognition site (AA^*) for $E1$ while the ligation product between foot F_1 and the upper end of T_1 contains recognition site (CC^*) for endonuclease $E3$. As such, both $E1$ and $E3$ will compete to perform cleavage on the common ligation product. (See Figure 6 (a) for detail.) It is possible that endonuclease $E3$ cuts switch S_2 away from anchorage T_1 , resulting in an idling step. However, there must also be non-zero probability that endonuclease $E1$ cuts foot F_1 away from anchorage T_1 , advancing the system to Step 2.

Step 2. Now foot F_1 has free end C and can swing around the ligation product between foot F_2 and anchorage T_2 and get ligated with the upper end C^* of anchorage T_3 . Note that now foot F_1 is in front of foot F_2 . The ligation of CC^* subsequently results in the cleavage of S_3 from T_3 .

Step 3. Switch S_3 has free end B and is ligated with the B^* end of anchorage T_2 , and the newly formed recognition site BB^* leads to the action of endonuclease $E2$ and results in the cleavage between foot F_2 and anchorage T_2 .

Step 4. Foot F_2 swings to in front of foot F_1 and is ligated with anchorage T_4 , resulting in the cleavage of switch S_4 from the lower end of anchorage T_4 .

Upon completion of Step 4, the walker has moved from anchorages T_1 and T_2 to anchorages T_3 and T_4 . This finishes a full inductual cycle, and hence the walker can continue moving down the track.

Correctness. To show the correctness of the design, we prove the following three properties of the walker: 1) the motion of the walker is unidirectional; 2) the walker never falls off the track; 3) the motion of the walker is never blocked. We omit detail here due to lack of space, but refer the reader to [19].

Implementation with conceptual enzymes. Figure 6 describes the implementation of

device II with these conceptual restriction enzymes $E1$, $E2$, and $E3$, which have similar cutting patterns as the one shown in Figure 1 (c). We require that $d_1 = d_2 = d_3$ and $e_1 = e_2 = e_3$, where d_i and e_i are the length parameters for E_i ($i = 1, 2$, and 3). In Figure 6 (a), two anti-parallel flows of reactions are depicted. Starting from the top, end A (of a switch) has sticky end sequence complementary to end A^* (lower end of an anchorage) and hence the two are ligated together. This creates a recognition site for endonuclease $E1$, and results in the cleavage of end C (of a foot) from end C^* (upper end of an anchorage). This downward flow of reactions can be fully reversed into the anti-parallel upward flow starting from the bottom with C^* and C and ends at the top with A and A^* . We note that due to the fully reversible nature of reactions, the reaction system has nonzero probability to explore all three states: the top one ($A, A^* :: C^*C$), the middle one ($AA^* :: C^*C$), and the bottom one ($AA^* :: C^*, C$), where $::$ represents the duplex portion of DNA connecting the two ends. Similar fully reversible anti-parallel flows of reactions involving $E2$ and $E3$ are depicted in Figure 6 (b). In contrast, reactions in Figure 6 (c) and 6 (d) are not fully reversible since neither ligation of AB^* nor that of BA^* can result in a recognition site for an endonuclease, and hence CC^* can not be cleaved. This irreversibility ultimately accounts for the unidirectionality of the motion of the walker. The downward reaction flow in Figure 6 (a), the upward reaction flow in (d), the downward reaction flow in (b), and the upward reaction flow in (c) correspond to Steps 1, 2, 3, and 4 in Figure ??, respectively.

Molecular implementation with real enzymes. The above conceptual enzymes can be mapped directly to real enzymes in Figure 7, where conceptual enzymes $E1$, $E2$, and $E3$ correspond to real enzymes Bpm I, Bsg I, and BpuE I, respectively. Table 3 describes the implementation with these real enzymes. Note that we have the following mapping from sequences in Figure 6 to the sequences in Table 3: $1 = C$, $u = TG$, $1^\diamond = GAG$, $2 = G$, $2^\diamond = CAG$, $\bar{3}^{\diamond R} = CTC$, $\bar{v}^R = AA$, and $\bar{3}^R = G$.

Practicality. One assumption we make about the enzyme is that the presence of a single strand between the recognition site and cleavage site of each endonuclease used above will neither alter the specificity nor totally inhibit the activity of that endonuclease. A theoretical modeling of the molecular structure of the enzyme and its interaction with the DNA strands would shed light on the practicality of this assumption. However, the final validation of this assumption relies on a rigorous experimental study. Though our preliminary experimental result is in agreement with this assumption, more work is still required to further substantiate this assumption.

5 Discussion

We have depicted the backbones of the walking devices as duplex DNA fragments for simplicity. However, this is not technically precise. One property we require of the backbone of a track is its rigidity, to ensure that the walker cannot skip anchorage(s) and “jump” ahead. Existing DNA lattices provide such a platform [5, 8, 16, 17]. We can easily embed the anchorages to a rigid DNA lattice and thus integrate a walking device to a lattice, with the latter provide the desired rigid backbone for the anchorages. In ad-

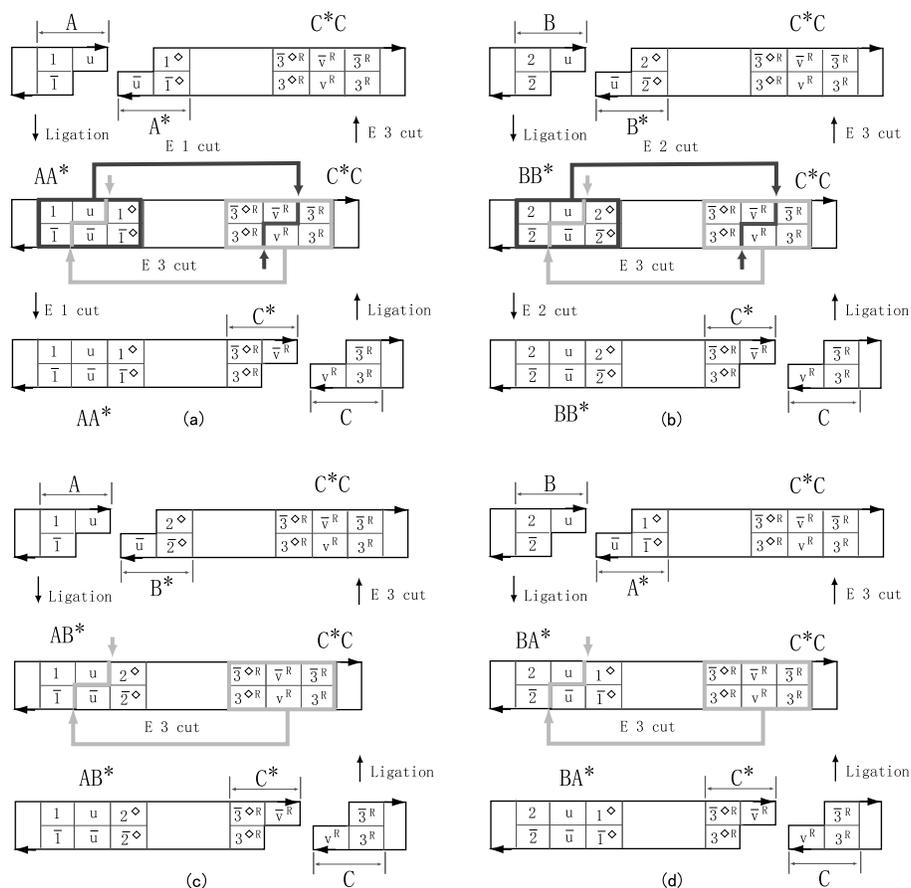


Fig. 6. Actions of conceptual enzymes used in the construction of device II. (a) Sequences 1, u , 1° , $\bar{1}$, \bar{u} , and $\bar{1}^\circ$ (sequences of AA^*) together constitute the recognition site (red box) for conceptual endonuclease $E1$, whose cleavage site is indicated with a pair of bold arrows. Sequences $\bar{3}^\circ R$, \bar{v}^R , $\bar{3}^R$, $3^\circ R$, v^R , and 3^R (sequences of C^*C) together constitute the recognition site (light gray box) for conceptual endonuclease $E3$, whose cleavage site is indicated with a pair of light gray arrows. (b) Two anti-parallel flows of reactions by $E2$ and $E3$. (c) and (d) Neither ligation of AB^* or BA^* results in cleavage of CC^* .

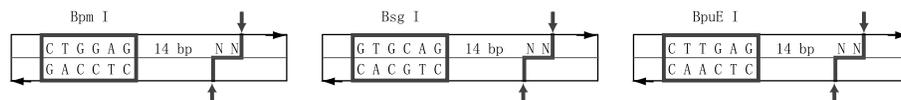


Fig. 7. Real enzymes used in the construction of device II. Endonuclease recognition sites and cleavage sites are indicated with bold boxes and pairs of bold arrows, respectively. N indicates the position of a base whose value does not affect recognition by an endonuclease

Table 3. Implementation of device II with endonucleases Bpm I, Bsg I, and BpuE I. Ligation sites and cleavage sites are denoted with – and ^, respectively. The bases that determine recognition sites in action are in upper case

Reactions	Enzymes	Sequences
$A + A^* :: C^*C \rightarrow$ $AA^* :: C^*C$	Ligase	5'...ctg-gag(n) ₁₁ ctcaag...3' 3'...g-acctc(n) ₁₁ gagtc...3'
$AA^* :: C^*C \rightarrow$ $AA^* :: C^* + C$	Bpm I	5'...CTGGAG(n) ₁₁ ctc aa^g...3' 3'...GACCTC(n) ₁₁ gag^tt c...3'
$AA^* :: C^* + C \rightarrow$ $AA^* :: C^*C$	Ligase	5'...ctggag(n) ₁₁ ctcaa-g...3' 3'...gacctc(n) ₁₁ gag-ttc...3'
$AA^* :: C^*C \rightarrow$ $A + A^* :: C^*C$	BpuE I	5'...c tg^gag(n) ₁₁ CTCAAG...3' 3'...g^ac ctc(n) ₁₁ GAGTTC...3'
$B + B^* :: C^*C \rightarrow$ $BB^* :: C^*C$	Ligase	5'...gtg-cag(n) ₁₁ ctcaag...3' 3'...c-acgtc(n) ₁₁ gagtc...3'
$BB^* :: C^*C \rightarrow$ $BB^* :: C^* + C$	Bsg I	5'...GTGCAG(n) ₁₁ ctc aa^g...3' 3'...CACGTC(n) ₁₁ gag^tt c...3'
$BB^* :: C^* + C \rightarrow$ $BB^* :: C^*C$	Ligase	5'...gtgcag(n) ₁₁ ctcaa-g...3' 3'...cacgtc(n) ₁₁ gag-ttc...3'
$BB^* :: C^*C \rightarrow$ $B + B^* :: C^*C$	BpuE I	5'...g tg^cag(n) ₁₁ CTCAAG...3' 3'...c^ac gtc(n) ₁₁ GAGTTC...3'
$AB^* :: C^* + C \rightarrow$ $AB^* :: C^*C$	Ligase	5'...ctgcag(n) ₁₁ ctcaa-g...3' 3'...gacgtc(n) ₁₁ gag-ttc...3'
$AB^* :: C^*C \rightarrow$ $A + B^* :: C^*C$	BpuE I	5'...c tg^cag(n) ₁₁ CTCAAG...3' 3'...g^ac gtc(n) ₁₁ GAGTTC...3'
$BA^* :: C^* + C \rightarrow$ $BA^* :: C^*C$	Ligase	5'...gtggag(n) ₁₁ ctcaa-g...3' 3'...cacctc(n) ₁₁ gag-ttc...3'
$BA^* :: C^*C \rightarrow$ $B + A^* :: C^*C$	BpuE I	5'...g tg^gag(n) ₁₁ CTCAAG...3' 3'...c^ac ctc(n) ₁₁ GAGTTC...3'

dition to the rigidity of the track, the structure and the size of the walker are also crucial factors to ensure that a foot can only explore the immediately neighboring anchorages. In device I, though it is hard to ensure this property for a two-footed walker (since in such a walker one foot might swing around the other foot in a similar fashion as in device II), this property can be rather straightforwardly guaranteed in a multi-footed walker with a rigid body. In device two, the two feet of the walker alternate their order along the track by swinging around each other and we hence only need to properly design the size of the body such that a foot can only reach a neighboring anchorage.

The designs of the devices assume that enzyme cleavage occurs only after the DNA strands are ligated. This assumption is in agreement with the experimental results observed in our recent construction of a unidirectional autonomous DNA walker [20]. In this device, we use two class II enzymes PflM I and BstAP I and the system operates at 37 °C. However, we note that this property does not hold true for all class II enzymes under all conditions. Indeed, Shapiro's group has observed that a class II enzyme Fok I can cleave GC rich DNA duplex strands with nicks present between Fok I recognition site and cleavage site under at low temperature (8 °C) [2].

Acknowledgement

The authors would like to thank Hao Yan and Thomas H. LaBean for helpful discussions. We are also grateful to the helpful comments from Alexander J. Hartemink, Yusu Wang, Hai Yu, and the anonymous reviewers. This work was supported by NSF under ITR Grant EIA-0086015 and ITR Grant 0326157, by NSF under QuBIC Grant EIA-0218376 and QuBIC Grant EIA-0218359, by NSF under EMT Grant CCF-0432038 and EMT Grant CCF-0432047, by DARPA/AFSOR under Contract F30602-01-2-0561, and by RGC under Grant HKBU2107/04E.

References

1. P. Alberti and J. L. Mergny. DNA duplex-quadruplex exchange as the basis for a nanomolecular machine. *Proc. Natl. Acad. Sci. USA*, 100:1569–1573, 2003.
2. Y. Benenson, R. Adar, T. Paz-Elizur, Z. Livneh, and E. Shapiro. DNA molecule provides a computing machine with both data and fuel. *Proc. Natl. Acad. Sci. USA*, 100:2191–2196, 2003.
3. Y. Chen, M. Wang, and C. Mao. An autonomous DNA nanomotor powered by a DNA enzyme. *Angew. Chem. Int. Ed.*, 43:3554–3557, 2004.
4. L. Feng, S. H. Park, J. H. Reif, and H. Yan. A two-state DNA lattice switched by DNA nanoactuator. *Angew. Chem. Int. Ed.*, 42:4342–4346, 2003.
5. T. H. LaBean, H. Yan, J. Kopatsch, F. Liu, E. Winfree, J. H. Reif, and N. C. Seeman. The construction, analysis, ligation and self-assembly of DNA triple crossover complexes. *J. Am. Chem. Soc.*, 122:1848–1860, 2000.
6. J. Li and W. Tan. A single DNA molecule nanomotor. *Nano Lett.*, 2:315–318, 2002.
7. D. Liu and S. Balasubramanian. A proton fuelled DNA nanomachine. *Angew. Chem. Int. Ed.*, 42:5734–5736, 2003.
8. C. Mao, W. Sun, and N. C. Seeman. Designed two-dimensional DNA holliday junction arrays visualized by atomic force microscopy. *J. Am. Chem. Soc.*, 121:5437–5443, 1999.
9. C. Mao, W. Sun, Z. Shen, and N. C. Seeman. A DNA nanomechanical device based on the B-Z transition. *Nature*, 397:144–146, 1999.
10. J. H. Reif. The design of autonomous DNA nanomechanical devices: Walking and rolling DNA. *Lecture Notes in Computer Science*, 2568:22–37, 2003. Published in Natural Computing, DNA8 special issue, Vol. 2, p 439-461, (2003).
11. N. C. Seeman. DNA in a material world. *Nature*, 421:427–431, 2003.
12. W. B. Sherman and N. C. Seeman. A precisely controlled DNA biped walking device. *Nano Lett.*, 4:1203–1207, 2004.
13. F. C. Simmel and B. Yurke. Using DNA to construct and power a nanoactuator. *Phys. Rev. E*, 63:041913, 2001.
14. F. C. Simmel and B. Yurke. A DNA-based molecular device switchable between three distinct mechanical states. *Appl. Phys. Lett.*, 80:883–885, 2002.
15. A. J. Turberfield, J. C. Mitchell, B. Yurke, Jr. A. P. Mills, M. I. Blakey, and F. C. Simmel. DNA fuel for free-running nanomachines. *Phys. Rev. Lett.*, 90:118102, 2003.
16. E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394:539–544, 1998.
17. H. Yan, S. H. Park, G. Finkelstein, J. H. Reif, and T. H. LaBean. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science*, 301:1882–1884, 2003.

18. H. Yan, X. Zhang, Z. Shen, and N. C. Seeman. A robust DNA mechanical device controlled by hybridization topology. *Nature*, 415:62–65, 2002.
19. P. Yin, A. J. Turberfield, and J. H. Reif. Designs of autonomous unidirectional walking DNA devices. Technical Report CS-2004-01, Duke University, Computer Science Department, 2004.
20. P. Yin, H. Yan, X. G. Daniell, A. J. Turberfield, and J. H. Reif. A unidirectional DNA walker moving autonomously along a linear track. *Angew. Chem. Int. Ed.*, 2004. In press.
21. B. Yurke, A. J. Turberfield, Jr. A. P. Mills, F. C. Simmel, and J. L. Neumann. A DNA-fuelled molecular machine made of DNA. *Nature*, 406:605–608, 2000.