

# Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker

Yu He and David R. Liu\*

**Multistep synthesis in the laboratory typically requires numerous reaction vessels, each containing a different set of reactants. In contrast, cells are capable of performing highly efficient and selective multistep biosynthesis under mild conditions with all reactants simultaneously present in solution<sup>1–4</sup>. If the latter approach could be applied in the laboratory, it could improve the ease, speed and efficiency of multistep reaction sequences. Here, we show that a DNA mechanical device—a DNA walker moving along a DNA track—can be used to perform a series of amine acylation reactions in a single solution without any external intervention. The products of these reactions are programmed by the sequence of the DNA track, but they are not related to the structure of DNA. Moreover, they are formed with speeds and overall yields that are significantly greater than those previously achieved by multistep DNA-templated small-molecule synthesis.**

The exquisite selectivity and efficiency of many biosynthetic pathways rely on the use of protein or nucleic acid templates to modulate the effective molarities of substrates. For combinatorial biosyntheses, including mRNA-templated peptide synthesis during ribosomal translation<sup>1,2</sup> as well as pathways that generate polyketides and non-ribosomal peptides<sup>3,4</sup>, this strategy enables the sequence of building blocks in a biosynthetic product to be determined by the sequence of mRNA nucleotides or enzyme modules in the template, enabling different multistep reaction products to be generated selectively from a single set of substrates with no external intervention.

In the absence of enzymes, laboratory reaction sequences to generate multistep synthetic products generally proceed in multiple reaction vessels, such that each vessel exposes a different set of substrates to a different set of reaction conditions. In principle, performing multistep reaction sequences in a single solution programmed by a template could significantly increase the ease, speed and overall efficiency of multistep syntheses. We previously engineered a DNA template that undergoes a series of secondary structure changes when subjected to a schedule of increasing temperatures<sup>5</sup>. These changes expose hybridization sites for DNA-templated synthesis, an approach to controlling chemical reactivity in which reactions between DNA-linked reagents are triggered by DNA hybridization, enabling a multistep reaction sequence to take place in one solution. Although successful, the generality of this approach is significantly limited by the requirement for a different arrangement of substrates for each step and also large changes in temperature during the reaction sequence (for example, 4 to 72 °C)<sup>5</sup>.

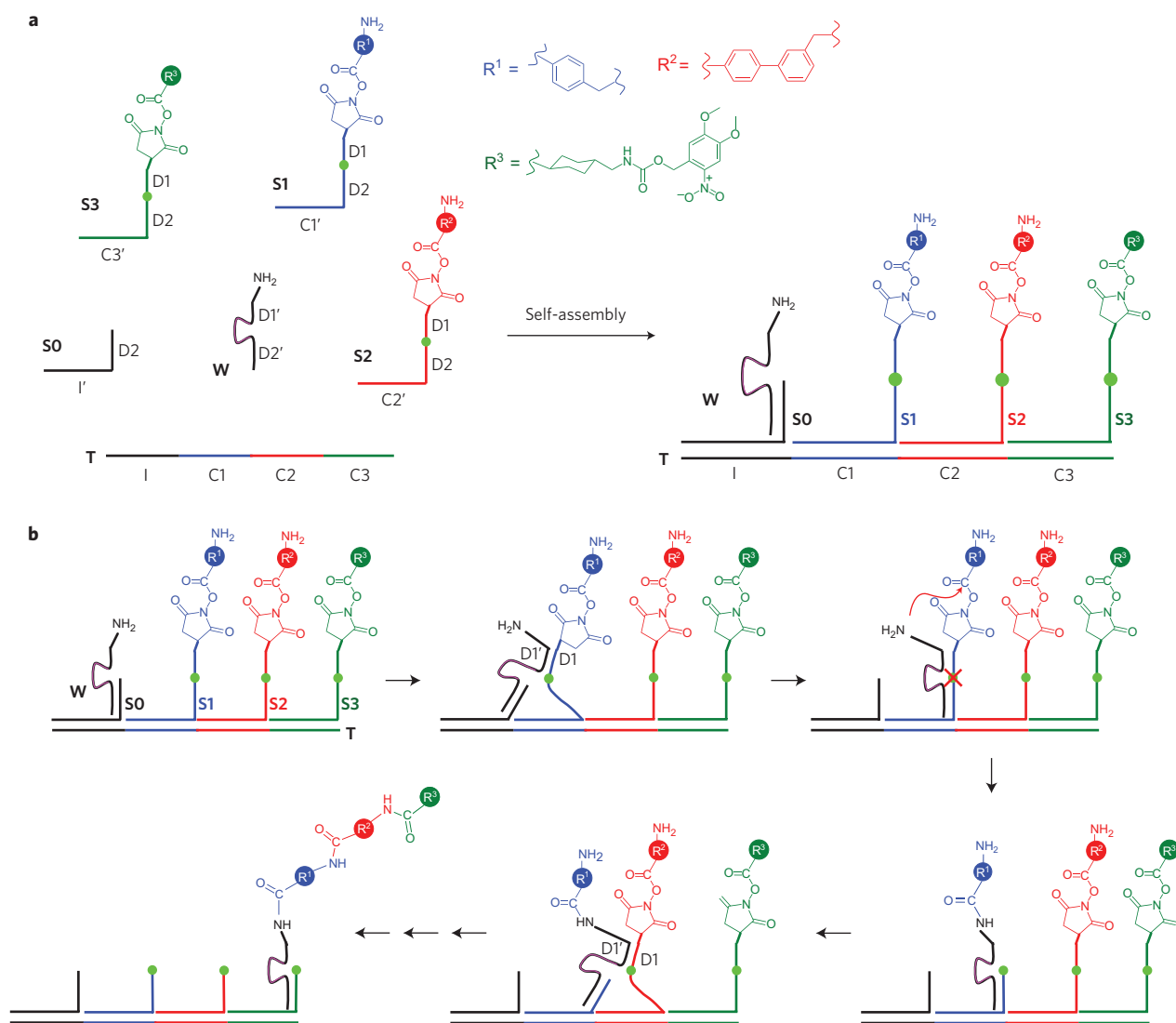
Recent advances in engineering of DNA-based devices<sup>6–9</sup> have resulted in nanoscale DNA machines that are capable of

autonomously changing their physical location over time<sup>10–16</sup>. For example, Mao and colleagues recently reported a DNA ‘nanowalker’ that moves unidirectionally along a DNA track from station to station<sup>12</sup>. We envisioned that the ability of a DNA walker to translocate could be integrated with DNA-templated synthesis to enable a specific chemical reaction to take place upon arrival of the walker at each station. Because the reaction product remains linked to the walker and serves as the starting material for the next reaction, this strategy results in a progressively more advanced synthetic reaction product as the walker moves along its track. Here, we report the development of an autonomous DNA walker that performs a series of amine acylation reactions as it moves from station to station along a DNA track. The resulting machine mediates multistep synthesis of an oligoamide in a single isothermal solution programmed by the sequence of a DNA track, conceptually resembling an artificial ribosome. The system can generate multistep products in overall yields that are much higher than those of previous DNA-templated small-molecule syntheses<sup>5,17</sup>, and in total reaction times of only a few hours. This work also represents one of the first DNA devices that manipulates the covalent structure of non-nucleic acid molecules<sup>18–20</sup>, and thus expands the functional scope of DNA-based devices.

The DNA-based ribosome mimetic (‘DNAsome’) system (Fig. 1a) consists of four sets of components. The first comprises track (T), analogous to an mRNA, which is a long single-stranded DNA molecule that contains an initiation site (I) followed by codon sites (C1, C2, C3), each of which encodes a different substrate (S1, S2, S3). The substrates form the next part, and are analogous to aminoacylated tRNAs, each contain (from 5′ to 3′) a synthetic small-molecule building block, a docking region (D1) of constant sequence, two consecutive central RNA nucleotides, a second constant docking region (D2) and a unique 21-nucleotide anticodon complementary to a template codon C1, C2 or C3. For the system described here, the small-molecule reagents are amino acid *N*-hydroxysuccinimidyl (NHS) esters that cannot undergo intramolecular cyclization due to the enforced separation of amine and ester groups. The third component is the initiator (S0), which is a truncated substrate that contains only D2 followed by an anticodon complementary to site I of the template. Finally there is the DNA walker (W), which is a DNA 38-mer containing a 5′ region complementary to D2 (D2′), the 10–23 RNA-cleaving DNase evolved by Joyce and colleagues<sup>21</sup>, a region complementary to D1 (D1′) and a 3′ primary amine group. The interval between any two adjacent stations corresponds to two turns of B-form DNA, which defines each walking step as ~7 nm.

When all components except the walker and initiator are combined in one solution, the substrates and templates self-assemble

Howard Hughes Medical Institute, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, USA. \*e-mail: drliu@fas.harvard.edu



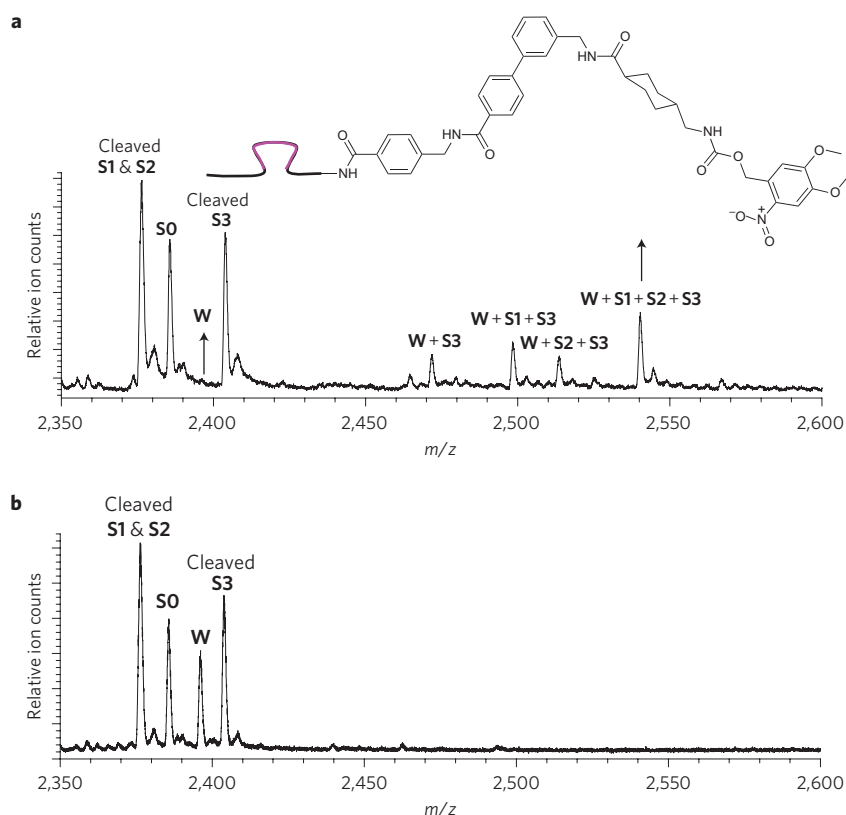
**Figure 1 | Overview of the DNAsome system.** **a**, The system described in this work comprises six DNA or DNA-linked molecules. Three substrates (**S1–S3**) and an initiator (**S0**) can hybridize on a single-stranded DNA track (**T**). Each substrate has an amino acid NHS ester at its 5' end and two ribonucleotides (green dot) in the middle of its DNA sequence. The DNA walker (**W**) contains a 3' amine group and an RNA-cleaving DNAzyme (purple line) that can cleave the ribonucleotides in the substrates. **b**, DNAsome-mediated multistep synthesis of a triamide product. All steps take place in a single solution under one set of reaction conditions without external intervention.

into a double-stranded complex. Upon addition of **S0** and **W**, the initiator and walker associate via D2:D2' hybridization and localize to the track–substrate complex through **S0**:I hybridization (Fig. 1a). The template-bound walker then hybridizes with the nearest D1 docking region, which is found on the **S1** substrate annealed to the first station (Fig. 1b). Hybridization between **W** and **S1** induces the favourable translocation of **W** from **S0** (which can hybridize only to the D2' region of **W**) to **S1** (which can hybridize to both the D1' and D2' regions of **W**). **W–S1** hybridization also triggers DNA-templated acylation of the walker's amine group with the NHS ester of **S1**, resulting in transfer of the first amino-acid building block from **S1** to **W**. The 10–23 DNAzyme core in the walker cleaves the ribonucleotide linkage in **S1**, allowing the 5' fragment of **S1** to dissociate (Fig. 1b). The system is now identical to the **W:S0:T** starting state, except that **W** has captured the first reaction product and has translocated from the initiation site to the first station.

Two subsequent cycles of walker translocation, DNA-templated amine acylation, DNAzyme-catalysed cleavage, and dissociation of the 5' fragment of the expended substrate result in the walker

resting at the final station on the template covalently linked to the final multistep reaction product (in this case, a triamide containing three amino-acid building blocks in a specific **T**-programmed order). Because each step of this cycle occurs spontaneously under identical conditions, the entire three-step reaction sequence proceeds autonomously, requiring no changes in temperature or pH, and no intervention from the researcher.

A major challenge of any effort to perform a multistep reaction sequence in a single solution is to prevent reactive substrates from undergoing any of the many possible reactions other than those on the desired pathway. In the case of transforming amino-acid building blocks into a specific oligoamide, each activated amino acid can only react with the nascent product, and not with each other. In nature, the ribosomal machinery modulates the effective molarity of aminoacylated tRNAs with respect to water or other nucleophiles to greatly increase the likelihood that an aminoacylated tRNA in the A-site of the ribosome couples with the nascent peptide in the P-site before hydrolysis or errant coupling takes place<sup>22,23</sup>. Like the ribosome, the DNAsome system displays multiple activated amino acids on a single track simultaneously. These amino



**Figure 2 | Analysis of reaction products generated by the DNAsome system. a**, Mass spectrometry analysis of the three-step DNAsome-mediated reaction sequence. See main text for a detailed description of reaction conditions. **b**, Mass spectrometry analysis of the experiment in **a** repeated using substrates lacking amino-acid NHS esters. See Supplementary Information for all expected and observed masses.

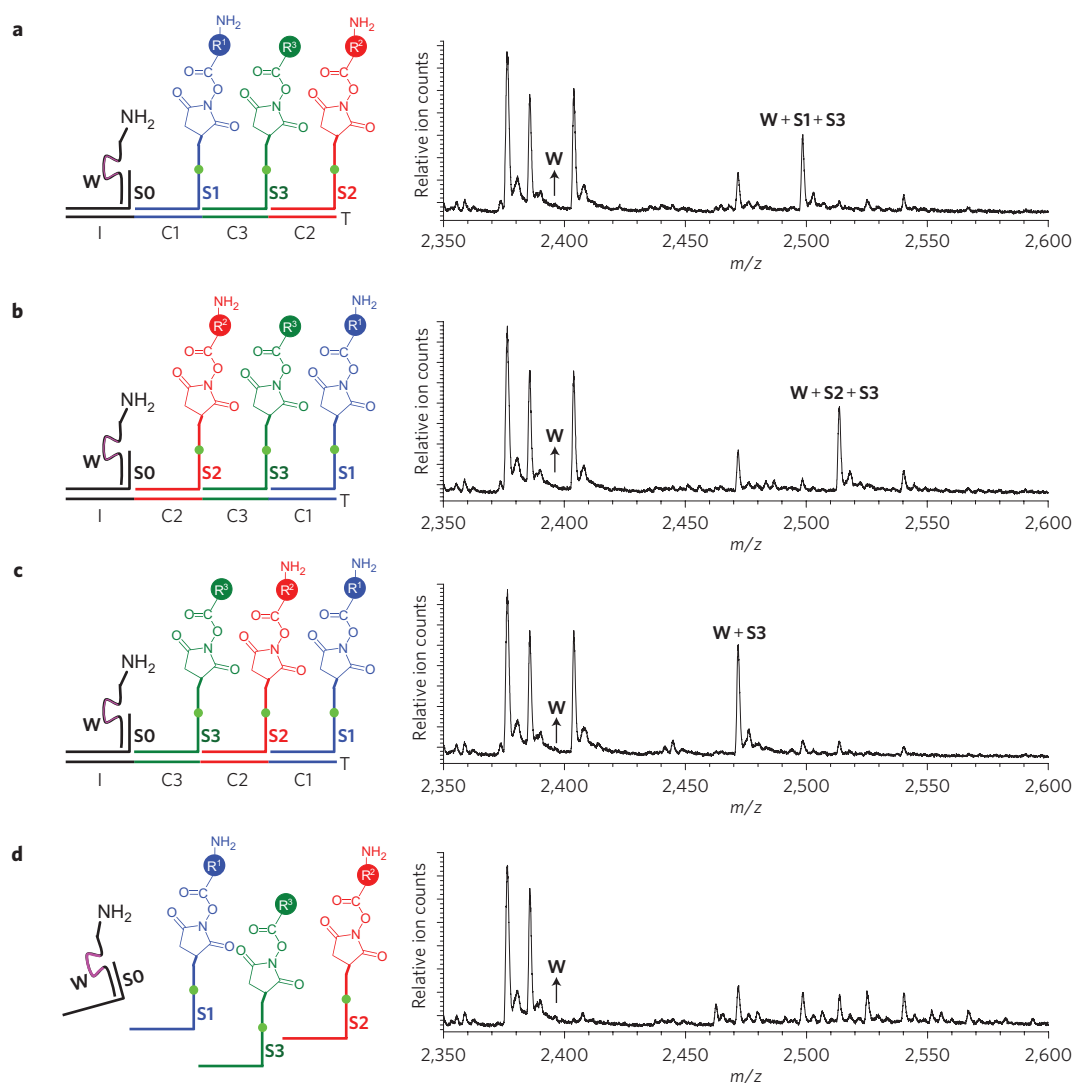
acids each contain a free amine group that can potentially react with the aminoacyl NHS ester on an adjacent substrate station, leading to misordered products or precluding the formation of the desired final product. We hypothesized, based on our previous observations<sup>5,24,25</sup>, that the double-stranded nature of the intervening track between adjacent substrates would effectively separate adjacent amino acids, decreasing their effective molarity and preventing side reactions. To test this hypothesis, we performed a model reaction in which a substrate containing a primary amine (**S3a**) and an NHS ester-linked substrate (**S2a**) were hybridized at adjacent stations of the same track (Supplementary Fig. S3). After incubation in 50 mM 3-(*N*-morpholino)propanesulphonic acid (MOPS) buffer, pH 7.5, with 10 mM Mg(OAc)<sub>2</sub> for 16 h, no significant product formation was observed, indicating that adjacently docked substrates are not prone to uncontrolled cross-reaction in the absence of the walker (Supplementary Fig. S3).

In a typical multistep synthesis experiment, **T**, **S1**, **S2** and **S3** (1.6 μM each) were combined at room temperature in aqueous buffer containing 5 mM MOPS and 10 mM Mg(OAc)<sub>2</sub> at pH 5.0. In a separate tube, **W** and **S0** (3.0 μM each) were combined in the same buffer. After 1 h at room temperature, the two solutions were combined in 50 mM aqueous sodium MOPS buffer at pH 7.5 (1.0 μM final concentration of each of the six molecules) to initiate the walking and multistep synthesis process. After 8 h at room temperature, the crude reaction was desalted by gel filtration, then analysed by high-resolution mass spectrometry (Fig. 2 and Supplementary Fig. S1). Following the conditions described above, **W** was completely consumed, and the three-step desired triamide arising from reaction of **W** with **S1**, **S2** and **S3** represented the major product (Fig. 2). We estimated the overall yield of the desired three-step product to be ~45% of detectable DNA-linked species by comparison with mass spectra of authentic product

standards mixed in various ratios (Supplementary Fig. S2). This overall yield is the highest of any three-step DNA-templated small-molecule synthesis reported to date<sup>5,17</sup>. Minor truncation products corresponding to the omission of a reaction (such as **W + S1 + S3**) were also observed, typically in ~10–25% yield each. These results indicate that the DNAsome system is capable of mediating an efficient three-step synthesis over several hours with no intervention.

To confirm that multistep synthesis mediated by the DNAsome proceeds in an ordered manner programmed by the nucleotide sequence of the track (in the above example, I–C1–C2–C3), we repeated the above experiment using three different DNA tracks, or omitting **T** altogether. Because **S3** lacks an amine group, it effectively serves as a translation terminator. When the experiment described above was repeated using an I–C1–C3–C2 track, **W** was again entirely consumed, and the predominant product observed by mass spectrometry was consistent with the reaction of **W + S1 + S3** (Fig. 3a). The absence of products containing **S2** is consistent with the ordered, sequence-dependent nature of DNAsome-mediated multistep synthesis. Similarly, using an I–C2–C3–C1 track resulted in the formation of a predominant product consistent with the reaction of **W + S2 + S3** (Fig. 3b), and using an I–C3–C2–C1 track yielded only **W + S3** product (Fig. 3c). Finally, omission of the track altogether resulted in the formation of a complex mixture of minor products consistent with the uncontrolled random intermolecular reaction of **W**, **S1**, **S2** and **S3** (Fig. 3d). Taken together, these results establish that artificial translation mediated by the DNAsome proceeds in a manner programmed by the order of codons on the DNA track.

In this work we integrated a DNA mechanical device and DNA-templated synthesis into a system that is capable of mediating, in a single solution, sequence-programmed, autonomous, multistep



**Figure 3 | Mass spectroscopy analysis of reactions identical to the one shown in Fig. 2a, but using different DNA tracks or with no DNA track. a–d,** The DNA tracks are I–C1–C3–C2 (a), I–C2–C3–C1 (b), I–C3–C2–C1 (c) and no DNA track (d). See Supplementary Information for all expected and observed masses.

ordered synthesis of oligoamides without natural enzymes. Because the device is powered by the favourable energetics of RNA cleavage, no changes in temperature, pH or other reaction conditions are needed over the course of the multistep synthesis. The system described here extends the capabilities of DNA-templated synthesis in significant ways. Although requiring the use of self-cleaving reagents that cannot react intramolecularly<sup>17</sup>, the multistep syntheses presented here are more efficient (~45% overall yield for three steps) and proceed more quickly (starting materials to three-step product within several hours) than existing methods for sequential multistep DNA-templated syntheses<sup>5,17</sup>. These results also raise the possibility of using a similar strategy to rapidly generate libraries of synthetic small molecules constructed through multiple consecutive DNA-templated synthesis steps in a single solution<sup>26,27</sup>. Finally, the system described here demonstrates how the ordered synthesis of products unrelated to nucleic acids and the avoidance of undesired side reactions can take place in a manner directed and catalysed entirely by nucleic acids, and therefore has conceptual relevance to early translation systems.

Received 19 April 2010; accepted 23 August 2010;  
published online 10 October 2010

## References

- Hanes, J. & Pluckthun, A. *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl Acad. Sci. USA* **94**, 4937–4942 (1997).
- Zahnd, C., Amstutz, P. & Pluckthun, A. Ribosome display: selecting and evolving proteins *in vitro* that specifically bind to a target. *Nature Methods* **4**, 269–279 (2007).
- Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805–1810 (2004).
- Sattely, E. S., Fischbach, M. A. & Walsh, C. T. Total biosynthesis: *in vitro* reconstitution of polyketide and nonribosomal peptide pathways. *Nat. Prod. Rep.* **25**, 757–793 (2008).
- Snyder, T. M. & Liu, D. R. Ordered multistep synthesis in a single solution directed by DNA templates. *Angew. Chem. Int. Ed.* **44**, 7379–7382 (2005).
- Seeman, N. C. From genes to machines: DNA nanomechanical devices. *Trends Biochem. Sci.* **30**, 119–125 (2005).
- Bath, J. & Turberfield, A. J. DNA nanomachines. *Nature Nanotech.* **2**, 275–284 (2007).
- Liedl, T., Sobey, T. L. & Simmel, F. C. DNA-based nanodevices. *Nano Today* **2**, 36–41 (2007).
- Beissenhirtz, M. K. & Willner, I. DNA-based machines. *Org. Biomol. Chem.* **4**, 3392–3401 (2006).
- Yin, P., Yan, H., Daniell, X. G., Turberfield, A. J. & Reif, J. H. A unidirectional DNA walker that moves autonomously along a track. *Angew. Chem. Int. Ed.* **43**, 4906–4911 (2004).
- Bath, J., Green, S. J. & Turberfield, A. J. A free-running DNA motor powered by a nicking enzyme. *Angew. Chem. Int. Ed.* **44**, 4358–4361 (2005).

12. Tian, Y., He, Y., Chen, Y., Yin, P. & Mao, C. D. Molecular devices—a DNAzyme that walks processively and autonomously along a one-dimensional track. *Angew. Chem. Int. Ed.* **44**, 4355–4358 (2005).
13. Yin, P., Choi, H. M. T., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451**, 318–322 (2008).
14. Omabegho, T., Sha, R. & Seeman, N. C. A bipedal DNA Brownian motor with coordinated legs. *Science* **324**, 67–71 (2009).
15. Gu, H. Z., Chao, J., Xiao, S. J. & Seeman, N. C. A proximity-based programmable DNA nanoscale assembly line. *Nature* **465**, 202–205 (2010).
16. Lund, K. *et al.* Molecular robots guided by prescriptive landscapes. *Nature* **465**, 206–210 (2010).
17. Gartner, Z. J., Kanan, M. W. & Liu, D. R. Multistep small-molecule synthesis programmed by DNA templates. *J. Am. Chem. Soc.* **124**, 10304–10306 (2002).
18. Chhabra, R., Sharma, J., Liu, Y. & Yan, H. Addressable molecular tweezers for DNA-templated coupling reactions. *Nano Lett.* **6**, 978–983 (2006).
19. Chen, Y. & Mao, C. Reprogramming DNA-directed reactions on the basis of a DNA conformational change. *J. Am. Chem. Soc.* **126**, 13240–13241 (2004).
20. Voigt, N. V. *et al.* Single-molecule chemical reactions on DNA origami. *Nature Nanotech.* **5**, 200–203 (2010).
21. Santoro, S. W. & Joyce, G. F. A general purpose RNA-cleaving DNA enzyme. *Proc. Natl Acad. Sci. USA* **94**, 4262–4266 (1997).
22. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
23. Trobro, S. & Aqvist, J. Mechanism of peptide bond synthesis on the ribosome. *Proc. Natl Acad. Sci. USA* **102**, 12395–12400 (2005).
24. Snyder, T. M., Tse, B. N. & Liu, D. R. Effects of template sequence and secondary structure on DNA-templated reactivity. *J. Am. Chem. Soc.* **130**, 1392–1401 (2008).
25. Gartner, Z. J. & Liu, D. R. The generality of DNA-templated synthesis as a basis for evolving non-natural small molecules. *J. Am. Chem. Soc.* **123**, 6961–6963 (2001).
26. Li, X. Y. & Liu, D. R. DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angew. Chem. Int. Ed.* **43**, 4848–4870 (2004).
27. Tse, B. N., Snyder, T. M., Shen, Y. H. & Liu, D. R. Translation of DNA into a library of 13000 synthetic small-molecule macrocycles suitable for *in vitro* selection. *J. Am. Chem. Soc.* **130**, 15611–15626 (2008).

### Acknowledgements

This work was supported by the Howard Hughes Medical Institute and National Institutes of Health/National Institute of General Medical Sciences (R01GM065865). The authors thank C. Dumelin and Y. Chen for insightful discussions and experimental assistance. The authors are grateful to Y. Shen for assistance with mass spectrometry.

### Author contributions

Y.H. and D.L. conceived and designed the project, analysed the data and wrote the manuscript. Y.H. performed the experiments.

### Additional information

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/naturenanotechnology](http://www.nature.com/naturenanotechnology). Supplementary information accompanies this paper at [www.nature.com/naturenanotechnology](http://www.nature.com/naturenanotechnology). Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>. Correspondence and requests for materials should be addressed to D.R.L.