



# Stereodivergent synthesis with a programmable molecular machine

**DOI:**

[10.1038/nature23677](https://doi.org/10.1038/nature23677)

**Document Version**

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

**Citation for published version (APA):**

Kassem, S., Lee, A. T. L., Leigh, D., Marcos Algaba, V., Palmer, L., & Pisano, S. (2017). Stereodivergent synthesis with a programmable molecular machine. *Nature*, *549*, 374-378. <https://doi.org/10.1038/nature23677>

**Published in:**

Nature

**Citing this paper**

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

**General rights**

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Takedown policy**

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.



## Stereodivergent synthesis with a programmable molecular machine

Salma Kassem,<sup>1</sup> Alan T. L. Lee,<sup>1</sup> David A. Leigh,<sup>1\*</sup> Vanesa Marcos,<sup>1</sup> Leoni I. Palmer<sup>1</sup> and Simone Pisano<sup>1</sup>

<sup>1</sup> *School of Chemistry, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*

**It has been convincingly argued<sup>1-3</sup> that molecular machines that manipulate individual atoms, or highly reactive clusters of atoms, with angstrom precision are unlikely to be realised. However, biological molecular machines routinely position rather less reactive substrates in order to direct chemical reaction sequences, from sequence-specific synthesis by the ribosome<sup>4</sup> to polyketide synthases,<sup>5-7</sup> where tethered molecules are passed from active site to active site in multi-enzyme complexes. Artificial molecular machines<sup>8-12</sup> have been developed for tasks that include sequence-specific oligomer synthesis<sup>13-15</sup> and the switching of product chirality,<sup>16-19</sup> a photo-responsive host molecule has been described that is able to mechanically twist a bound molecular guest,<sup>20</sup> and molecular fragments have been selectively transported in either direction between sites on a molecular platform through a ratchet mechanism.<sup>21</sup> Here we report on an artificial molecular machine that moves a substrate between different activating sites to achieve different product outcomes from chemical synthesis. This molecular robot can be programmed to stereoselectively produce, in a sequential one-pot operation, an excess of any one of four possible diastereoisomers from the addition of a thiol and an alkene to an  $\alpha,\beta$ -unsaturated aldehyde in a tandem reaction process. This includes diastereoisomers that cannot be made selectively<sup>22</sup> through conventional iminium-enamine organocatalysis. We anticipate that future generations of programmable molecular machines may play significant roles in chemical synthesis and molecular manufacturing.**

The unloaded form of programmable molecular machine **1** features a terminal alkene that enables attachment of an  $\alpha,\beta$ -unsaturated aldehyde substrate to an arm that can be steered by a pH-responsive

rotary switch embedded in a rigid platform bearing two chiral activating sites—prolinol silyl ethers—of opposite handedness (Figs 1 and 2). Control of the rotary switch allows the substrate to be positioned in close proximity to a selected activating site at any given stage of the machine operation. Condensation of the substrate with an activating site initiates an iminium or enamine promoted reaction. The asymmetric tandem addition of a nucleophile and an electrophile to an  $\alpha,\beta$ -unsaturated aldehyde under chiral prolinol silyl ether catalysis is widely used in organic synthesis to generate complex molecular structures with excellent enantio- and diastereocontrol.<sup>23,24</sup> However, the approach typically only allows access to the *syn*- diastereoisomers; the corresponding *anti*- isomers generally cannot be selectively generated through tandem iminium-enamine organocatalysis.<sup>25,26</sup> The stereochemical outcome of the tandem transformations promoted by **1** is dictated by the position of the arm at each stage. Accordingly, the machine can be programmed to selectively produce any of the full matrix of stereoisomeric products through different sequences of inputs (Fig. 1).

<Fig. 1 here>

The chemical structure of molecular machine **1** is shown in Fig. 2. The design features the two prolinol silyl ethers of opposite handedness connected to the quinoline stator of an acyl hydrazone rotary switch<sup>27,28</sup> *via* triazole linkages. The substrate, an  $\alpha,\beta$ -unsaturated aldehyde, is attached to the arm through an ester linkage. The ester group is sufficiently robust to withstand the assembly process and machine operation yet will allow the product to be released from the machine, by reductive cleavage, at the end of the program of reactions. Incorporation of the quinoline stator into the platform ensures that the substrate and the chosen activating site are held in fixed relative orientations in both configurations of the hydrazone switch,<sup>21</sup> which can be varied from 4:1 *E:Z* to 1:9 *E:Z* under the conditions of machine operation. The *para*-substituted phenyl ring adds rigidity to the arm, preventing the substrate from overreaching to the undesired activating site, while the four-carbon chain affords sufficient degrees of freedom for the substrate to interact with the chosen activating site without having to adopt

unfavourable strained conformations. The programs were performed at dilute (mM) concentrations of the machine in order to favour intra-machina over intermolecular reactions.<sup>13,14,21</sup>

<Fig. 2 here>

Four different programs (**A**, **B**, **C** and **D**) for the sequential one-pot operation of loaded molecular machine **2** are shown in Fig. 2. The reagents and conditions for each stage are essentially the same across all programs; the programs differ in whether a stage is used or not and the order in which they are applied within a program. The assembly is initiated by a cross-olefin-metathesis reaction (with acrolein dimethyl acetal) that loads the substrate onto the arm (Fig. 2, stage **I**), followed by deprotection of the masked aldehyde (Fig. 2, stage **II**, CF<sub>3</sub>COOH (2.2 equiv.)). Figs 3a and 3b show the partial <sup>1</sup>H NMR spectra of machine **2** before and after stage **I**, respectively. Downfield shifts of proton H<sub>a</sub>, from 4.78 ppm to the aldehyde region (9.51 ppm), and alkene protons H<sub>b</sub> and H<sub>c</sub> are indicative of hydrolysis of the dimethyl acetal to the α,β-unsaturated aldehyde (Fig 3c). A decision on the state of the switch (Fig. 2, stage **III**, *E*-**3**(2H<sup>+</sup>) or *Z*-**3**(3H<sup>+</sup>)) then differentiates programs **A** and **B** from programs **C** and **D**, and dictates the handedness of the activating site for the addition of the nucleophile (R<sup>1</sup>SH) (Fig. 2, stage **IV**, iminium activation, R<sup>1</sup> = (CH<sub>2</sub>)<sub>2</sub>(CF<sub>2</sub>)<sub>7</sub>CF<sub>3</sub>). If the decision on the switch state is *E*- (programs **A** and **B**), no additional reagents are added at this stage; if the decision is *Z*- (programs **C** and **D**) an additional input of CF<sub>3</sub>COOH (3 equiv.) is used to switch the hydrazone configuration.

Figs 3c-f show partial <sup>1</sup>H NMR spectra acquired during key stages of program **D** (see Supplementary Information for <sup>1</sup>H NMR spectra at key stages of each program **A-D**). Configurational isomerisation of the hydrazone rotary switch (*E*-**3**(2H<sup>+</sup>) to *Z*-**3**(3H<sup>+</sup>)) was monitored by <sup>1</sup>H NMR through the characteristic<sup>21,27,28</sup> upfield shift of proton H<sub>j</sub> (15.86 ppm in Fig. 3b to 14.38 ppm in Fig. 3c) and downfield shift of H<sub>i</sub> (4.52 ppm to 4.71 ppm). Subsequent addition of R<sup>1</sup>SH (stage **IV**) results in a change in H<sub>a</sub> (a doublet at 9.51 ppm in Fig. 3c to a broad triplet at 9.79 ppm in Fig. 3d; Fig. 3 inset) typical of 1,4-addition into an α,β-unsaturated aldehyde.

&lt;Fig. 3 here&gt;

Upon completion of the first step of the tandem transformation, a second decision on the state of the switch (Fig. 2, stage **V**) dictates the handedness of the activating site used for the addition of the electrophile ( $R^2$ ) (Fig. 2, stage **VI**, enamine-activation,  $R^2 = H_2C=C(SO_2Ph)_2$ ). If the switch is in the *E*-state, a change in the configuration of the switch (*E*-to-*Z*) is made by the addition of  $CF_3COOH$  (6 equiv.) (stage **V**, program **B**). Conversely, if the switch is in the *Z*-state and *Z*-to-*E* isomerisation is desired,  $Et_3N$  (7 equiv.) is added, deprotonating the pyridinium group of the hydrazone switch and restoring the thermodynamically favoured *E*-state of the switch (stage **V**, program **D**; Fig. 3e). Alternatively, the state of the switch can be chosen to remain unchanged (stage **V**, programs **A** and **C**). The presence of signals at 9.77 ppm and 9.64 ppm in the  $^1H$  NMR spectrum upon addition of  $R^2$  (Fig. 3f) is indicative of the formation of the second chiral centre, generating both *syn*- and *anti*-diastereoisomers. The mechanism of operation of the molecular machine during program **D** is shown in Fig. 4.

Reductive cleavage of the substrate (Fig. 2, stage **VII**) reduces the aldehyde and ester groups, releasing the substrate from the molecular machine to afford product **4**. The outcome of each program (**A-D**) was analysed by chiral HPLC and compared to authentic standards of each of the four possible stereoisomers of **4** (Supplementary Information, section S1.3). For each program, the conversions and diastereomeric and enantiomeric ratios of the product mixture were determined by (chiral) HPLC and  $^1H$  NMR spectroscopy and the absolute configuration of the major stereoisomer identified (Fig. 2).

Each of the four stereoisomers of **4** is the major product of a different program of the molecular machine (Fig. 2): *syn*- diastereoisomers (*2R,3R*)-**4** from **A** (73:27 e.r., 69:31 d.r.) and (*2S,3S*)-**4** from **D** (80:20 e.r., 70:30 d.r.); *anti*- diastereoisomers (*2S,3R*)-**4** from **B** (59:41 e.r., 63:37 d.r.) and (*2R,3S*)-**4** from **C** (74:26 e.r., 70:30 d.r.). The stereoselectivities are limited by the fidelity of the switch state (~10:90 to

~80:20 *E:Z*) during stages of each operation cycle, yet are comparable to those obtained for *syn*-diastereoisomers through conventional organocatalysis (Supplementary Information, section S1.3).

<Fig. 4 here>

The absolute configuration of C3 (carbon-3) in the product of all four programs is determined by the handedness of the activating site at stage **IV** (iminium activation): (*3R*)-configuration for reactions at the (*S*)-activating site (programs **A** and **B**); (*3S*)-configuration for reactions at the (*R*)-activating site (programs **C** and **D**). However, contrary to the original design concept, it is not the handedness of the activating site but the state of the switch during stage **VI** (enamine activation) that determines the relative configuration of C2 with respect to C3: when the switch is in the *E*-configuration (programs **A** and **D**) the addition of R<sup>2</sup> to C2 occurs *syn*- to the substituent at C3, whereas the *anti*- diastereoisomer is generated when the switch is in the *Z*-state (program **B** and **C**).

The contrasting behaviour of the two configurational states of the machine presumably results from the transiently formed macrocyclic intermediates produced upon enamine formation having different conformations, sizes and shapes on the different sides of the machine (Fig. 4). The issue illustrates the conceptual complexity involved in designing multi-component molecular machines in which the different parts have interacting roles.<sup>29</sup> In this case it is simply fortuitous that in the programmed operation of **2** one macrocycle favours a *syn*- product and the other *anti*-.

The stereodivergent synthesis of any of four possible diastereoisomeric products by a programmable molecular machine, including *anti*-diastereoisomers not selectively available through conventional iminium-enamine tandem organocatalysis, illustrates the potential of artificial molecular machines in chemical synthesis. This first generation machine augurs well for the development of small-molecule robots that can be programmed to manipulate substrates to control synthesis in a form of mechanosynthesis (that is, the use of mechanical constraints to direct reactive molecules to specific

molecular sites<sup>30</sup>), in a manner reminiscent of the way that molecular construction is carried out in biology.

## References:

1. Smalley, R. E. Of chemistry, love, and nanobots. *Sci. Am.* **285**(3), 76–77 (2001).
2. Whitesides, G. M. The once and future nanomachine. *Sci. Am.* **285**(3), 78–83 (2001).
3. Jones, R. A. L. *Soft Machines: Nanotechnology and Life* (Oxford Univ. Press, 2004).
4. Yonath, A. Hibernating bears, antibiotics, and the evolving ribosome (Nobel Lecture). *Angew. Chem. Int. Ed.* **49**, 4340–4354 (2010).
5. Maier, T., Leibundgut, M. & Ban, N. The crystal structure of a mammalian fatty acid synthase. *Science* **321**, 1315–1322 (2008).
6. Brignole, E. J., Smith, S. & Asturias, F. J. Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nat. Struct. Mol. Biol.* **16**, 190–197 (2009).
7. Chan, D. I. & Vogel, H. J. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem. J.* **430**, 1–19 (2010).
8. Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).
9. Abendroth, J. M., Bushuyev, O. S., Weiss, P. S. & Barrett, C. J. Controlling motion at the nanoscale: Rise of the molecular machines. *ACS Nano* **9**, 7746–7768 (2015).
10. Pan T. & Liu, J. Catalysts encapsulated in molecular machines. *ChemPhysChem.* **17**, 1752–1758 (2016).
11. Cheng, C. & Stoddart, J. F. Wholly synthetic molecular machines. *ChemPhysChem.* **17**, 1780–1793 (2016).
12. Astumian, R. D. How molecular motors work—insights from the molecular machinist's toolbox: the Nobel Prize in Chemistry 2016. *Chem. Sci.* **8**, 840–845 (2017).
13. Lewandowski, B. et al. Sequence-specific peptide synthesis by an artificial small-molecule machine. *Science* **339**, 189–193 (2013).
14. De Bo, G., Kuschel, S., Leigh, D. A., Lewandowski, B., Pappmeyer, M. & Ward, J. W. Efficient assembly of threaded molecular machines for sequence-specific synthesis. *J. Am. Chem. Soc.* **136**, 5811–5814 (2014).
15. Meng, W. et al. An autonomous molecular assembler for programmable chemical synthesis. *Nat. Chem.* **8**, 542–548 (2016).
16. Wang, J. & Feringa, B. L. Dynamic control of chiral space in a catalytic asymmetric reaction using a molecular motor. *Science* **331**, 1429–1432 (2011).
17. Mortezaei, S., Caterineu, N. R. & Canary, J. W. A redox-reconfigurable, ambidextrous asymmetric catalyst. *J. Am. Chem. Soc.* **134**, 8054–8057 (2012).
18. Zhao, D., Neubauer, T. M. & Feringa, B. L. Dynamic control of chirality in phosphine ligands for enantioselective catalysis. *Nat. Commun.* **6**, 6652 (2015).
19. Mortezaei, S., Catarineu, N. R. & Canary, J. W. Dial-in selection of any of four stereochemical outcomes among two substrates by in situ stereo-reconfiguration of a single ambidextrous catalyst. *Tetrahedron Lett.* **57**, 459–462 (2016).
20. Muraoka, T., Kinbara, K. & Aida, T. Mechanical twisting of a guest by a photoresponsive host. *Nature* **440**, 512–515 (2006).
21. Kassem, S., Lee, A. T. L., Leigh, D. A., Markevicius, A. & Solà, J. Pick-up, transport and release of a molecular cargo using a small-molecule robotic arm. *Nat. Chem.* **8**, 138–143 (2016).
22. Krautwald, S. & Carreira, E. M. Stereodivergence in asymmetric catalysis. *J. Am. Chem. Soc.* **139**, 5627–5639 (2017).
23. Mielgo, A. & Palomo, C.  $\alpha,\alpha$ -Diarylprolinol ethers: new tools for functionalization of carbonyl compounds.

*Chem.–Asian J.* **3**, 922–948 (2008).

24. Jensen, K. L., Dickmeiss, G., Jiang, H., Albrecht, L. & Jørgensen, K. A. The diarylprolinol silyl ether system: a general organocatalyst. *Acc. Chem. Res.* **45**, 248–264 (2012).

25. Marigo, M., Schulte, T., Franzén, J. & Jørgensen, K. A. Asymmetric multicomponent domino reactions and highly enantioselective conjugated addition of thiols to  $\alpha,\beta$ -unsaturated aldehydes. *J. Am. Chem. Soc.* **127**, 15710–15711 (2005).

26. Zhao, G.-L., Rios, R., Vesely, J., Eriksson, L. & Córdova, A. Organocatalytic enantioselective aminosulfonylation of  $\alpha,\beta$ -unsaturated aldehydes. *Angew. Chem., Int. Ed.* **47**, 8468–8472 (2008).

27. Su, X. & Aprahamian, I. Switching around two axles: controlling the configuration and conformation of a hydrazone-based switch. *Org. Lett.* **13**, 30–33 (2011).

28. Ray, D., Foy, J. T., Hughes, R. P. & Aprahamian, I. A switching cascade of hydrazone-based rotary switches through coordination-coupled proton relays. *Nat. Chem.* **4**, 757–762 (2012).

29. Kay, E. R. & Leigh, D. A. Rise of the molecular machines. *Angew. Chem., Int. Ed.* **54**, 10080–10088 (2015).

30. Drexler, K. E. *Engines of Creation: The Coming Era of Nanotechnology* (Anchor Books, New York, 1986).

## Figure legends

**Figure 1** | **a**, Flowchart showing the programmable synthesis of any one of four stereoisomers by molecular machine **1**. **b**, The molecular machine's response to command inputs in program **D**. The stereochemistry of each chiral centre in the product was anticipated to show dependence on the handedness of the activating site utilized in the step that forms that centre. However, the results (Fig. 2) show that while the absolute configuration is determined by the activating site handedness (*R* or *S*) during iminium activation (stage **IV**), the relative configuration (*syn*- or *anti*-) is determined by the switch state (*E* or *Z*) during enamine activation (stage **VI**). Inset: switching process.

**Figure 2** | Programmable stereodivergent synthesis of all possible stereoisomers of **4** from a one-pot tandem iminium-enamine promoted reaction sequence by molecular machine **1**. Reagents and conditions: **I** acrolein dimethyl acetal (100 equiv.), Hoveyda-Grubbs Catalyst™ 2<sup>nd</sup> generation (20 mol%), CH<sub>2</sub>Cl<sub>2</sub> (1.5 mM), 96 h, 38 %; **II** CF<sub>3</sub>COOH (2.2 equiv.), r.t., 10 min, CD<sub>2</sub>Cl<sub>2</sub> (1 mM); **III** (*E*→*Z*) CF<sub>3</sub>COOH (3.0 equiv.), r.t., 5 h, CD<sub>2</sub>Cl<sub>2</sub> (1 mM); **IV** R<sup>1</sup>SH (100 equiv.), 0 °C, 30–41 h, CD<sub>2</sub>Cl<sub>2</sub> (1 mM); **V** (*E*→*Z*) CF<sub>3</sub>COOH (6.0 equiv.), 0 °C, 20 h, CD<sub>2</sub>Cl<sub>2</sub> (1 mM); **V** (*Z*→*E*) Et<sub>3</sub>N (7.0 equiv.), 120 μL CD<sub>3</sub>CN, 0 °C, 6 h, CD<sub>2</sub>Cl<sub>2</sub> (1 mM); **VI** R<sup>2</sup> (200 equiv.), 0 °C, 24 h, CD<sub>2</sub>Cl<sub>2</sub> (1 mM); **VII** NaBH<sub>4</sub>, MeOH, 0 °C, 2 h, CD<sub>2</sub>Cl<sub>2</sub> (1 mM), then LiAlH<sub>4</sub>, THF, –78 °C to r.t., 1 h. Stages **II–VI** are carried out as a one-pot sequence. R<sup>1</sup> = (CH<sub>2</sub>)<sub>2</sub>(CF<sub>2</sub>)<sub>7</sub>CF<sub>3</sub>. R<sup>2</sup> = H<sub>2</sub>C=C(SO<sub>2</sub>Ph)<sub>2</sub>. Program sequences. **A**: **I–II–IV–VI–VII**. **B**: **I–II–IV–V–VI–VII**. **C**: **I–II–III–IV–VI–VII**. **D**: **I–II–III–IV–V–VI–VII**. d.r. (diastereomeric ratio) and e.r. (enantiomeric ratio) determined by chiral stationary phase HPLC (see Supplementary Information).

**Figure 3** | Partial <sup>1</sup>H NMR spectra (600 MHz, 295 K, CD<sub>2</sub>Cl<sub>2</sub>) of program **D** at distinctive stages of operation. **a**, Molecular machine **1**. **b**, Stage **I**: Loaded molecular machine **2** followed by sequential one-pot addition of: **c**, Stages **II**, **III** (*E*→*Z*): CF<sub>3</sub>COOH (5.0 equiv.), r.t., 5 h, >98 % conversion to *E*:*Z*(10:90)-**3**(3H<sup>+</sup>); **d**, Stage **IV**: R<sup>1</sup>SH (100 equiv.), 0 °C, 41 h, >98 % conversion to *E*:*Z*(10:90)-(*S*)-R<sup>1</sup>S-**3**(3H<sup>+</sup>); **e**, Stage **V** (*Z*→*E*): Et<sub>3</sub>N (7.0 equiv.), 0 °C, 6 h, >98 % conversion to *E*:*Z*(80:20)-(*S*)-R<sup>1</sup>S-**3**; **f**, Stage **VI**: R<sup>2</sup> (200 equiv.) 0 °C, 24 h, >98 % conversion to *E*:*Z*(80:20)-(*S*)-R<sup>1</sup>S-(*S*)-R<sup>2</sup>H-**3**. Inset: proton H<sub>a</sub> before and after additions of R<sup>1</sup>SH and R<sup>2</sup>. The *syn* and *anti* diastereoisomers were assigned by comparison with known compounds (Supplementary information, Section S1.3). Dashed lines connect resonances indicative of the configuration of the hydrazone switch (H<sub>j</sub> and H<sub>i</sub>) and the chemical transformation of the substrate (H<sub>a</sub>, H<sub>b</sub> and H<sub>c</sub>). Signals from traces of residual solvent and excess R<sup>2</sup> are shown in grey. \* unreacted *Z*-**3**(3H<sup>+</sup>). For full proton assignments of **1** and **2**, see Supplementary Information.

**Figure 4** | Mechanism of operation of molecular machine **1** during program **D** (**I–II–III–IV–V–VI–VII**). Stages **II–VI** carried out as a one-pot sequence (see Fig. 2 for reagents and conditions) with each operation proceeding in >98 % conversion to the next stage as evidenced by <sup>1</sup>H NMR spectroscopy (fidelity of switching *E*-to-*Z* ~10:90 *E*:*Z*; *Z*-to-*E* ~80:20 *E*:*Z*). Intermediates *Z*-**3**(3H<sup>+</sup>), *Z*-(*S*)-R<sup>1</sup>S-**3**(3H<sup>+</sup>), *E*-(*S*)-R<sup>1</sup>S-**3** and *E*-(*S*)-R<sup>1</sup>S-(*S*)-R<sup>2</sup>H-**3** observed by <sup>1</sup>H NMR (Figs 3c, 3d, 3e and



3f, respectively). Square brackets indicate intermediates not observed by  $^1\text{H}$  NMR (*Z*-**3**(*Im*-3H<sup>+</sup> and *E*-(*S*)-R<sup>1</sup>S-**3**(*En*)). R<sup>1</sup> = (CH<sub>2</sub>)<sub>2</sub>(CF<sub>2</sub>)<sub>7</sub>CF<sub>3</sub>. R<sup>2</sup> = CH<sub>2</sub>C(SO<sub>2</sub>Ph)<sub>2</sub>.

### Supplementary Information

Detailed synthetic procedures, operation methods and full characterisation data.

### Author Contributions

VM devised the concept. SK, VM and LIP carried out the experimental work. SP and LIP performed model studies. SK, VM and ATLL designed the operation experiments. DAL directed the research. All the authors contributed to the analysis of the results and the writing of the manuscript.

### Corresponding Author

david.leigh@manchester.ac.uk.

### Notes

The authors declare no competing financial interests.

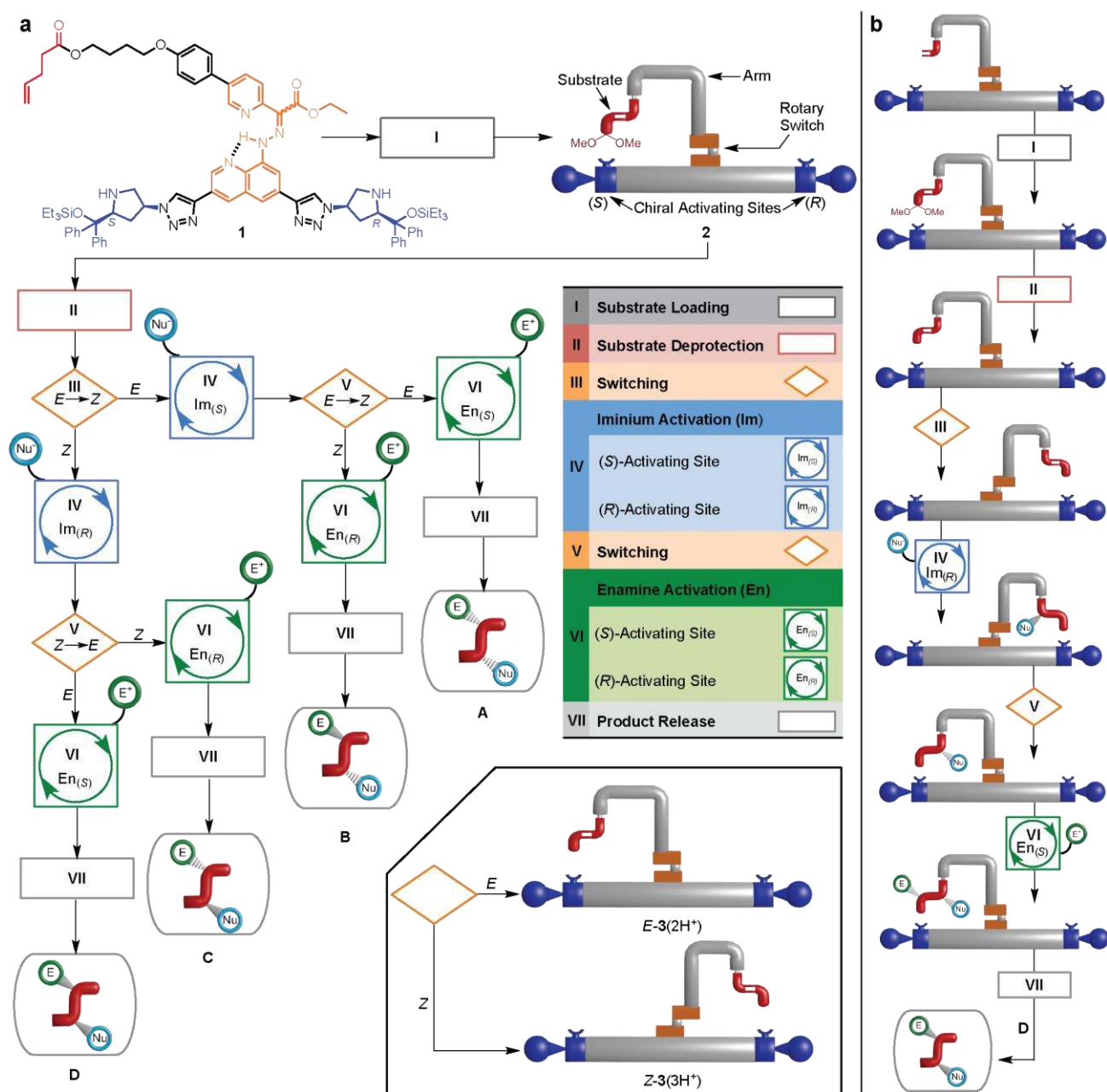
### Acknowledgements

We thank the Engineering and Physical Sciences Research Council (EPSRC) (EP/H021620/1 & 2) and the European Research Council (ERC) (Advanced Grant No. 339019) for funding, and the EPSRC National Mass Spectrometry Service Centre (Swansea, UK) for high resolution mass spectrometry. DAL is a Royal Society Research Professor.

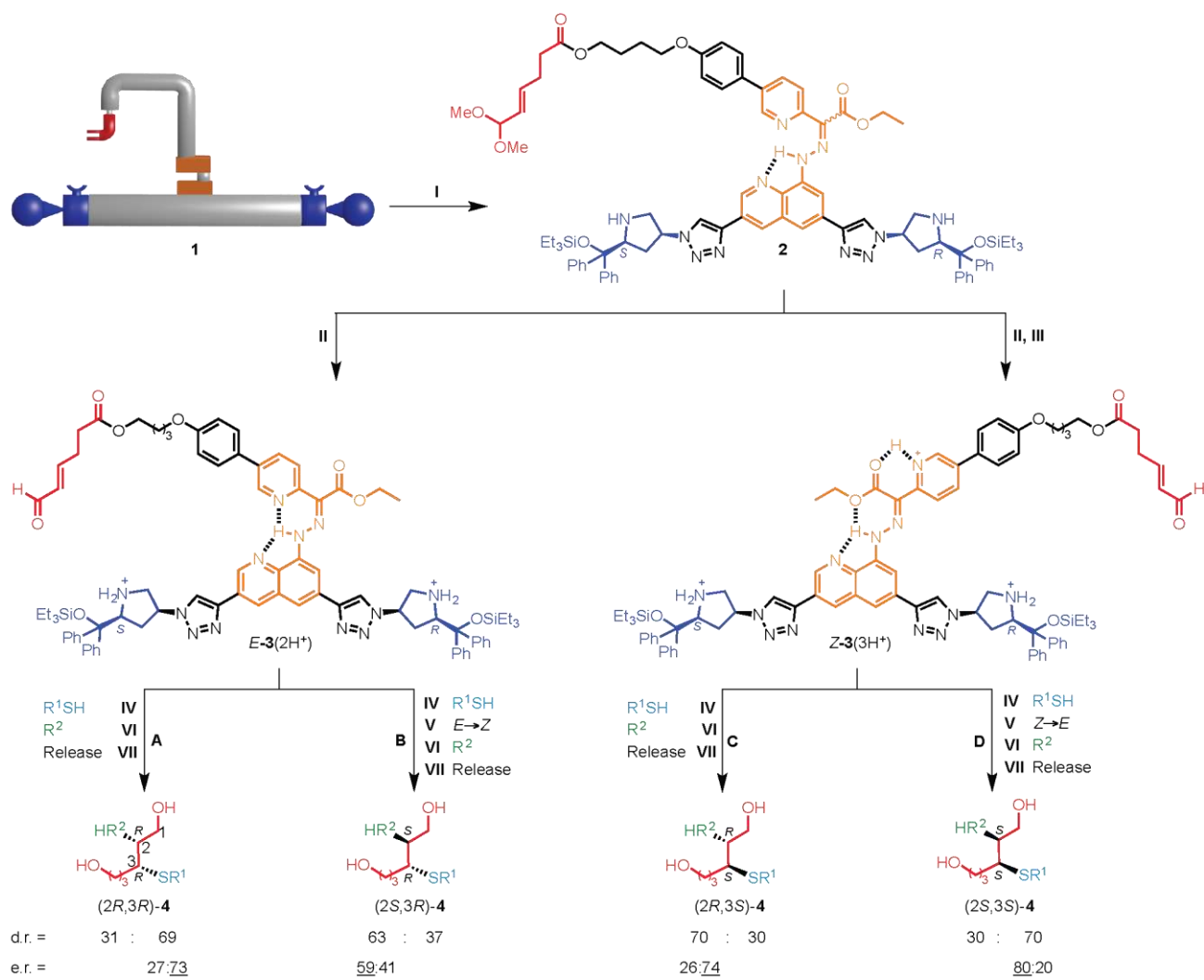
### Data availability statement

The data that support the findings of this study are available within the paper and its Supplementary Information, or are available from the Mendeley Data repository with the identifier doi 10.17632/bj4yktgnfj.1.

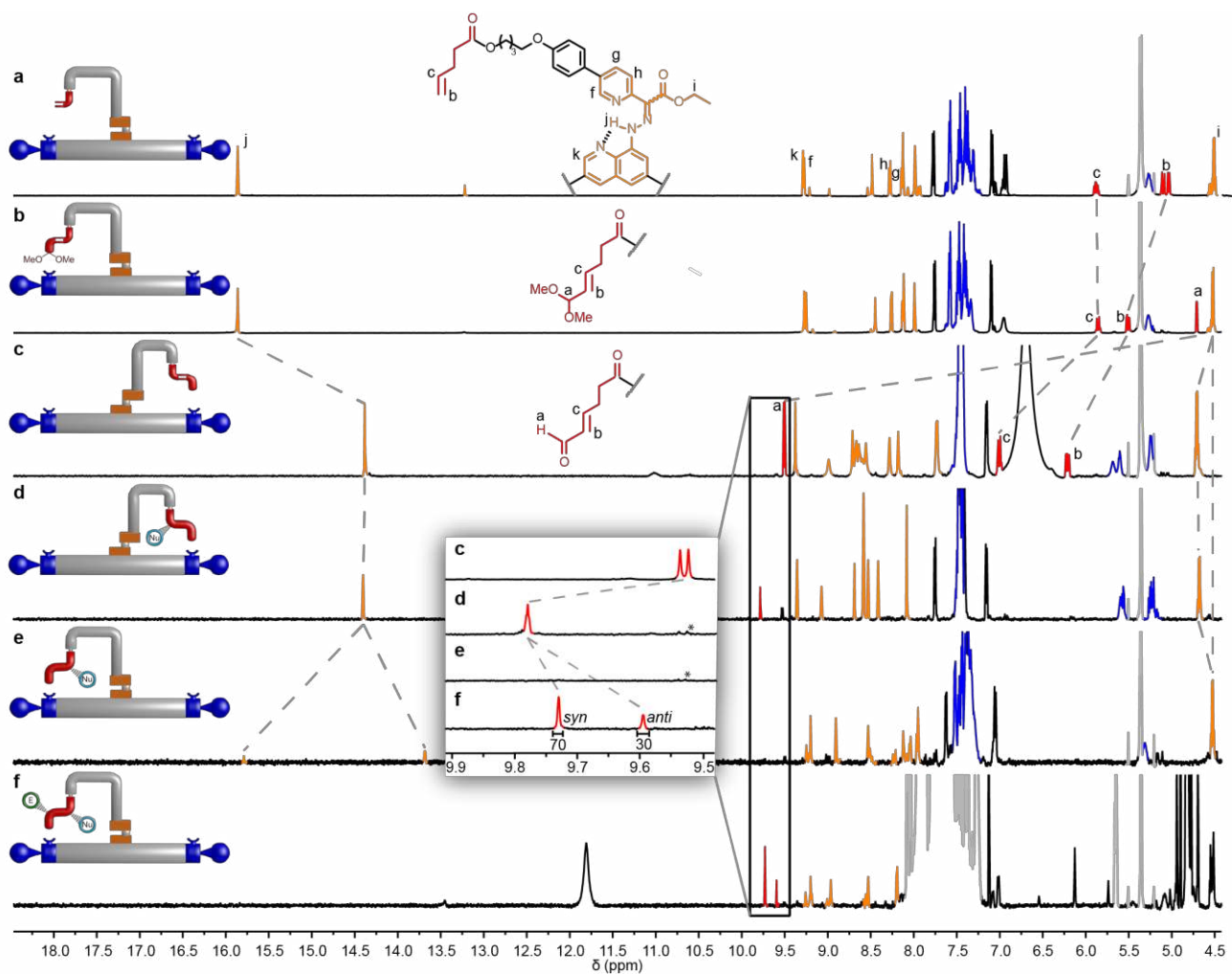
Figures



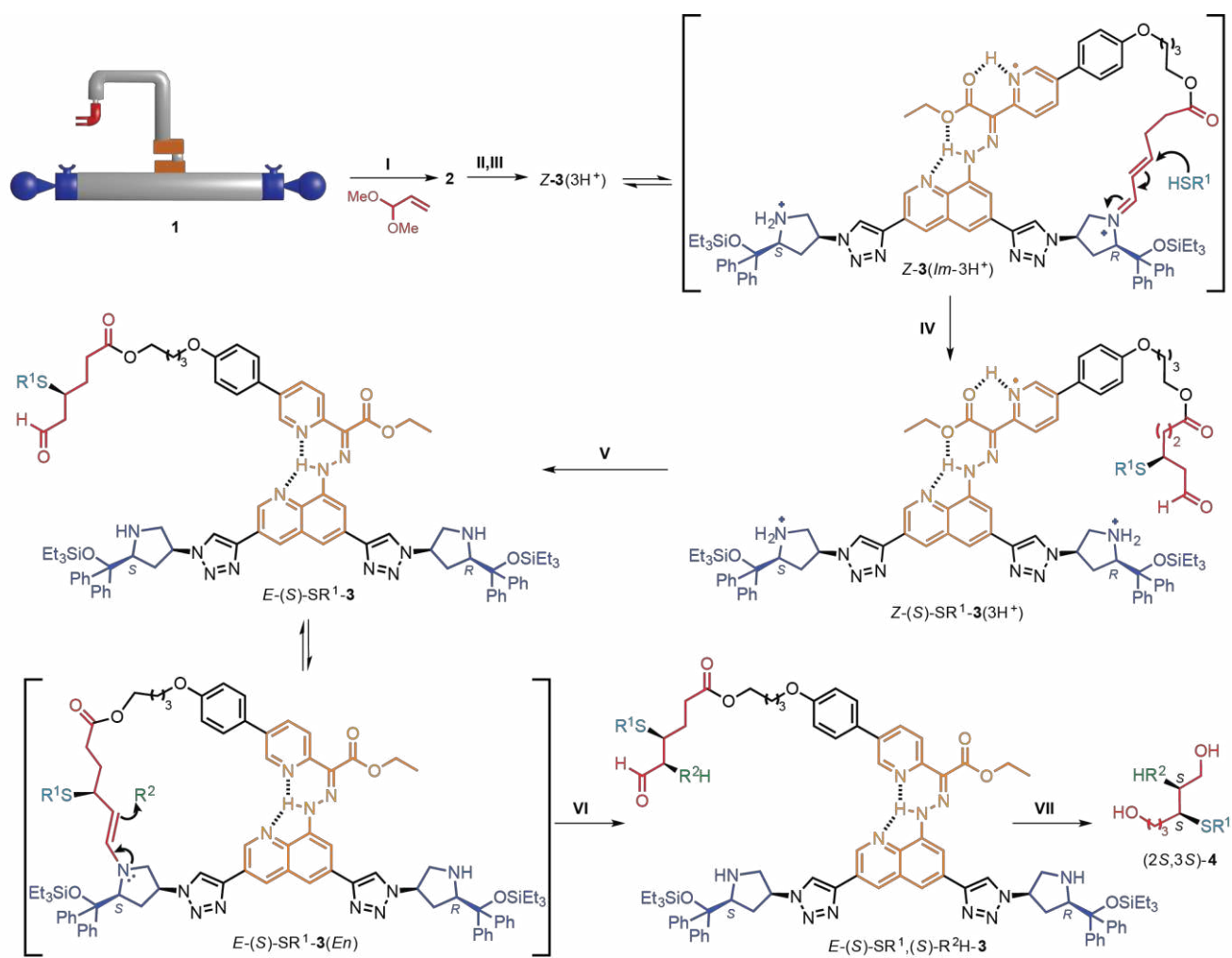
<Figure 1>



<Figure 2>



<Figure 3>



<Figure 4>