DOI: 10.1002/cbic.201300248

A Rotaxane Turing Machine for Peptides

Claire Margaret Wilson,[b] Andrea Gualandi,[a] and Pier Giorgio Cozzi*[a]

Life is based on molecular machines and their ability to process information. Cells process information, reading and translating the genetic code, in a manner similar to a universal Turing machine.[1] They are capable of autonomous conversion of an input-data-encoding molecule to an output molecule according to a set of rules defined by a molecular program. The logic of life is carried through specific nucleotide sequences and is transcribed into defined sequences of amino acids that make up peptides. The ribosome is the cellular molecular machine employed in this process.[2] Such machines are found in prokaryotic (~2.6 MD) and eukaryotic (~4.3 MD) cells.

In a manner similar to a Turing machine, the sequence of amino acids in a peptide is assembled by the forward movement of the ribosome along an mRNA strand. Information from the strand is interpreted, and amino acids are incorporated step by step. Less-complex artificial molecular machines[3] can be ingeniously designed and synthesised with the ability to perform similar actions. They can also be programmed to store information,[4a] perform mechanical work[4b] and carry out specific reactions.[4c] Complex properties such as catalytic behaviour can be regulated by artificial switches,[4d] and even more intricate molecular properties, such as handedness, can be influenced by a chiral molecular machine.[4e] The Turing machine operated by scanning a data tape, whose striking analogy to information-encoding biopolymers inspired several designs for molecular DNA computers. It is now possible to build molecular machines with the ability to set up specific reactions in a programmed sequence.[5] In fact, Leigh and co-workers have recently published an example of an artificial read up reactor with many similarities to a Turing machine. The machine contains the sequence of reactions to be performed, and when it reads this sequence, the desired peptidic bonds are formed. The Turing machine is described mathematically as consisting of: 1) a finite set of states, *S*, 2) a finite alphabet, *I*, 3) a starting state, and 4) a partial function from $S\times I_0$ to $S\times I_n$ (*R*, *L*) where *R*, *L* are the movement of the conveyor in both directions.

Amino acids (chemical entities, *I*) are linked through peptidic bonds by this miniaturised version of a ribosome. The artificial molecular machine (**4**) is a rotaxane[6] containing a functionalized macrocycle (**2**; Scheme 1).

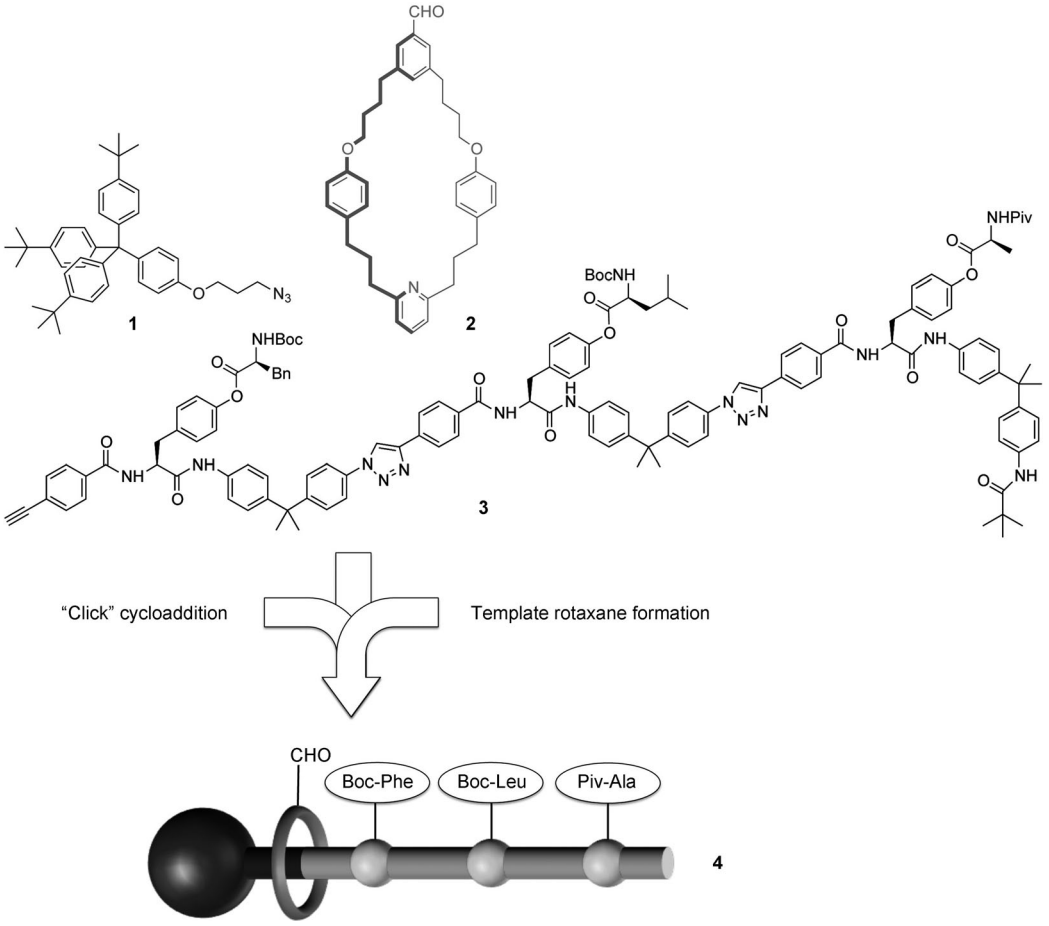
[a] *Dr. A. Gualandi, Prof. P. G. Cozzi*
Dipartimento di Chimica "G. Ciamician", Alma Mater Studiorum
Università di Bologna
Via Selmi 2, 40126 Bologna (Italy)
E-mail: piergiorgio.cozzi@unibo.it

[b] *C. M. Wilson*
Centre for Synthesis and Chemical Biology
School of Chemistry and Chemical Biology, University College Dublin
Dublin 4 (Ireland)

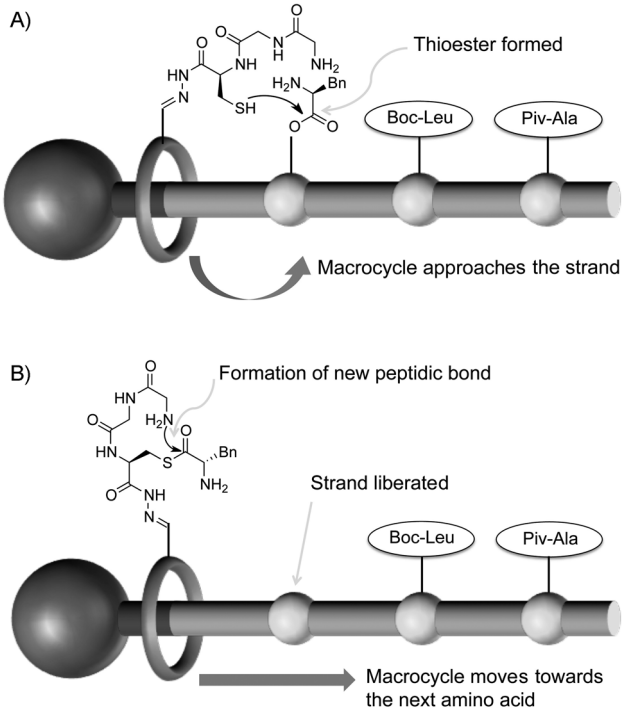
The macrocycle determines the order of the amino acids in the peptide chain by its motion along the axis of the molecular machine (*R*, *L* movement). In this way, the microsynthesiser can process information as it operates. The axis of the rotaxane, which bears three amino acids attached by weak phenolic ester linkages, acts as the strand for the machine. Spacers are inserted to allow movement of the macrocycle-bound reactive arm along the thread. In addition, they control the folding of the thread. The "moonwalking" of the macrocycle along the strand allows reaction with the amino acid building blocks in the correct sequence (the "operation" $S\times I$ of the machine).

The formation of such a complex machine would be not possible without powerful methods for assembling the rotaxane. The macrocycle (**2**) contains a copper-coordinating pyridine. The strand is threaded through the macrocycle by the selective and versatile Cu^{I}-catalysed cycloaddition of the terminal alkyne with the azide-bearing stopper group (**1**).[7] The active template strategy invented by the Leigh group[8] afforded the desired compound in 30% yield. A moiety capable of bearing the reactive peptidic chain must also be inserted in the macrocycle. A bond of suitable strength is also required to allow the simple detachment of the peptidic sequence from the macrocycle after the desired reactions have been performed. In this case, a hydrazone exchange is used to introduce a cysteine derivative bearing the reactive arm and the site for peptide elongation [a trityl (Trt)-protected thiol group and a *tert*-butoxycarbonylcarbamate (Boc)-protected amine at the end of a glycylglycine residue respectively].

Removal of the protective Boc and Trt protecting groups is the trigger that activates the machine. Microwave heating in acetonitrile/dimethylformamide (3:1) in the presence of *N*,*N*-diisopropylethylamine allows the formation of peptidic bonds. In addition, tris(2-carboxyethyl)phosphine is added as a reducing agent to cleave any disulfide bonds formed through thiol oxidation. With the thiol of the cysteine group deprotected, the presence of the organic base makes the thiol nucleophilic and capable of undergoing transacylation as the macrocycle approaches the first reactive site (S_0 to S_1). The first reactive site is the amino acid phenolic ester, which blocks the macrocycle from travelling further along the axis of the rotaxane. The thioester formed then transfers the amino acid to another site on the macrocycle (Scheme 2A) in a method ubiquitous in bioconjugation, known as native chemical ligation.[9] Once the covalent bond connecting the amino acid to the strand is broken, the macrocycle can travel further along the axis of the rotaxane and moves towards the second amino acid on the strand (S_1 to S_2; Scheme 2B). The O→S acyl transfer/S→N acyl transfer processes continue as the macrocycle travels along the axis. When the last amino acid has been incorporated into



Scheme 1. Synthesis of the functionalised molecular machine **4** from a stopper group (**1**), functionalized macrocycle (**2**) and axis (**3**).



Scheme 2. Machinery **4** at work.

the macrocycle, the macrocycle is free to detach from the strand with the hydrazine-linked peptide.

The artificial molecular machine synthesises the peptide from the C to the N terminus, although in a relatively long time (36 h) compared to a ribosome. Other bond-forming reactions are possible with this machine, although some adjustment is required. However, one problem must still be solved, and will probably be addressed in the future design of molecular Turing ribosomic machines. The information that the axis bears is lost as the molecule exits the machine. Another machine is required for the synthesis of another sequence. Of course, this is not a desirable situation. A possible solution is in the core design of the machine, that is, in the choice of linker between the macrocycle and the reactive arm that takes the amino acids from the thread. The linkage used by Leigh in this molecular machine can be easily cleaved, thus allowing for analysis of the resulting peptide. The use of a reversible bond would allow the machine to be readjusted for the synthesis of another peptide. The Turing machine could then read in the reverse and in the forward direction. The next question would be how to control the sequence of amino acids. In a reversible Turing machine, the formation of the amino acid sequence could be promoted by another arm, temporarily installed on the macrocycle. Such a reversible ribosomic Turing machine, with the ability to read back and forth and to introduce further

instructions onto the strand, could possibly be based on a catenane or made by introducing selective "gates" to trap the macrocycle at the end of the sequence. This will be another exciting and difficult chapter to look forward to in this interesting field.

Acknowledgements

University of Bologna, PRIN (Rome), and Fondazione del Monte are acknowledged for the support of our research.

Keywords: molecular machines · peptides · ribosomes · rotaxanes

[1] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 6th ed., Freeman, New York, **2006**.
[2] A. Yonath, *Angew. Chem.* **2010**, *122*, 4438; *Angew. Chem. Int. Ed.* **2010**, *49*, 4340.
[3] E. R. Kay, D. A. Leigh, F. Zerbetto, *Angew. Chem.* **2007**, *119*, 72; *Angew. Chem. Int. Ed.* **2007**, *46*, 72.
[4] a) C. P. Collier, E. W. Wong, M. Belohradský, F. M. Raymo, J. F. Stoddart, P. J. Kuekes, R. S. Williams, J. R. Heath, *Science* **1999**, *285*, 391; b) T. Muraoka, K. Kinbara, T. Aida, *Nature* **2006**, *440*, 512; c) C. Monnereau, P. H. Ramos, A. B. Deutman, J. A. Elemans, R. J. Nolte, A. E. Rowan, *J. Am. Chem. Soc.* **2010**, *132*, 1529; d) H. J. Yoon, J. Kuwabara, J.-H. Kim, C. A. Mirkin, *Science* **2010**, *330*, 66; e) J. Wang, B. L. Feringa, *Science* **2011**, *331*, 1429.
[5] a) B. Lewandowski, G. De Bo, J. W. Ward, M. Papmeyer, S. Kuschel, M. J. Aldegunde, P. M. E. Gramlich, D. Heckmann, S. M. Goldup, D. M. D'Souza, A. E. Fernandes, D. A. Leigh, *Science* **2013**, *339*, 189. For a highlight, see: b) P. R. McGonigal, J. F. Stoddard, *Nat. Chem.* **2013**, *5*, 260.
[6] W. L. Yang, Y. J. Li, H. B. A. Liu, L. F. Chi, Y. L. Li, *Small* **2012**, *8*, 504.
[7] A. C. Fahrenbach, J. F. Stoddart, *Chem. Asian J.* **2011**, *6*, 2660.
[8] K. D. Hänni, D. A. Leigh, *Chem. Soc. Rev.* **2010**, *39*, 1240.
[9] L. Raibaut, N. Ollivier, O. Melnyk, *Chem. Soc. Rev.* **2012**, *41*, 7001.

Received: April 17, 2013
Published online on ■■ ■■, 0000