# Molecular Computing: The Activities Of The MolCoNet Network

Giancarlo Mauri
Università di Milano-Bicocca, Milano, Italy

## Introduction

The MolCoNet network is centered on computing with bio-sequences, in particular on algorithmic and information-theoretic aspects. The main goal is to foster joint work and exchange of results and knowledge among the partners and to assess the state-of-the art, especially concerning theoretical models of molecular computing, such as splicing systems and membrane systems, the understanding of computational processes in living organisms and the design and implementation of DNA-based algorithms for solving difficult problems. The network includes 13 members from 10 different European countries. The next sections will give a short overview of the main research areas these teams were involved in.

## Molecular Computing: Algorithmic and information theoretic aspects

The experiment that started the area of DNA computing [1], demonstrated the potential of performing computations on a nano-scale. Inspired by an unexpected similarity between the functioning of polymerase enzymes and that of a Turing machine, Adleman investigated if and how bio-molecules such as DNA, RNA, proteins can be used as computational data carriers in chemical solutions. Since then, the area of DNA computing evolved through a series of different experiments, mainly aimed initially at solving other computationally hard problems, as the satisfiability problem (SAT), the subset sum problem, or breaking the data encryption standard (DES). Rapid advances in biotechnology made it possible, e.g., to scale-up the inputs or to use RNA, or proteins. Gradually, other kinds of applications than addressing hard mathematical problems were considered. Far from being "real" applications, or challenging the supremacy of electronic computers, what these experiments showed was that nano-level structures and processes can be manipulated in a very controlled, a-priori established way. So, it became evident that we need new algorithmic techniques and new models of computation to fully exploit the potential of molecular computing.

The theory of splicing systems introduced in 1987 by Tom Head, is by now a mature area with well-established research topics [2]. In 1998, G. Paun introduced a new theoretical framework for molecular computations, modeling the cell structure and taping its potential for computations: the theory of membrane systems [3]. This area attracted the attention of many teams in Molconet and elsewhere and developed in a very rapid pace, spreading from Europe to USA and Japan.

## Self-assembly

Through all this experimental and theoretical work, the molecular computing community gradually recognized that a completely new computing paradigm has been born: computing= data design + self-assembly

Indeed, the crucial ingredient of molecular computations, provided for free by Nature, is the process of self-assembly: two complementary strands of DNA always bind to each other to form a double strand. In this way, computation reduces to designing the DNA strands in such a way that both the input and the algorithm itself are encoded in the DNA strands.

Being able to control the process of self-assembly in molecular computations brings a whole new perspective and opens up the possibility of nano-scale design and nano-constructions. The molecular computing community was quick to investigate this possibility and several exciting experiments have been reported. E.g., it was shown that DNA complexes can be directed to form even Sierpinski triangles on a nano-scale. Taking this further, experiments showed the possibility of building ever more complex nano-structures: DNA-based scissors that can open and close, nanotubes (with potential for electronics layout) or even DNA-based motors that can ``walk" along a track built from DNA as well.

Sophisticated formal results on combinatorics of words and theory of codes have been developed to deal with problems related to DNA-strand design for molecular computing experiments. It was clear already from Adleman's experiment that the strands encoding the input problem have to be designed in such a way that they can only bind to each other in the desired way, have about the same melting temperature, and so on [4].

**Computation in living cells**

The research on DNA and molecular computing has also stressed the fact that living cells are very sophisticated devices that manipulate information. So, we need new models to understand how they treat information. The theory of pointer rewriting systems for example has been introduced to formalize the process of gene assembly in ciliates [5]. The theoretical results (on permutations, strings, and graphs) provide a formal framework for reasoning and giving predictions about an involved biological process. MolCoNet contributed to the development of this very promising area of research: thanks to the exchange of information among the partners, some of them started to cooperate on this subject.

**Modeling biological systems**

Systems biology is but one new dynamic direction that attracts nowadays the attention of biologists, mathematicians, and computer scientists and in which many techniques and insights from Molecular Computing may prove invaluable. The goal is to understand the behavior of a complex system like the living cell. Some of the teams in MolCoNet started to cooperate on this subject, producing some models of mechanisms such as mechanosensitive channels in prokaryotes, or the p53 pathway, that seems very promising from a biological point of view, based on membrane systems (see Sec.2).

**Future Perspectives**

Based on the idea of combining nano-scale design and self-assembly, and on breakthroughs in other fields such as physics, electronics, or material sciences, the possibility of building involved nano-constructs became real. This opens up an enormous potential to exciting new applications in nano-science and nano-technology.

In parallel, the field of Molecular Computing also evolved towards computing in vivo, i.e., computational processes that occur both within and between living cells, in order for them to function and survive in dynamic environments. Understanding these cellular processes will inspire new algorithmic techniques and adaptable hardware and software architectures for solving difficult and ill-defined problems across many application areas. On the other hand, insights into in-vivo computational processes will deepen the understanding of the functioning of the cell and ultimately enable researchers to harness living cells as computational devices and to program them as "nano-bots" for particular tasks such as targeted drug delivery, chemical factories, nano-structures, bio-film scaffolding and self-assembling.

Based on the bits and pieces we currently know about living cells, it is hoped we will soon be able to build an integrated understanding of how entire cells or even entire organisms function, adapt, and survive in dynamic environments.

**REFERENCES**


[1] L. M. Adleman, Molecular computation of solutions to combinatorial problems. Science, 266, 1994:1021-1024.

[2] Gh. Paun, G. Rozenberg, A. Salomaa, DNA Computing: New Computing Paradigms, Springer-Verlag, 1998.

[3] Gh. Paun, Membrane Computing. An Introduction, Springer-Verlag, Berlin, 2002.

[4] G. Mauri, C. Ferretti, Word design for DNA computing: a survey, LNCS 2943, Springer-Verlag, 2003, 37–46

[5] A. Ehrenfeucht, T. Harju, I. Petre, D. M. Prescott, and G. Rozenberg, Computation in Living Cells: Gene Assembly in Ciliates, Springer (2003).