

Discussion animée par Misha Gromov (IHES)

Introduction Eric Westhof

Graziano Vernizzi (CEA-Saclay)  
Des noeuds aux pseudonoeuds.

Philippe Dumas (IBMC)  
Dynamique de systèmes biologiques non linéaires

Introduction Edouard Belaga

Yaakov Benenson (Weizmann Institute)  
Automata and Antisense.

Jacques Ninio (Paris VI)  
La mémoire.

Alain Denise (Orsay)  
Bioinformatique des génomes.

Dear Eric,

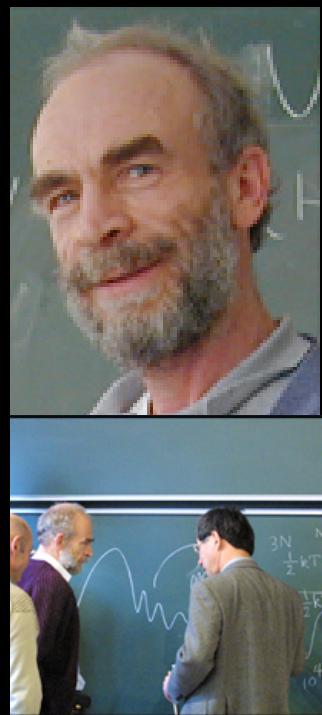
I am sorry I can not attend your meeting as I have not quite recovered from the surgery.

I hope very much you will keep the initiative and I will be most happy to come to your next meeting in Strasbourg. It is a great idea, bringing together biologists and mathematicians. Making such a gathering is hard, I know by personal experience, but this is rewarding to all participants. Eventually, when a good rapport is established, biologists will benefit from mathematicians' help in structuring experimental data and/or in optimizing experiment design.

Mathematicians, in their turn, will open to new kind of problems -some intellectually beautiful, some not so smooth but practically useful, and all challengingly hard to solve. And long before significant results are achieved, biologists and mathematicians will have good time talking together and learning from each other new ways of looking at the great old problem: "What is the nature and the structure of life?"

I wish you all enjoy a great day,

Misha



**HUMAN FRONTIER SCIENCE PROGRAM (HFSP)**  
12 quai St Jean, 67080 STRASBOURG Cedex, FRANCE  
E-mail: grant@hfsp.org  
Web site: <http://www.hfsp.org>

**OPPORTUNITIES FOR INTERDISCIPLINARY RESEARCH**

The Human Frontier Science Program (HFSP) supports international collaborations in basic research with emphasis placed on novel, innovative and interdisciplinary approaches to fundamental investigations in the life sciences. Applications are invited for grants to support projects on complex mechanisms of living organisms.

**CALL FOR LETTERS OF INTENT FOR RESEARCH GRANTS: AWARD YEAR 2006**

The HFSP research grant program aims to stimulate novel, daring ideas by supporting collaborative research involving biologists together with scientists from other disciplines such as chemistry, physics, mathematics, computer science and engineering. Recent developments in biological, physical, applied sciences and new disciplines such as bioinformatics and nanoscience open up new approaches to understanding the complex mechanisms underlying biological functions in living organisms. Preliminary results are not required in research grant applications. Applicants are expected to develop new lines of research and to demonstrate that the proposed research must be distinct from applicants' other research funded by other sources. HFSP supports only international, collaborative teams, with an emphasis on encouraging scientists early in their careers.

**International teams of scientists interested in submitting applications for support must first submit a letter of intent online via the HFSP web site. The guidelines for potential applicants and further instructions are available on the HFSP web site ([www.hfsp.org](http://www.hfsp.org)).**

Research grants provide 3 years support for teams with 2 – 4 members, with not more than one member from any one country, unless more members are absolutely necessary for the interdisciplinary nature of the project, which is an essential selection criterion. Applicants may also establish a local interdisciplinary collaboration as a component of an international team (see below). The principal applicant must be located in one of the member countries\* but co-investigators may be from any other country. Clear preference is given to intercontinental teams.

**TWO TYPES OF GRANT ARE AVAILABLE:**

**Young Investigators:** Grants are for teams of scientists who are all within 5 years of establishing an independent laboratory and within 10 years of obtaining their PhDs. Successful teams will receive up to \$450,000 per year for the whole team. Scientists involved in a local interdisciplinary collaboration are considered as 1.5 team members for budgetary purposes.

**Program Grants:** are for independent scientists at all stages of their careers, although the participation of younger researchers is especially encouraged. Program grants provide up to \$450,000 per year for the whole team. Scientists involved in a local interdisciplinary collaboration are considered as a single team member for budgetary purposes.

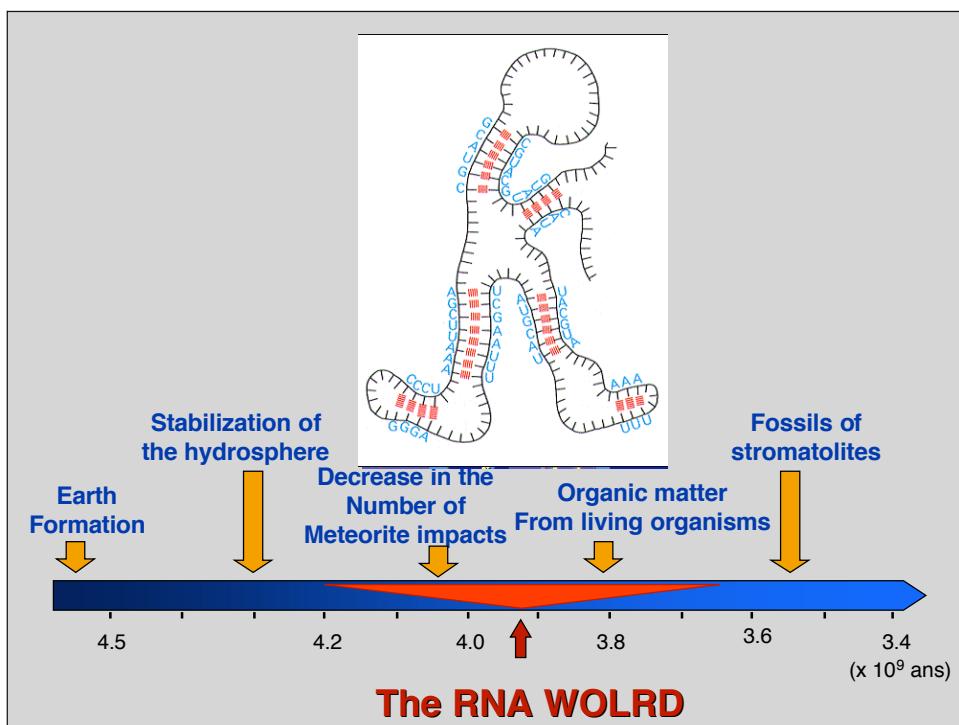
**Important Deadlines :**  
**Compulsory pre-registration for password: 21 MARCH 2005**  
**Submission of Letters of Intent: 31 MARCH 2005**

\*Members are Australia, Canada, the European Union (including the 10 new member countries), France, Germany, Italy, Japan, the Republic of Korea, Switzerland, the United Kingdom and the United States.

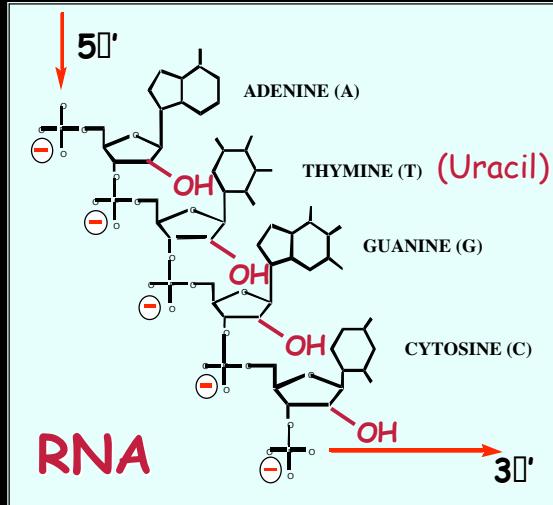
New full member countries for award year 2006 are Australia and the Republic of Korea

## Why RNA ?

- The only molecule which can be both genome and catalyst;
- Origin of life some 3.9 billion years ago;
- RNA molecules pervade every single step in molecular biology;
- A very frequent target of antibiotics.



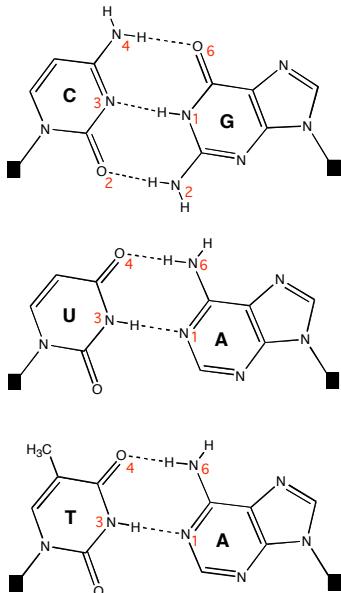
## Nucleic acids are negatively charged biopolymers ...



### Why mathematicians and RNA ?

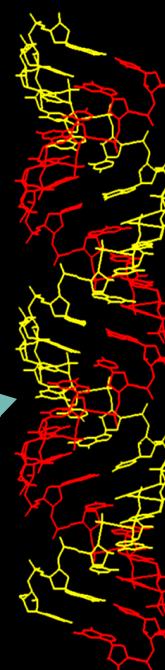
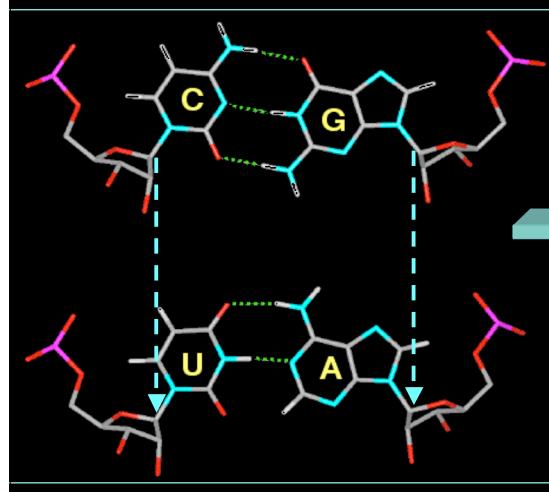
- Only 4 letters (A, G, C, U);
- They form complementary base pairs ( $G=C$ ;  $A-U$ );
- Secondary structures can be computed based on energy tables;
- Most appropriate for molecular evolution (experimentally & theoretically).

**Horizontal  
Interactions**  
**Base pairing.**  
**In helices**  
**Complementary**  
**Watson-Crick**

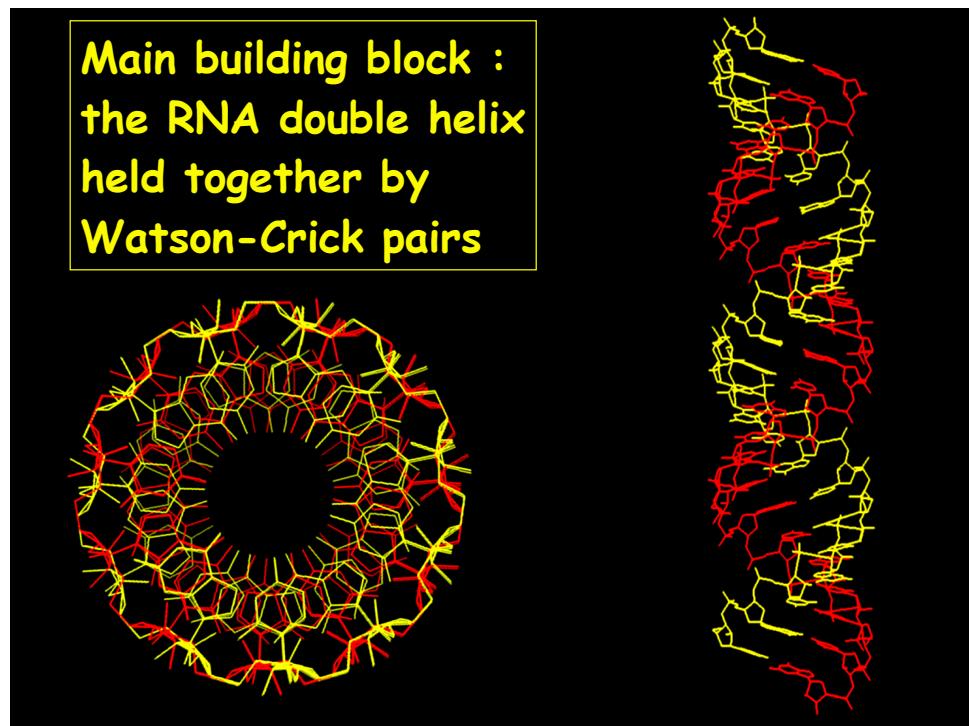


$55^\circ$        $55^\circ$   
 10.5 Å

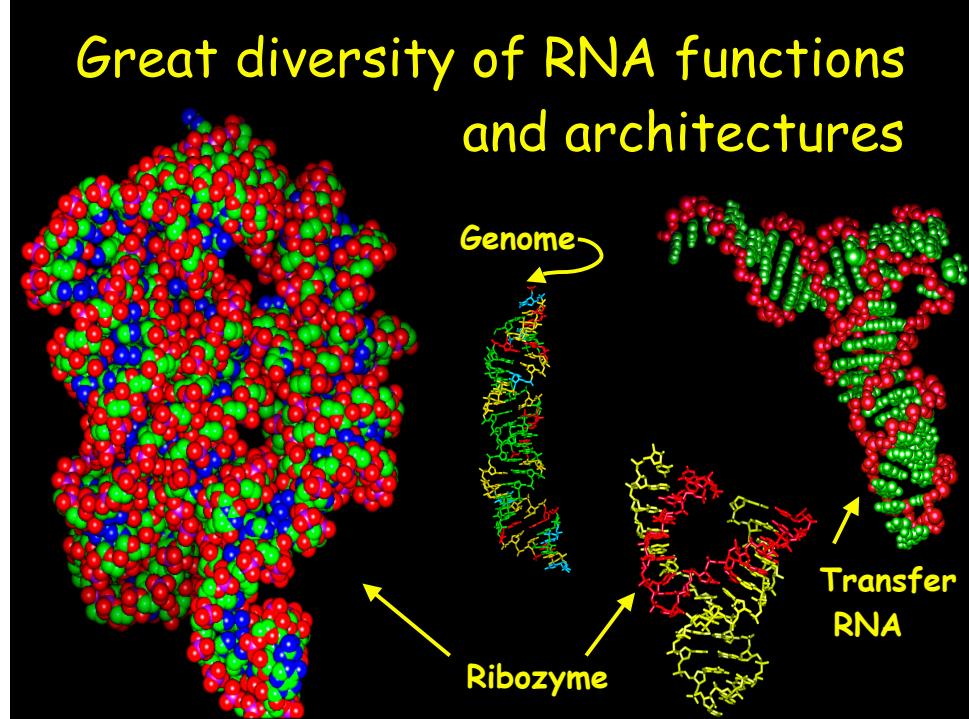
**Watson-Crick pairs are  
isosteric**



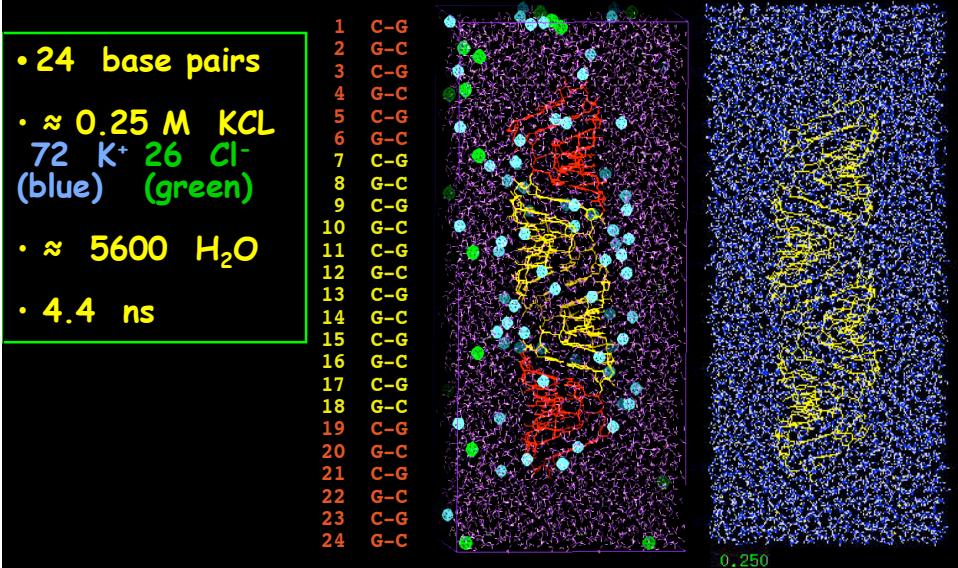
**Main building block :  
the RNA double helix  
held together by  
Watson-Crick pairs**



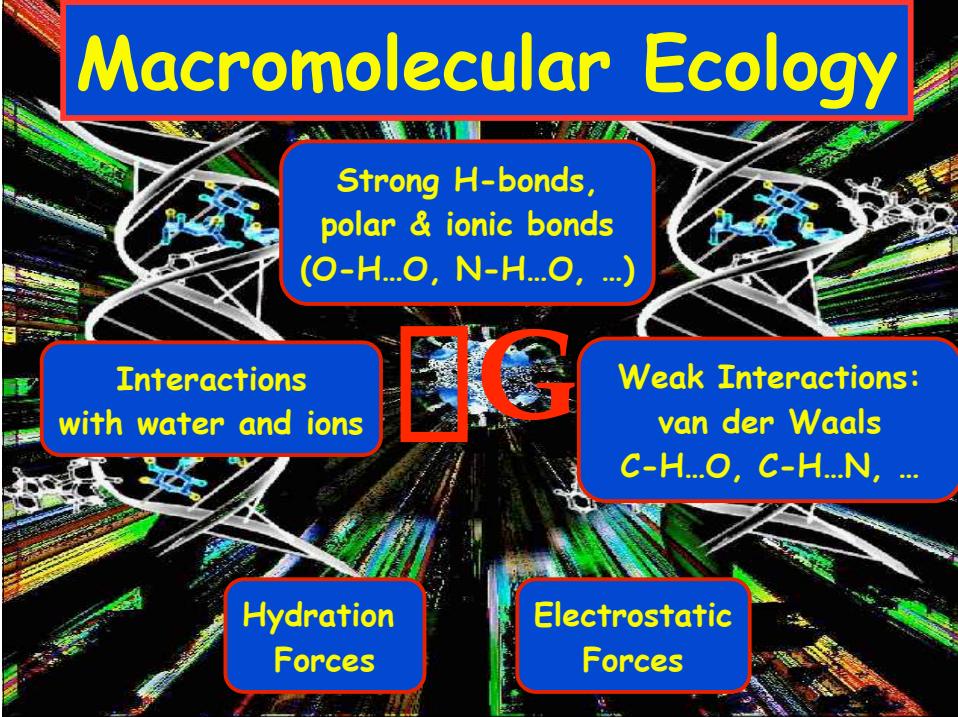
**Great diversity of RNA functions  
and architectures**

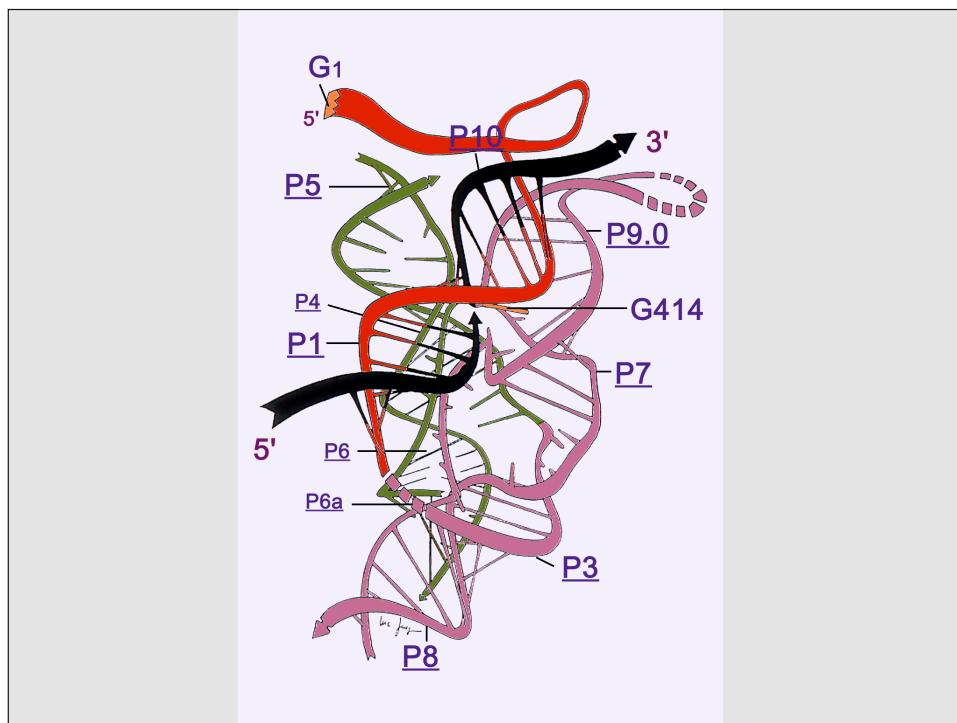
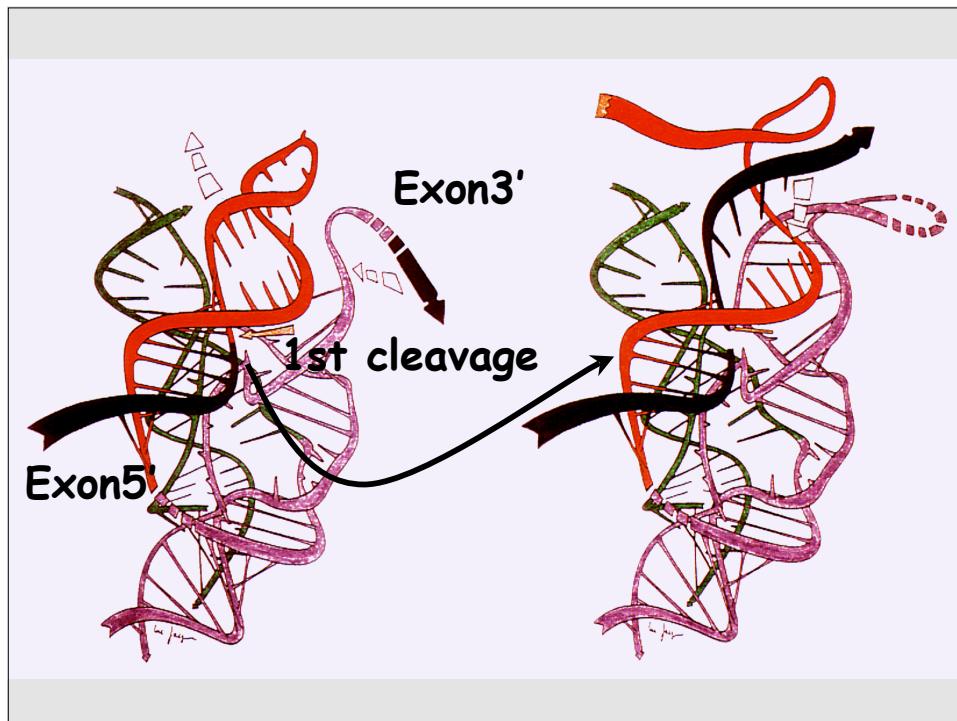


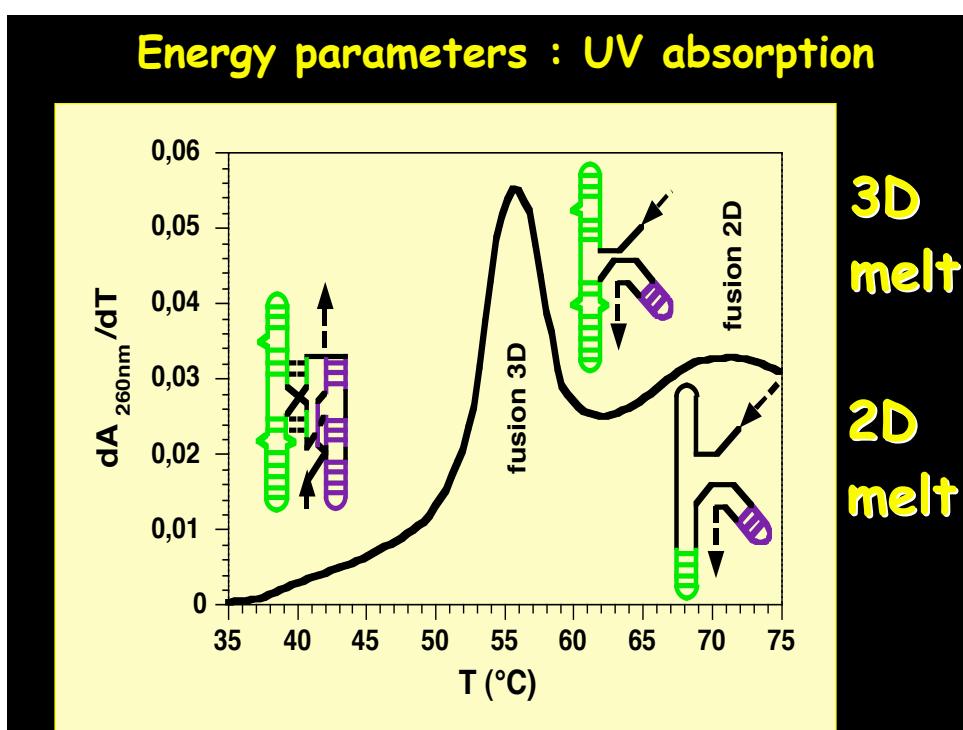
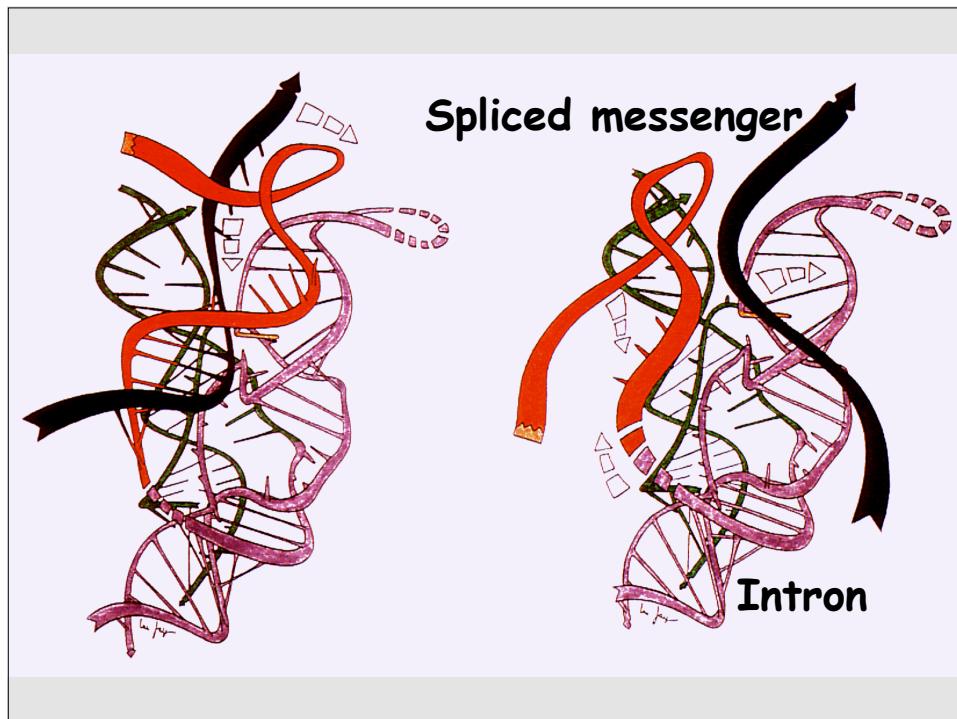
## Molecular dynamics simulations (P. Auffinger)



## Macromolecular Ecology





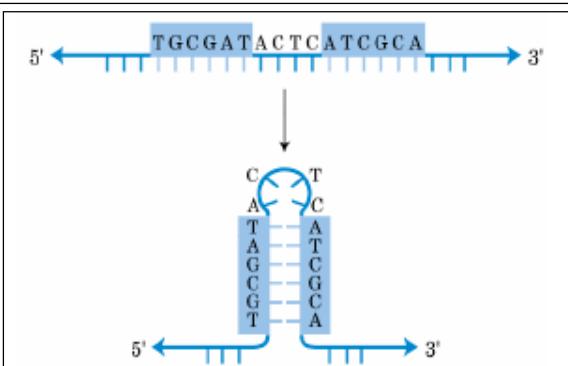


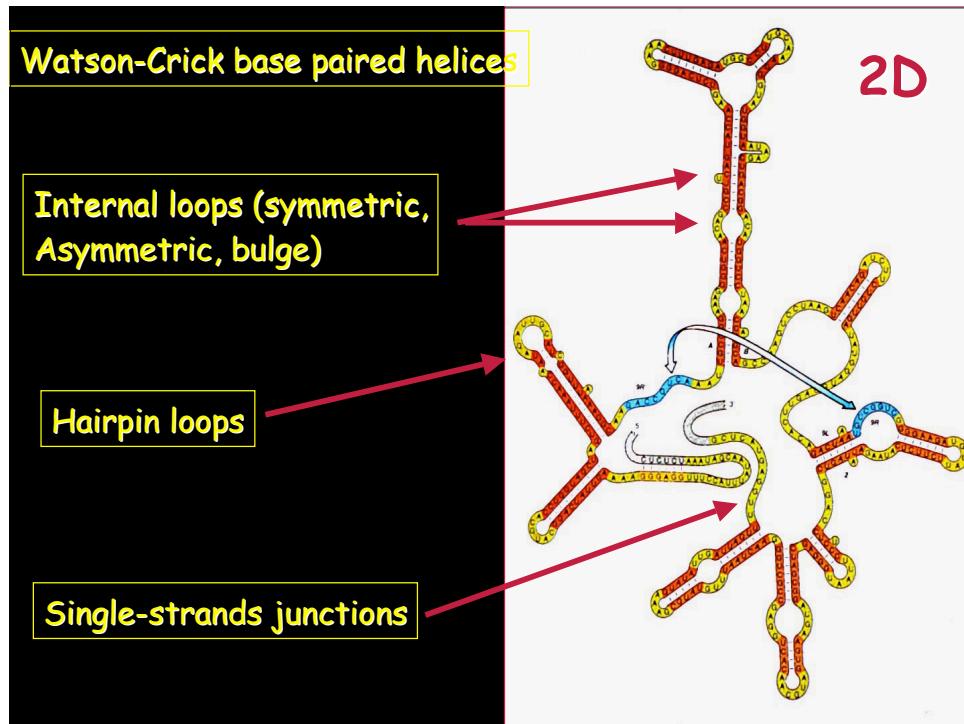
## How to fold a RNA sequence ?

1. Biological approach.

2. Theoretical approach.

(Single-stranded) nucleic acids fold on themselves





- Grammatically correct string of parentheses  
.....((.....)).((((.....)))......)  
AGCTACGGAGCGATCTCCGAGCTTTCGAGAAACCTCTATTAGC
- Planar graph
- Arch diagram
- Mountain diagram

# How to fold a RNA sequence ?

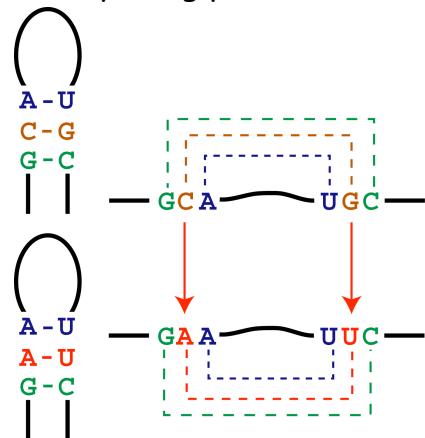
## 1. Biological approach.

### Comparative Sequence Analysis

- Structures diverge slower than sequences
- Molecules with similar functions and different nucleotide sequences will form similar structures.
- Predicts secondary and tertiary structure from underlying sequence.
- Correctly identifies high percentage secondary structure pairings and a smaller number of tertiary interactions.
- Primarily a manual method

## RNA alignments

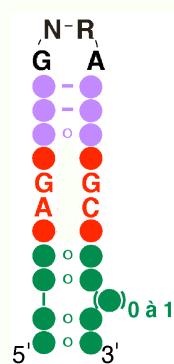
- RNA sequences are aligned/compared differently because sequence variation in RNA maintain base-pairing patterns



- Thus an alignment should exhibit **covariation** at interacting basepairs

Comparisons of sequences  
lead  
to the 2D structures

((((.....((((....)))).....)).))  
GCUGGAGGGAAAGCAAUUUAGCACG-GC  
GCUUCAGUAGAGCGAUCUAGCGGAAGU  
AGGAAGUGGGAGUAAUCCAGCAUCGCU  
CUUCGAGGUUCGAAAAGAGUGCAGA-GG



## How to fold a RNA sequence ?

### 2. Theoretical approach.

#### Theoretical approach : Basic Model

- RNA linear structure:  $R=r_1 r_2 \dots r_n$  from  $\{A,C,G,U\}$
- RNA secondary structure: pairs  $(r_i, r_j)$  such that  $0 < i < j < n+1$ .
- Goal: secondary structures with minimum free energy.

## Implementing Model Restrictions

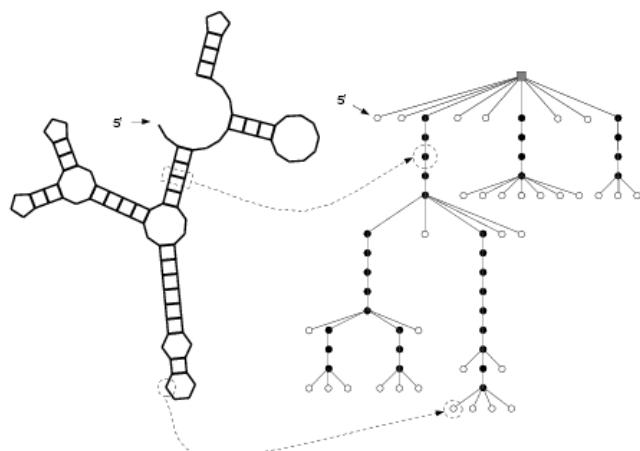
- No knots: pairs  $(r_i, r_j)$  and  $(r_k, r_l)$  such that  $i < k < j < l$ . RNA does contain knots.
- No “close” base pairs:  $j - i \geq t$  for some  $t > 0$ .
- Complementary base pairs: A-U, C-G (wobble pair G-U).

## Tinoco-Uhlenbeck postulate

- Assumption: The energy of each base pair is independent of all of the other pairs and the loop structure.
- Consequence: Total free energy is the sum of all of the base pair free energies.

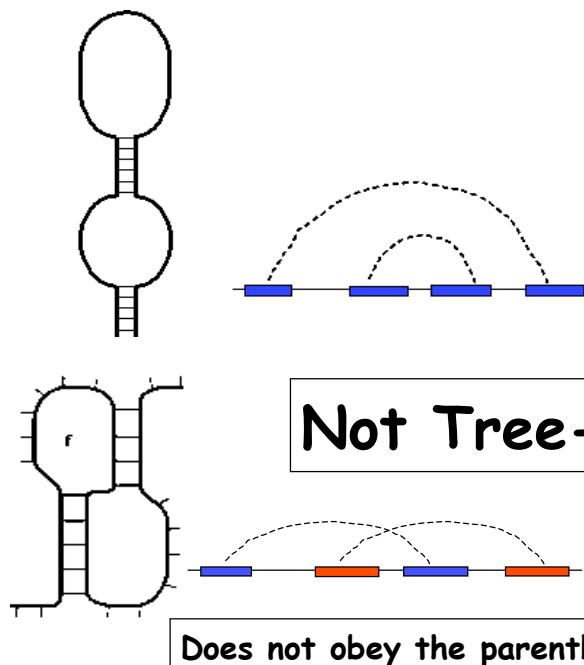
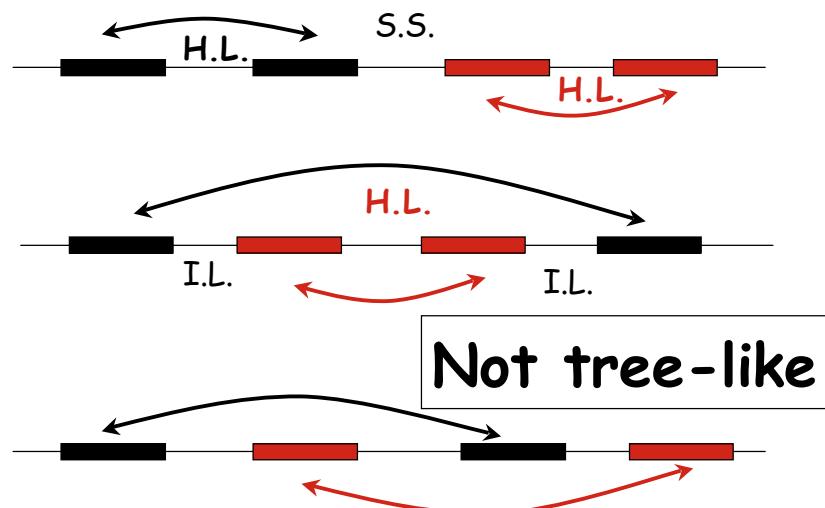
## Independent Base Pairs

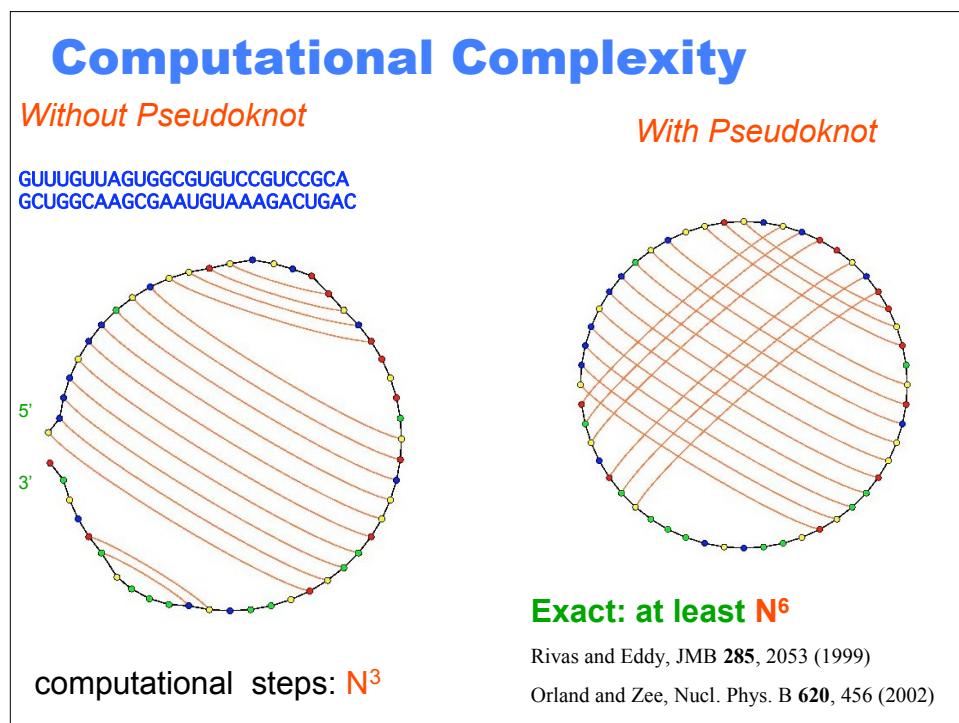
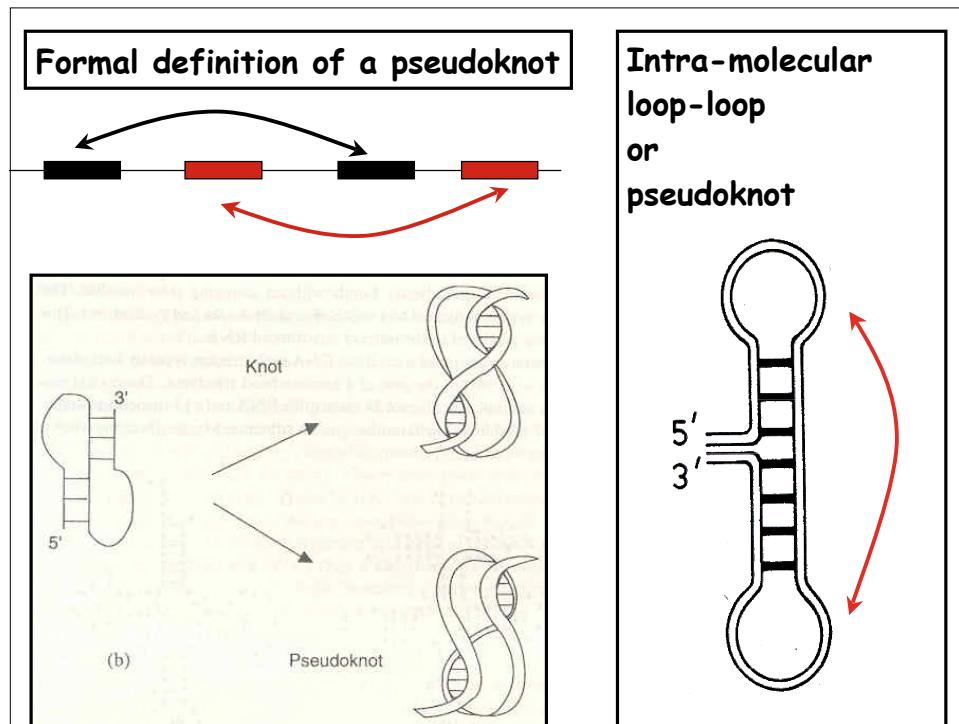
- Use solutions for smaller strings to determine solutions for larger strings.
- This is **precisely** the kind of decoupling required for dynamic programming algorithms to work.



Secondary structure as its tree representation

**Only three ways to pair four segments**

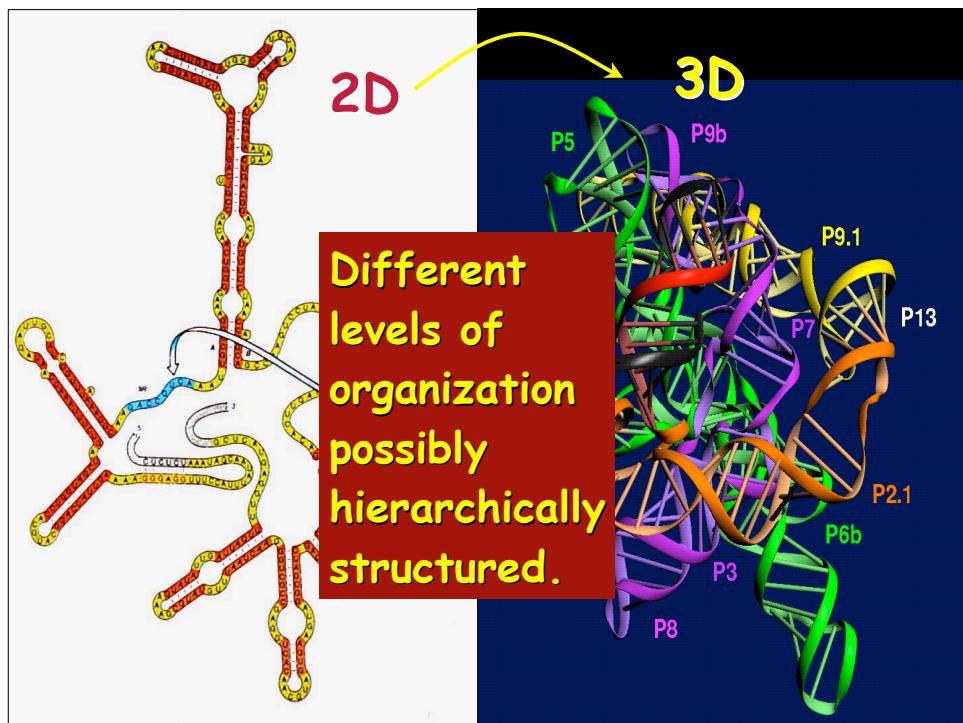




## SANKOFF's problem

align and derive the 2D structure  
from a set of non-aligned sequences

NP-complete !



**Les êtres vivants sont en fait  
des structures historiques.  
F.Jacob.**

**All biological sequences result from  
structural constraints,  
biological adaptation,  
genetic drift,  
and historical contingencies.**

Discussion animée par Misha Gromov (IHES)

Introduction Eric Westhof

Graziano Vernizzi (CEA-Saclay)  
Des noeuds aux pseudonoeuds.

Philippe Dumas (IBMC)  
Dynamique de systèmes biologiques non linéaires

Introduction Edouard Belaga

Yaakov Benenson (Weizmann Institute)  
Automata and Antisense.

Jacques Ninio (Paris VI)  
La mémoire.

Alain Denise (Orsay)  
Bioinformatique des génomes.