Computational Chemistry in Synthetic Biology?

Patrick Merlot
patrick.merlot@kjemi.uio.no

CTCC seminar

Norwegian Centre of Excellence

March 18, 2011

Outline

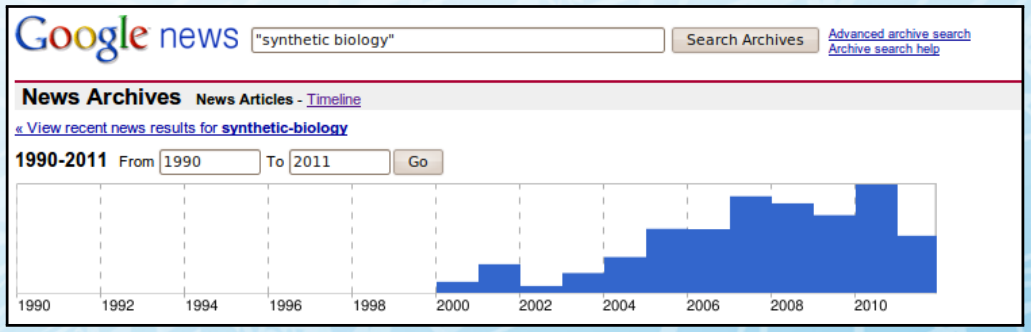
- What is Synthetic Biology?
- How to computationally design ”improved” proteins
- What are the challenges for Computational chemistry?

Outline

- **What is Synthetic Biology?**
  - Objectives and Foundations
  - History of Cells
  - Steps for synthetising new proteins
  - Challenges of synthetic biology
- How to computationally design ”improved” proteins
- What are the challenges for Computational chemistry?

What is Synthetic Biology?

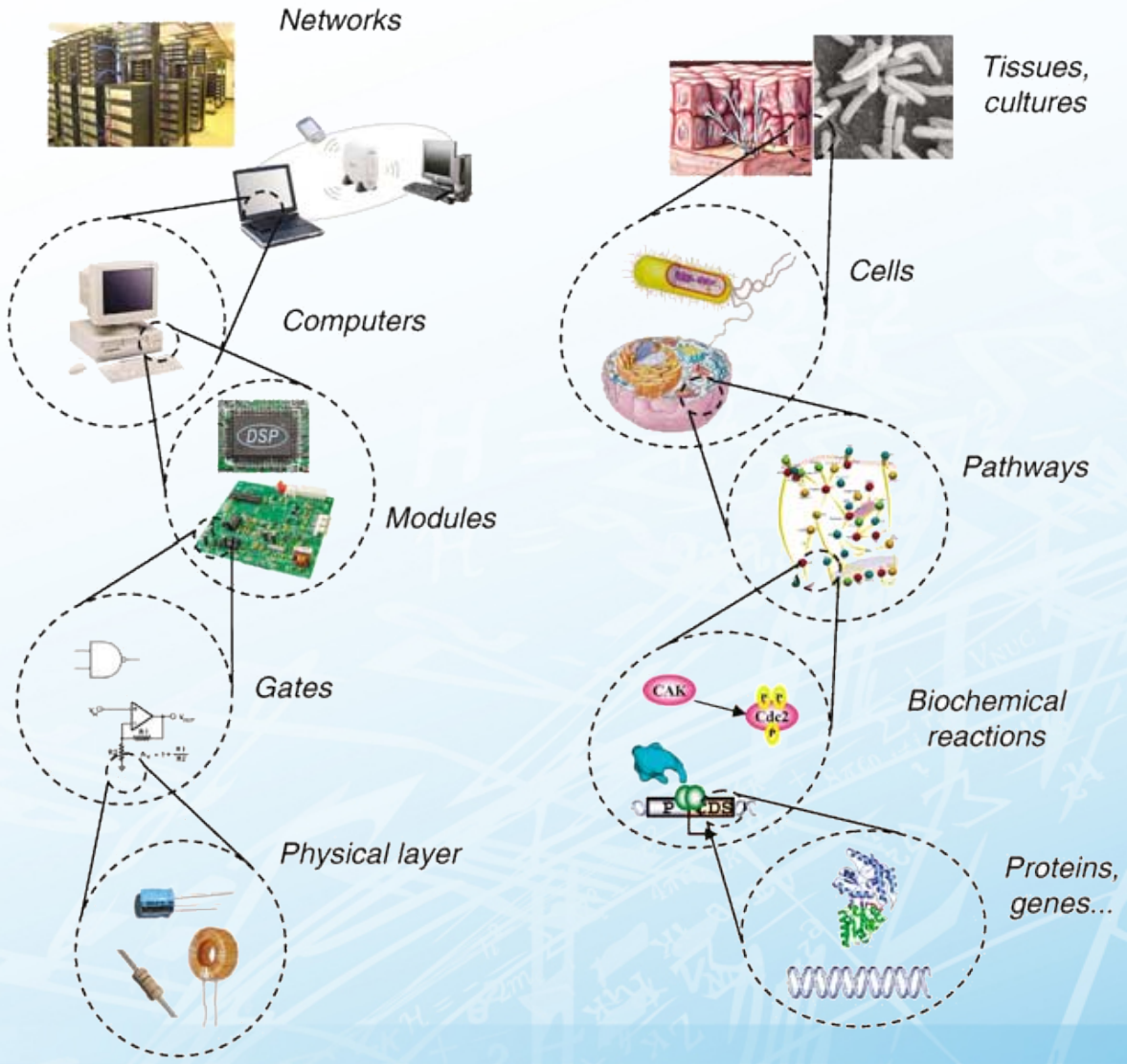
- **Goals**:
  - **To produce new biological systems to carry out some desired and well-defined functions**
  - To develop an engineering technology for the living systems (faster, cheaper, more efficient for a wide usage of engineering principles)
- **How?**
  1. by designing a new protein structure fitting the needs
  2. by generating DNA sequences that fold to the desired structure
  3. by introducing this synthetic DNA in a cell that will process it
- a recent scientific field



Analogy to Computer Engineering

Foundations:

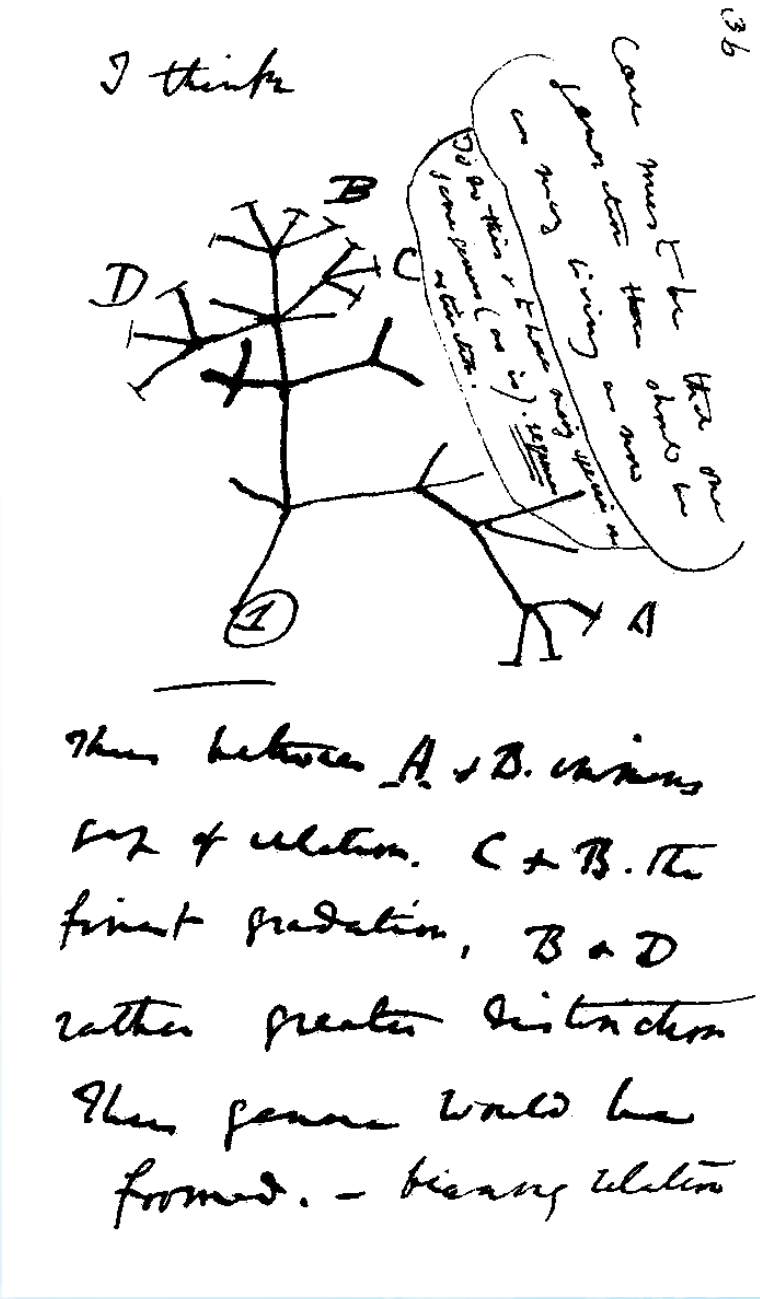
- "Cells are processors"
- "DNA is a (low-level) programming language"



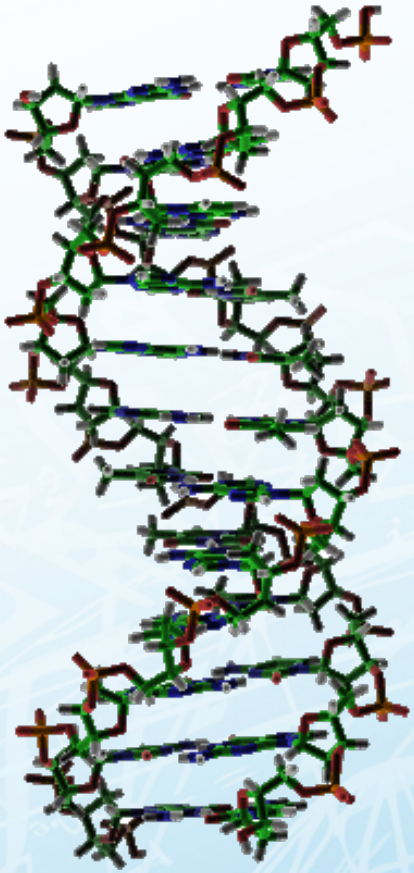
Computing vs Genetics

Digital code	A T G C
Open standards	(e.g. codons)
Modular code	Genes
Error protection	DNA repair
Data compression	Overlapping ORFs
Redundant backups	Double helix, Copy number
Self diagnostics	Apoptosis
Firewalls	Species
Operating systems	Ribsomes

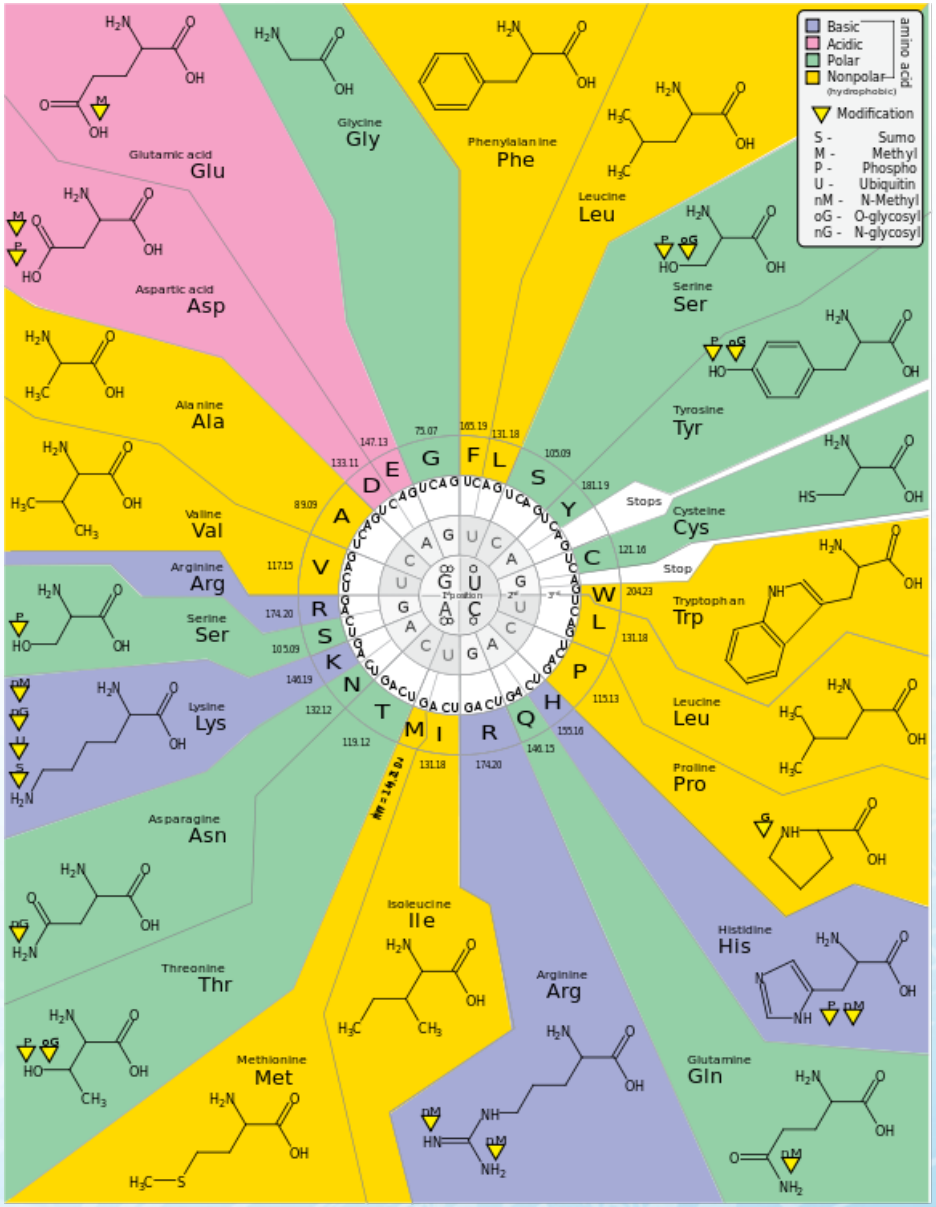
History of Cells - Theory of Evolution (1859)



History of Cells - DNA Structure (1953)



History of Cells - DNA mapping to proteins

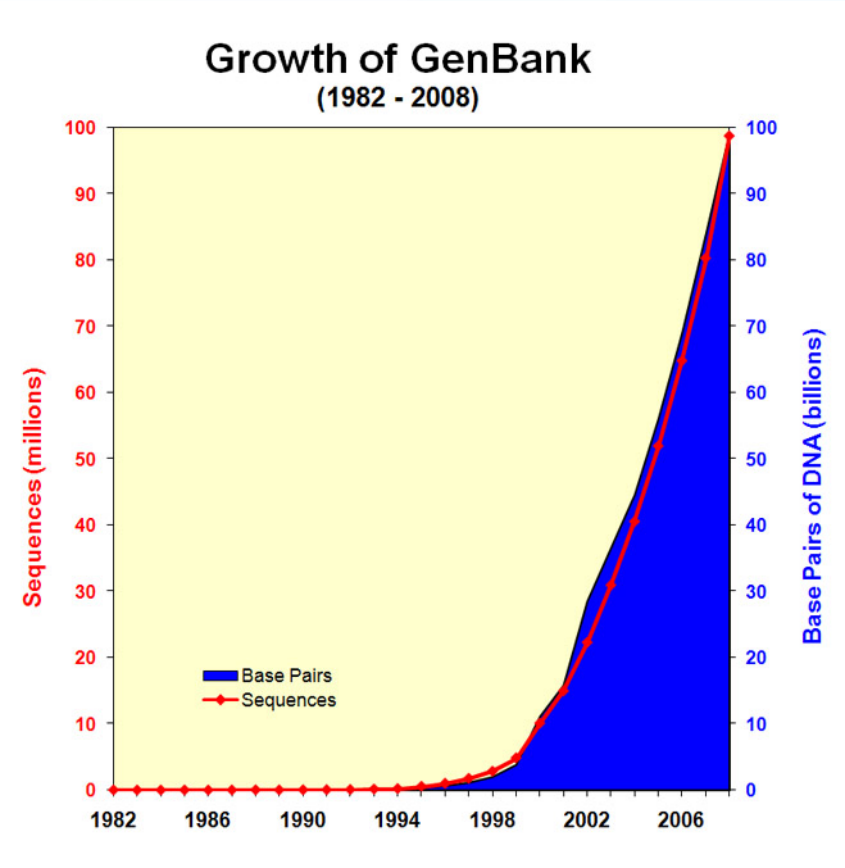


How to work with it?

- READING
- COMPREHENSION
- WRITING

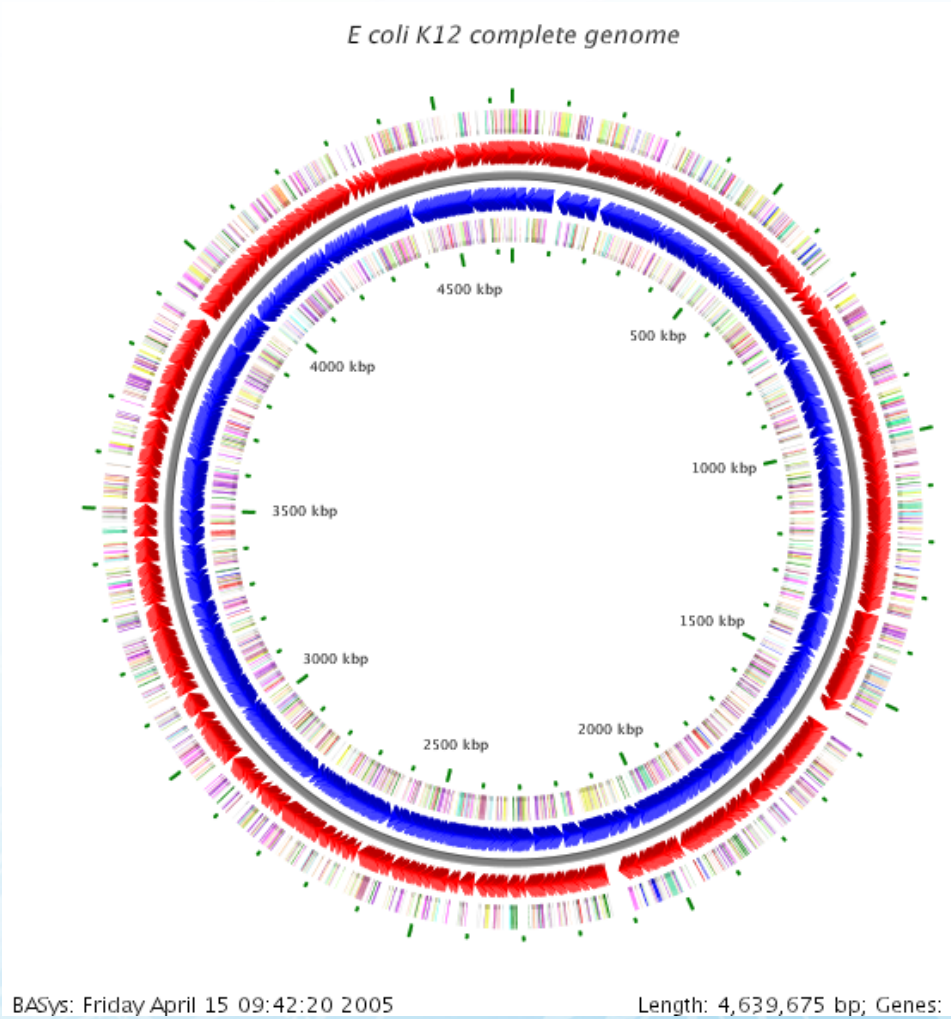
Reading DNA

- in the 70's: dangerous and $< 500 bp/day$
- Few numbers:
  - e-Coli bacteria: 5 Mega base-pairs (bp), compact powerful ~ 5000 genes
  - human: 6000Mbp, ~ 25000 genes
- since 1998, Automated sequencer: $\sim 0.5 Mbp/day$
- impressive growth of genBank (1981 to 2008)



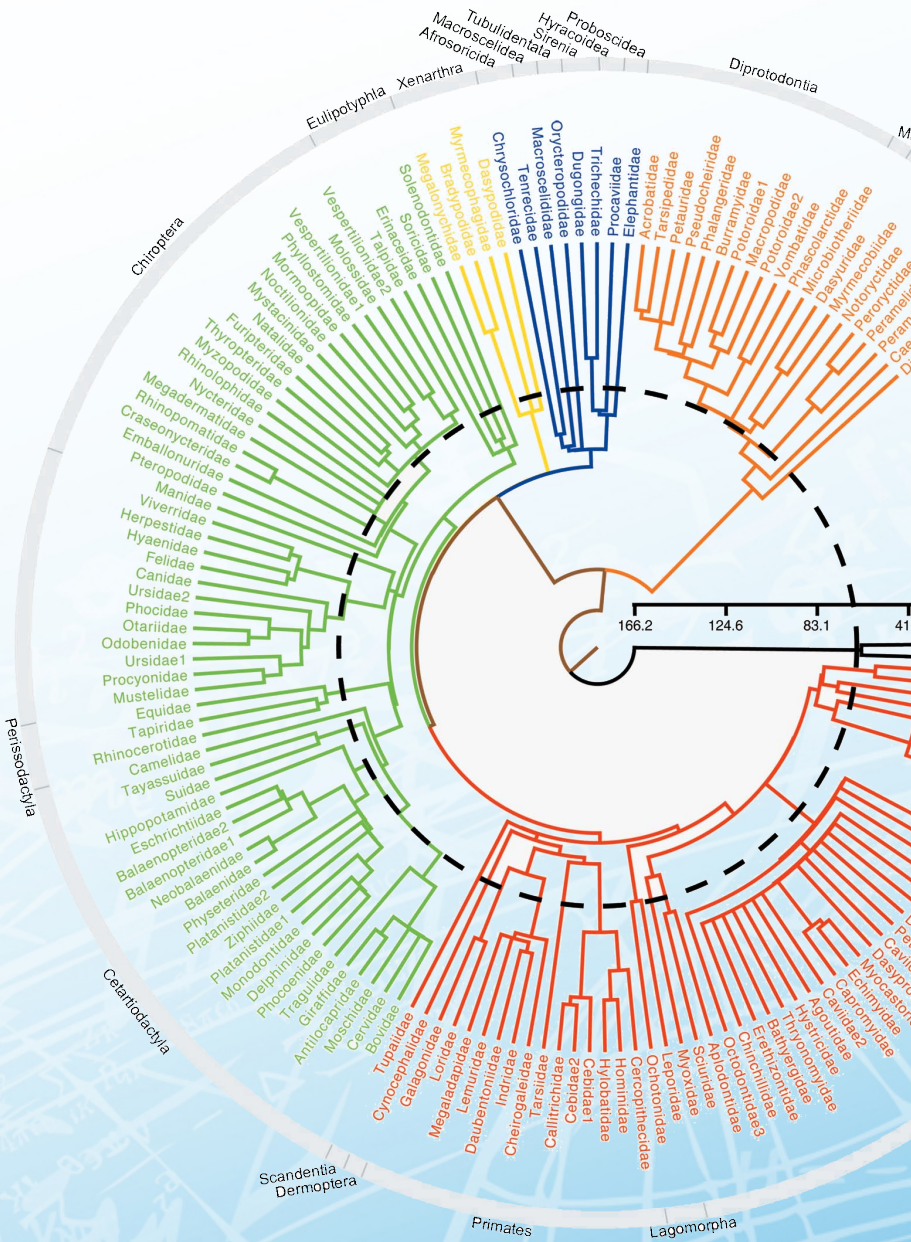
Comprehension of DNA

- They can make **a whole bacterial genomic map**



Comprehension of DNA

- They can sequence species ($>$ 5000 species decoded and listed in **GenBank**)
- They can show **relationship between proteins** of a system (e-Coli)



Writing DNA

- To understand it... need to build it
- early 70's: start writing our "applications" from scratch (identified enzymes to cut, to glue,... DNA)
- **Importance of DNA printers**
  - **Many DNA manufacturing companies** (GeneArt, BlueHeron, Medprobe in Oslo, ...)
  - Really cheap and accessible when: 10 Mbp/hour for $< \$100$

Writing DNA (Editors/Software)

Microsoft Research

Search Microsoft Research

Videos Projects Publications

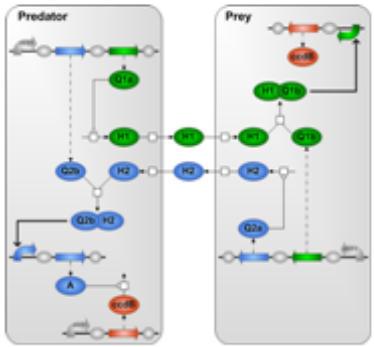
Home | Our Research | Connections | Careers

Find us on Facebook
Follow us on Twitter

Worldwide Labs | Research Areas | Research Groups

> Projects > A Programming Langauge Genetic Engineering of Living Cells

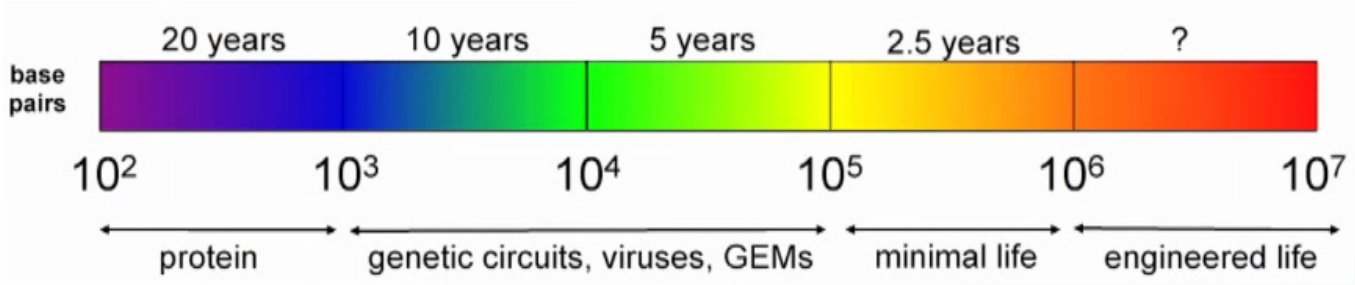
A Programming Language for Genetic Engineering of Living Cells

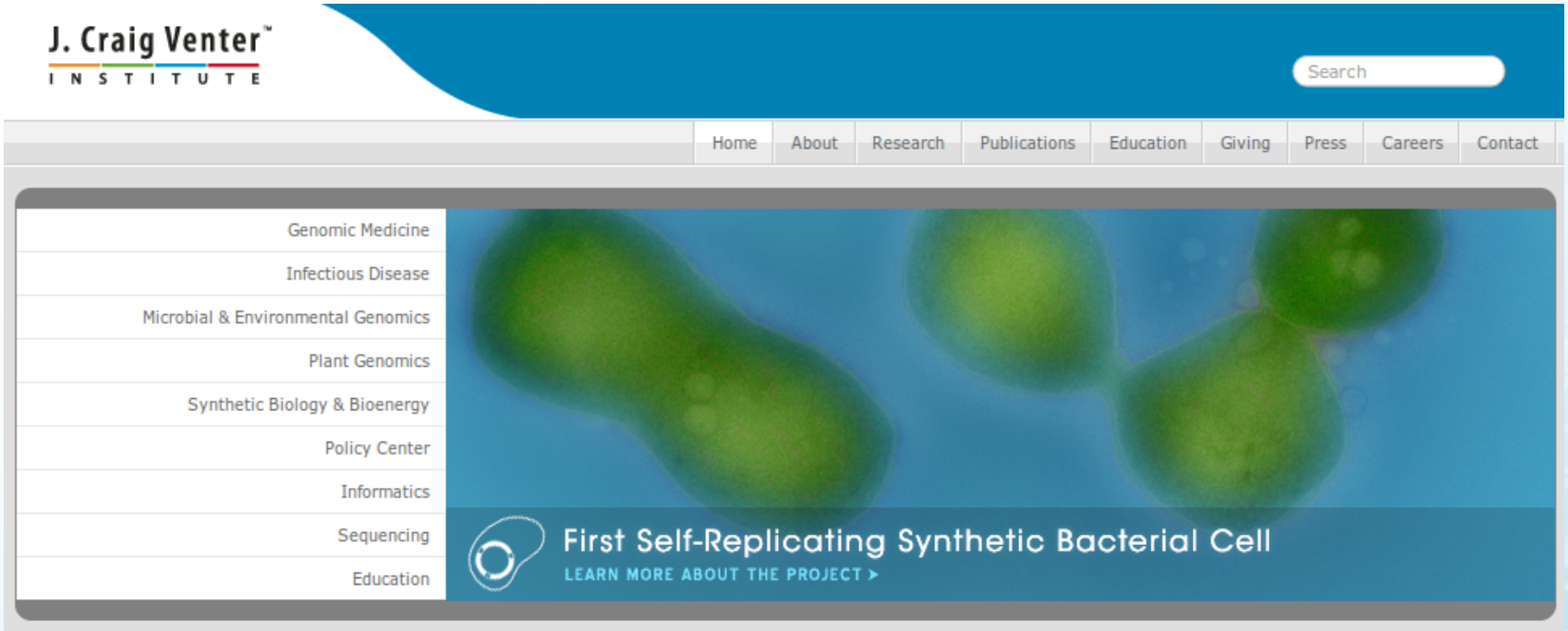


Synthetic biology aims at producing novel biological systems to carry out some desired and well-defined functions. An ultimate dream is to design these systems at a high level of abstraction using engineering-based tools and programming languages, press a button, and have the design translated to DNA sequences that can be synthesised and put to work in living cells.

- Development of true genetic programming languages
- debugging "in vitro"
- **"An ultimate dream is to design these systems press a button, and have the design translated to DNA sequences that can be synthesised and put to work in living cells. "**

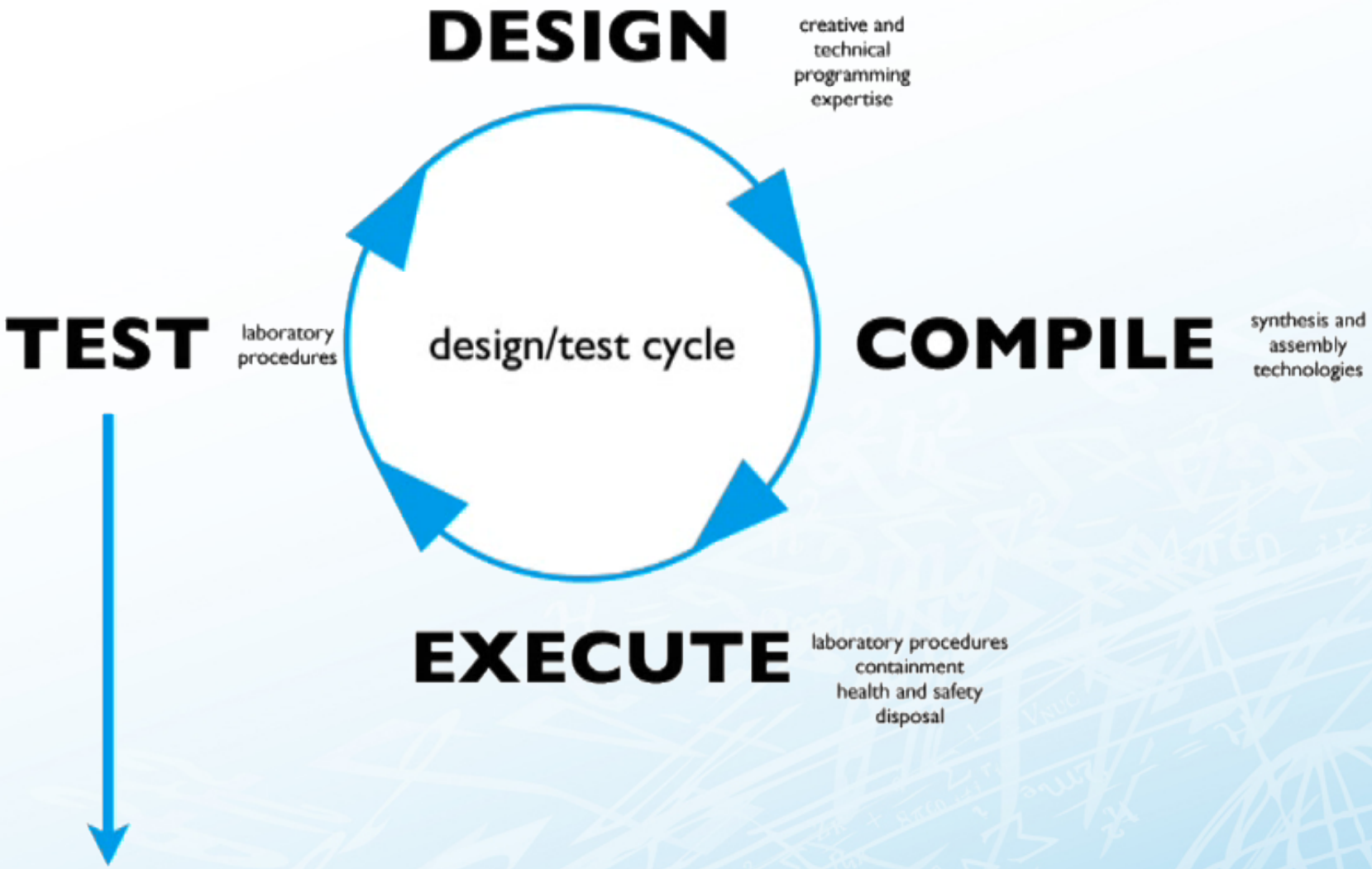
The Human genome coming soon...





- Quick development of synthetic biology, the next IT industry
- The problem today: **not HOW** to engineer something, but **for WHAT**! (iGem, Fold.It, DIYBio)

How to contribute?

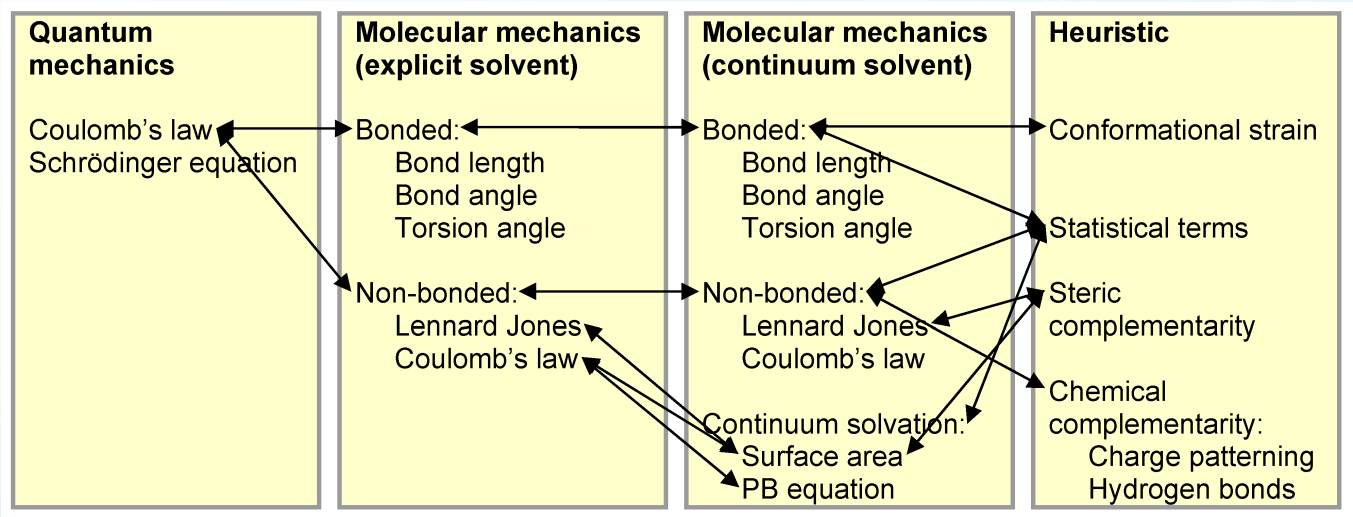


Outline

- What is Synthetic Biology?
- **How to computationally design "improved" proteins**
- What are the challenges for Computational chemistry?

Computational protein design algorithms

- **Predicting Structure from DNA Sequence?** (from amino acids native conformation)
  Not so easy!!! .The chain can fold up in too many ways! (**www.Fold.It**)
- **models of protein energetics** to evaluate how mutations affect the structure and function of proteins
  - **molecular mechanics**
  - **statistical potential** (i.e. known-proteins in databank)
  - and other empirical terms



- QM/MM

Outline

- What is Synthetic Biology?
- How to computationally design "improved" proteins
- **What are the challenges for Computational chemistry?**

Challenges for Computational chemistry: improving designed enzymes

- **Structural characterization**: to determine success of active site geometry (e.g. Crystal structures)
- **Directed evolution** to identify DNA sequence missing from the design that increases activity
- **Molecular Force Fields**: for better control of side chain conformation, loop conformation, transition state binding orientation.
- **Molecular dynamics with explicit solvent**: for assessing conformational stability and population/orientation of transition state in the active site.
- **quantum chemistry** and **QM/MM hybrid methods** in directly determining the magnitude of transition state energy barriers in the context of the designed active site and ultimately in the context of the designed protein

Summary and Conclusion

- Synthetic Biology: the next IT industry
- Needs some improvements of the technology to make it accessible to the public
- Needs some improvements in identifying and designing "what to improve": needs **Computational chemistry**

Thank you for your attention!