Integration of Biological Knowledge

New Developments in Biosensing Technologies

Ricardo Dias

rpdias@ciencias.ulisboa.pt

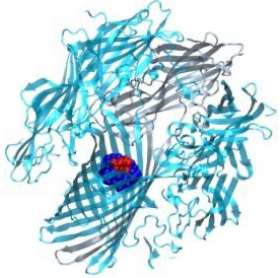
Bugworkers

Ciências ULisboa Faculdade de Ciências da Universidade de Lisboa

BioISI Biosystems and Integrative Sciences Institute

Genomics Bringing Knowledge to you: anyTime anyWhere

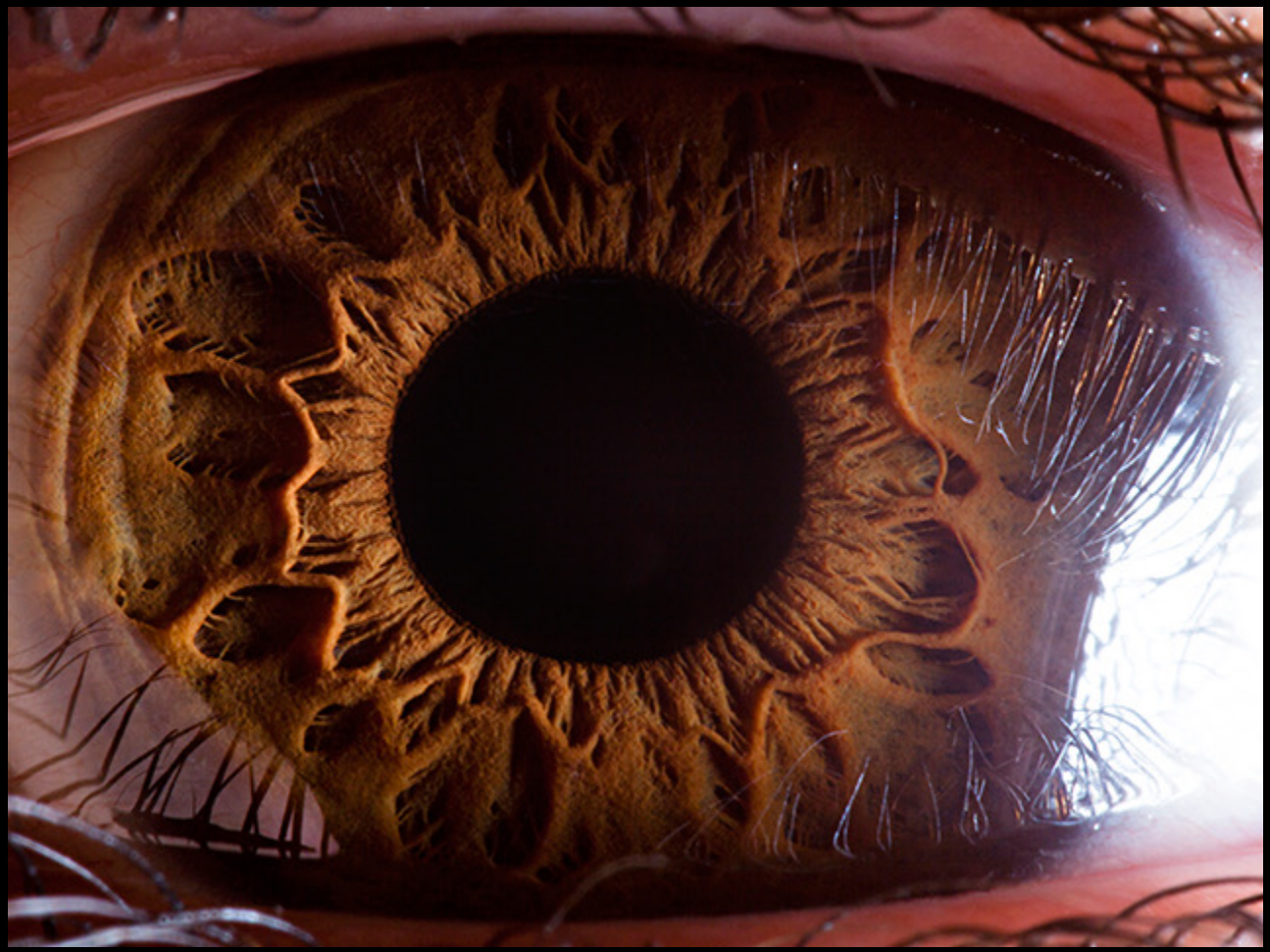
Today



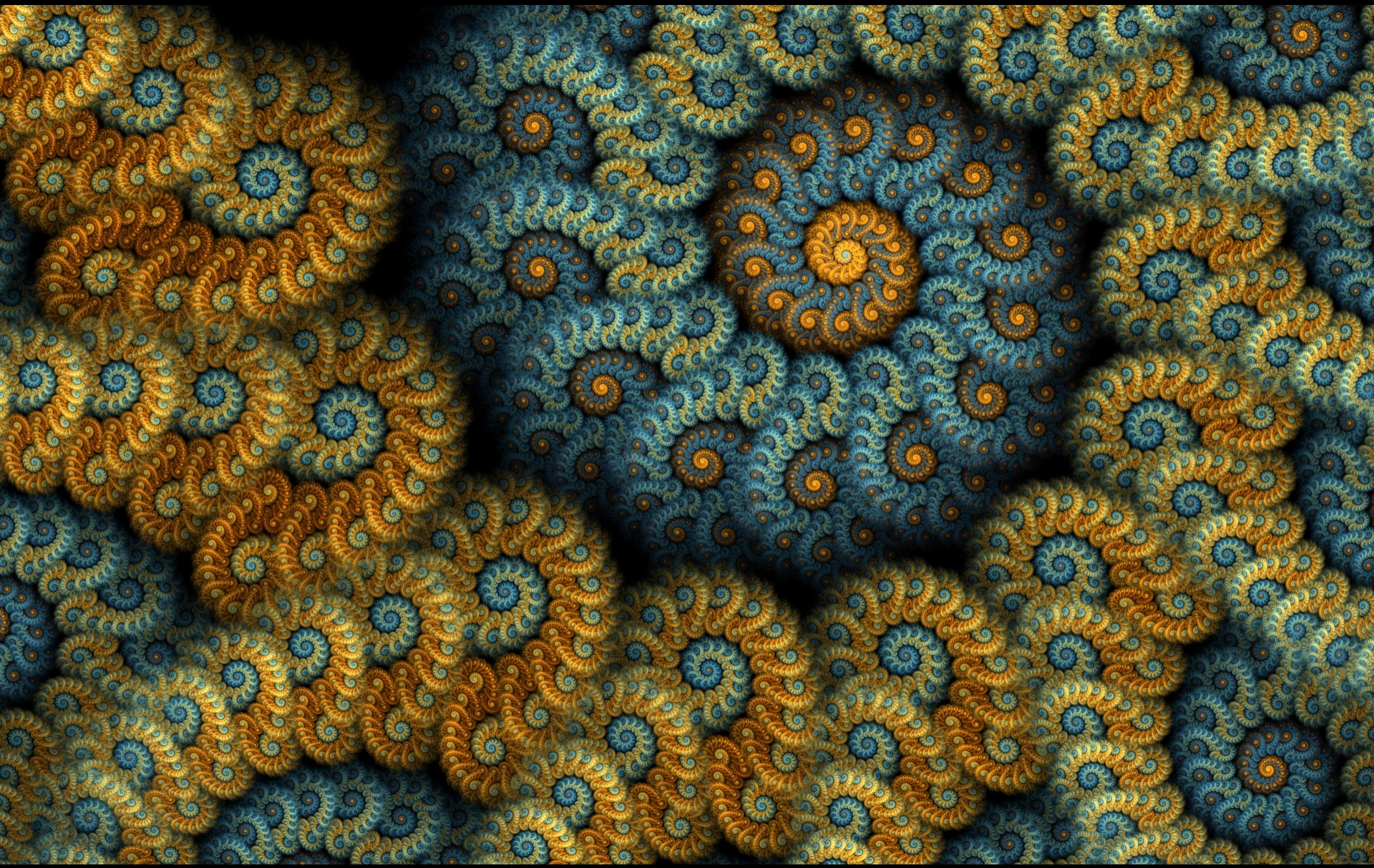
- Intro
- Web of Life - integration of Biological Knowledge
- The need of bio-cyber-physical sensing devices
- Why nanopores as bio-cyber-physical signal transducers ?
- Challenges for producing large scale Knowledge in real-time ?

Biological Knowledge
Vs
Resilience


© William C Graham



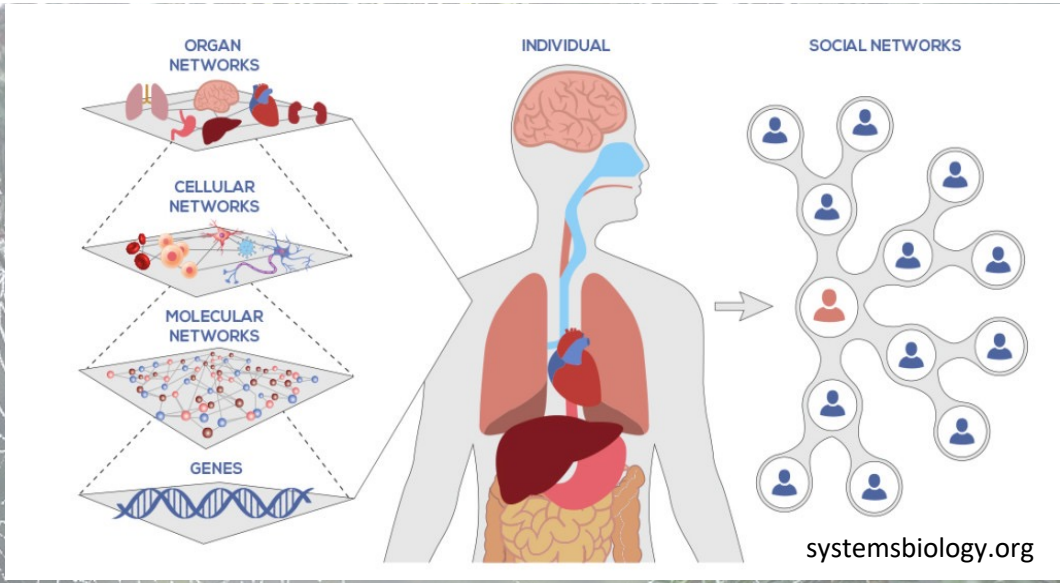

©Heather Pierce



New Developments in Genomics Technologies

Web of life

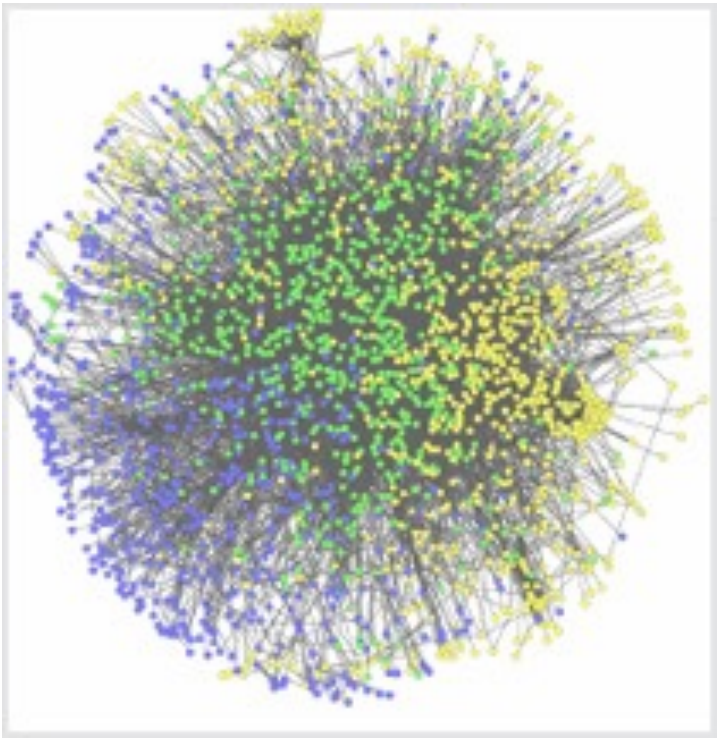
- Multi-component
- Multi-specie
- Multi-Interaction
- Multi-level
- Multi-network



- Multi-dimensional
  - Attribute
  - Time
  - Space

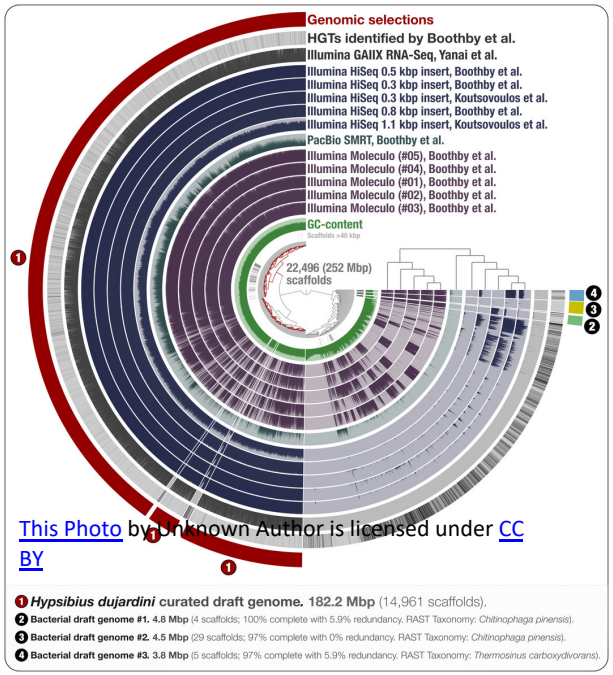
Integrated Multi level Systematic Approach for Knowledge Production

- **Biome characterization**
  - System
  - Dimensional
  - Component
- **Interactions**
  - Space and time
  - Resolution at Single -unit; -Component



New Emergent Concept on Genomic

- **Genome is different:**
  - From cell-to-cell
  - Along the time
  - Multi Component Information
  - Source of Cumulative information



Non-Biological Interconnected Web

The need of bio-cyber-physical sensing devices

Digitization of biological components and systems

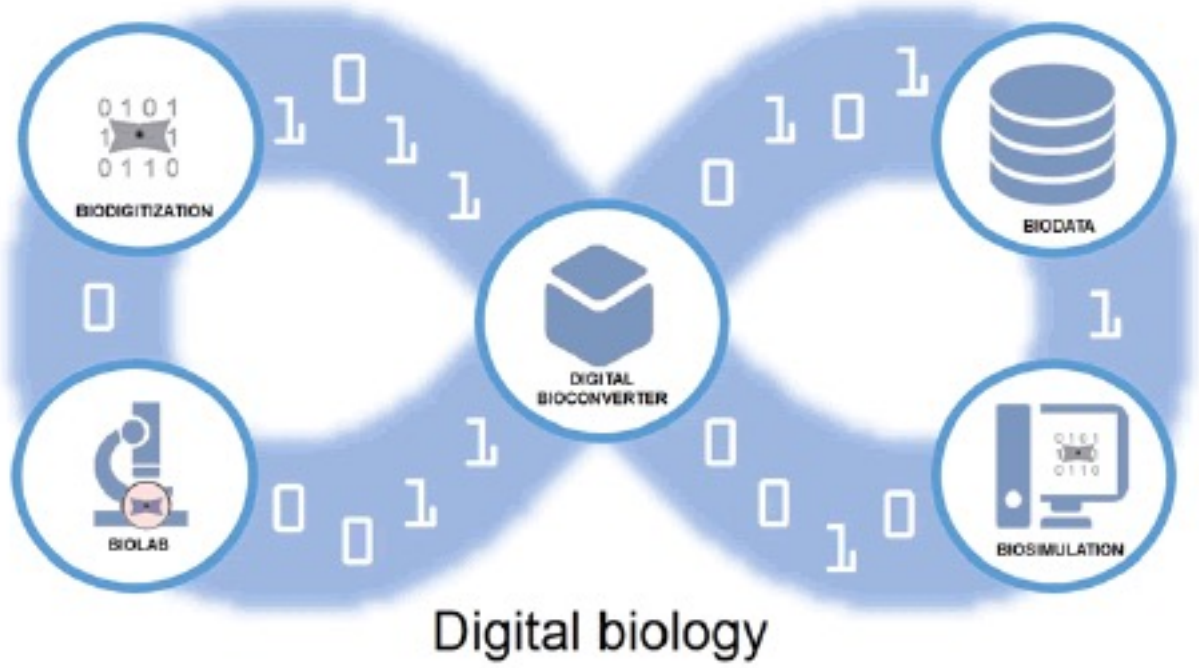
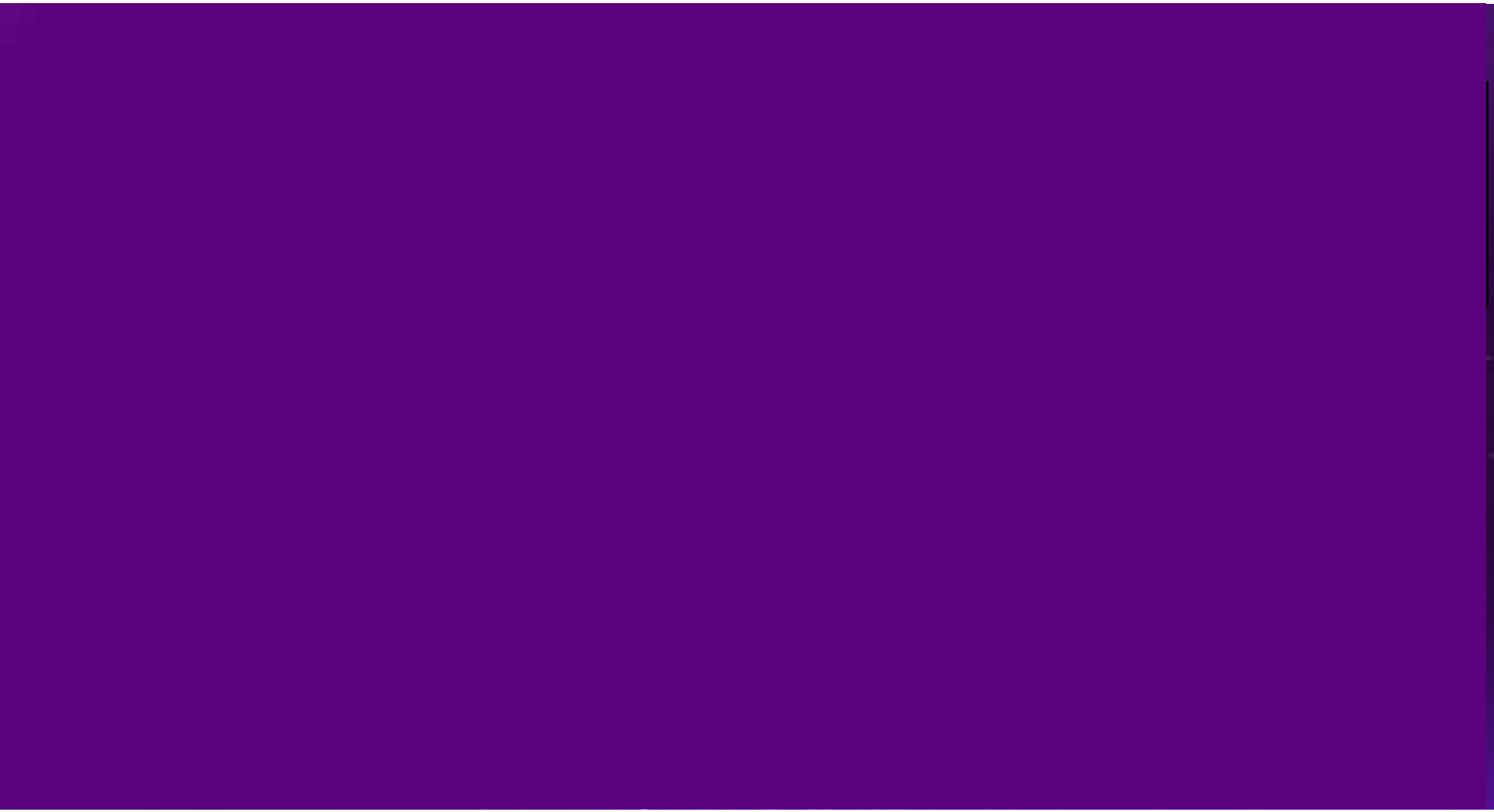


Figure 1: The digital biology loop, with the digital bioconverter, a tool for instantiating data driven biosimulations into biomolecules and cells for analysis at the lab bench, digitization and further simulation and analysis.

STOmics

SpaTial Enhanced REsolution Omic-Sequencing



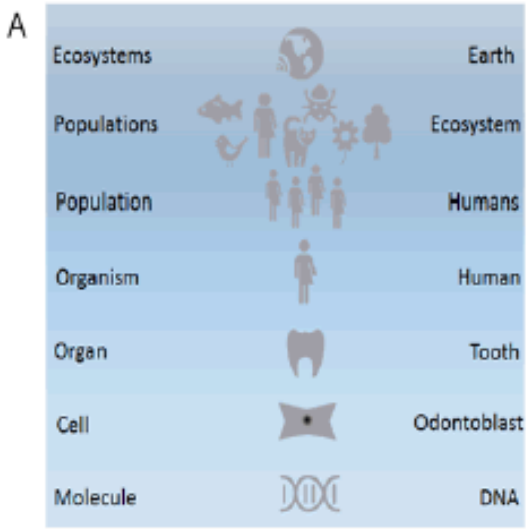
STOmics

3D construction

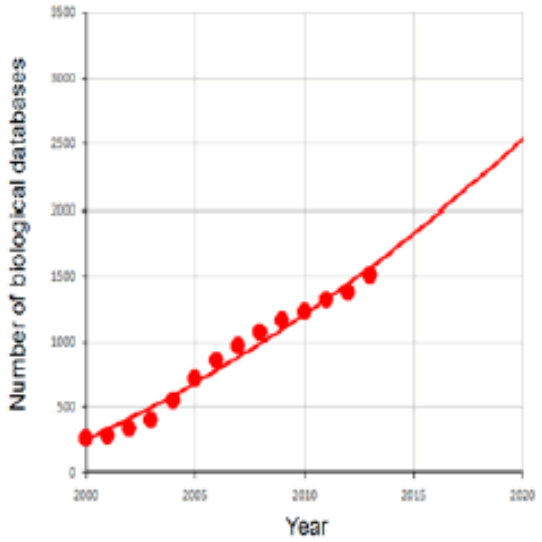
Multi Component

Multi Organism / Specie

Multi -Level -Time and -Space



Levels of biological organization



Growth of biological databases

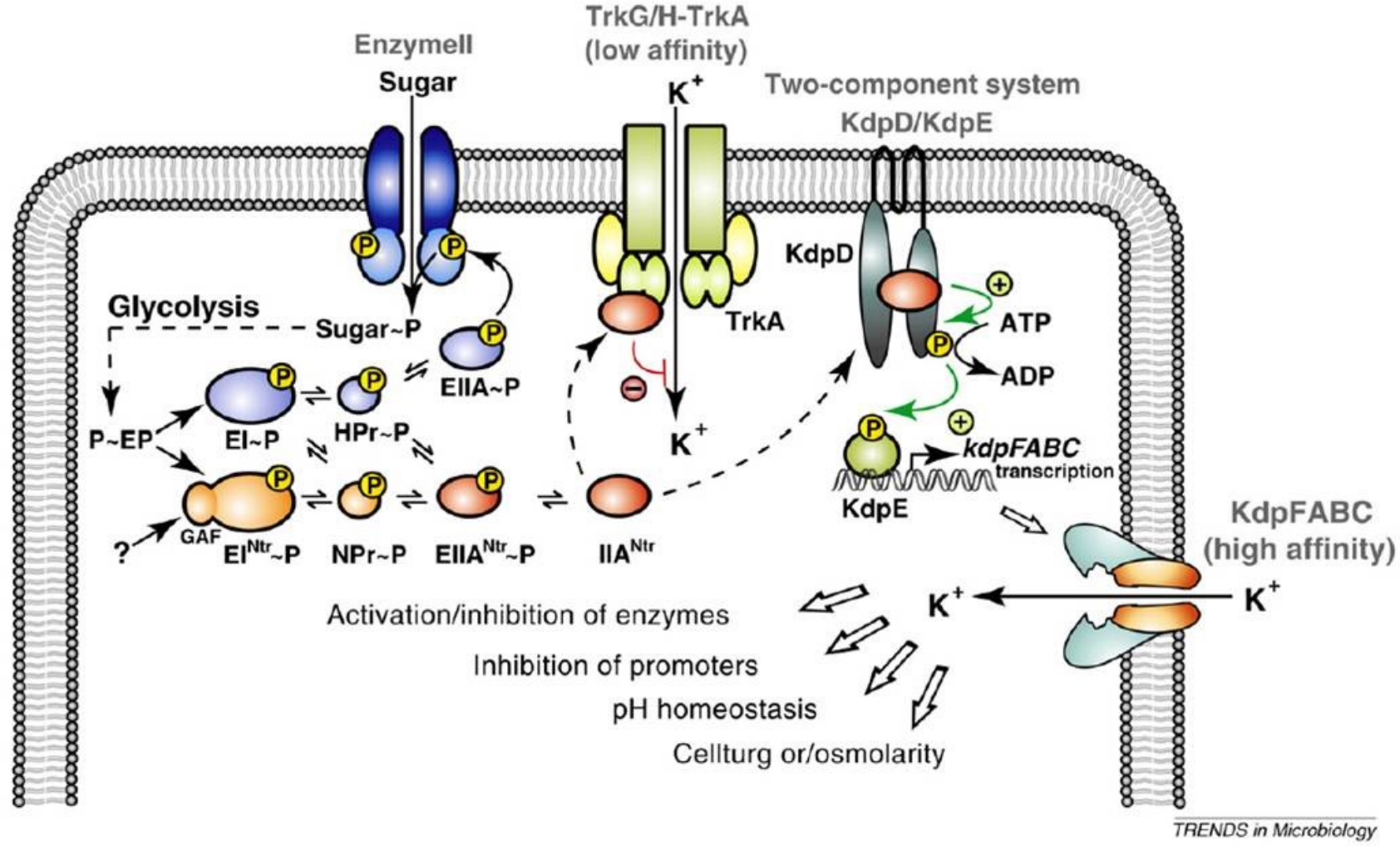
Sources: ONT_2017; Journal of Bioengineering & Biomedical Science ISSN: 2155-9538

Technology requirements

- Real-time data acquisition
- Multi-site
- Onsite
- Multi component
- Resolution at single (molecule) level

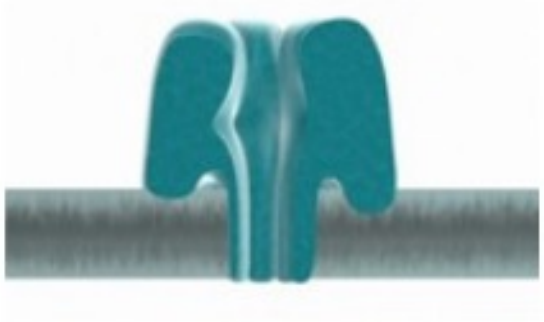
Why nanopore as bio-cyber-physical signal transducers ?

Environmental Sensing and Signal Transduction

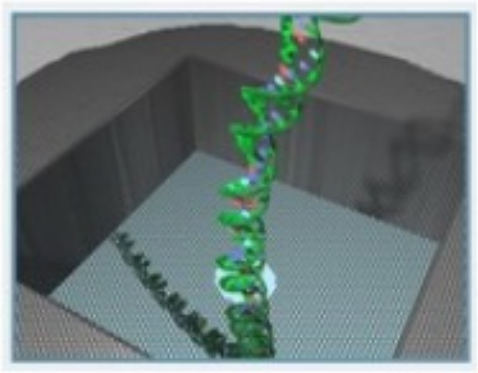


Nanopore used for molecular sensing

A nanopore: a nano-scale hole.



- Biological: a pore-forming protein (e.g. α-Hemolysin) in a membrane (e.g. lipid bilayer)

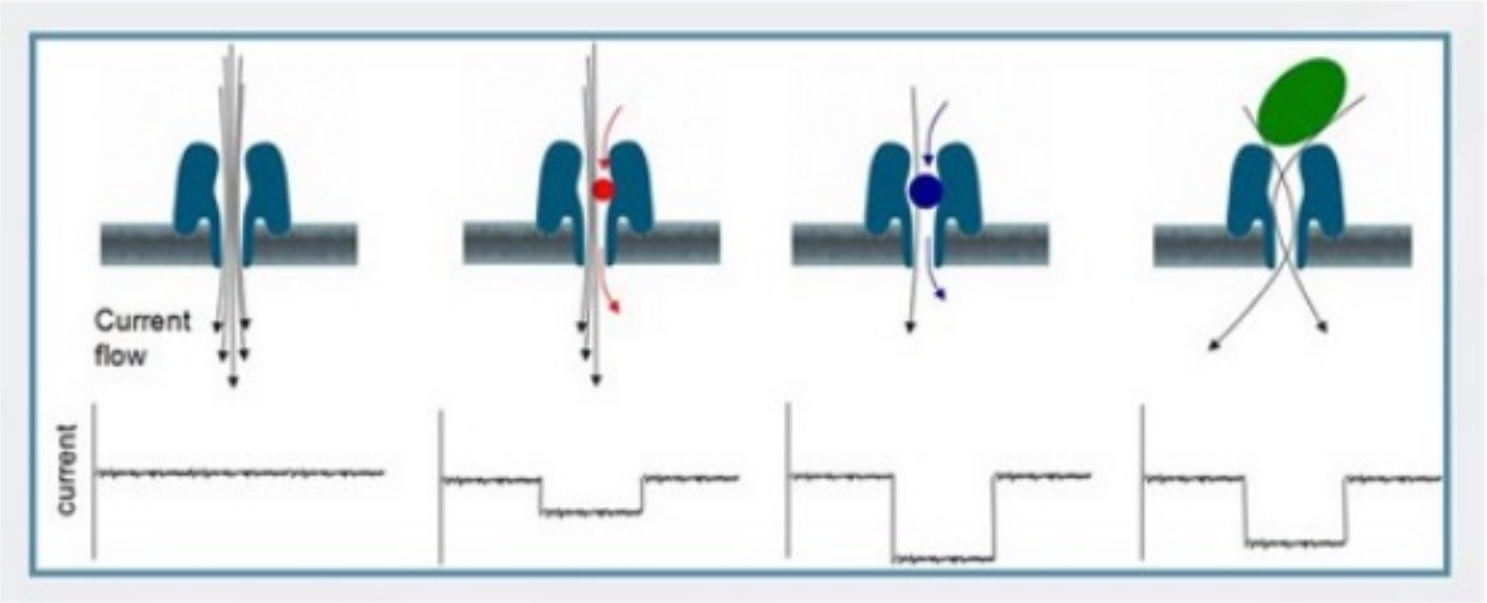


- Solid-state: in synthetic materials (e.g. silicon nitride or graphene)

- Hybrid: formed by a pore-forming protein set in synthetic material

Nanopore sensing

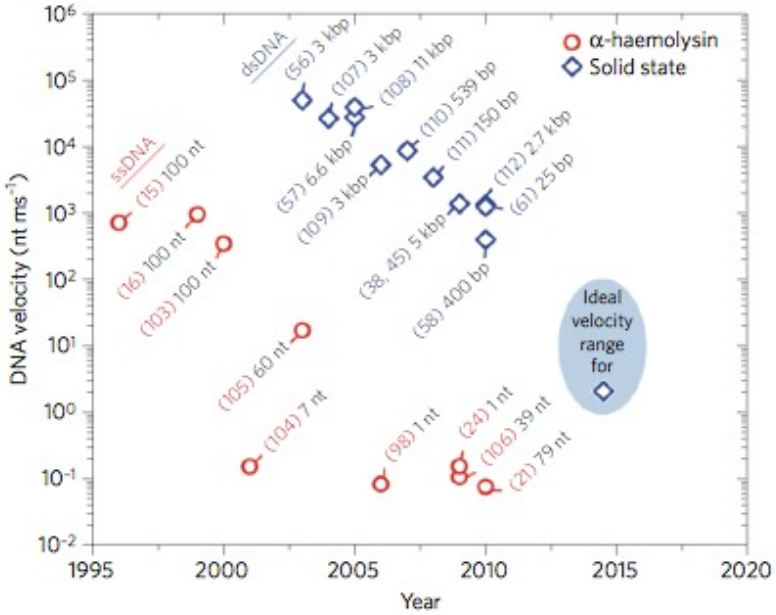
Ionic current passed through membrane by setting a voltage across the membrane.

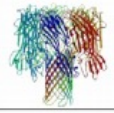


- Disruption in current detected when analyte passes through the pore or near its aperture.

- Characteristic disruption indentifies the molecule in question.

Trends in nanopore-based sensing
Biological nanopores

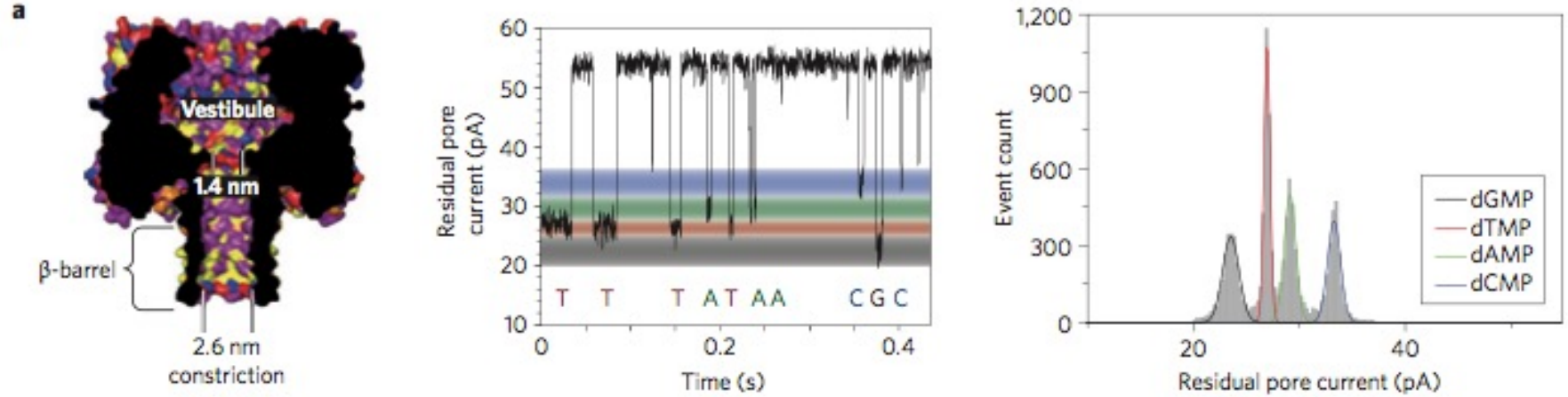


Biological Nanopore	Structure: Side View	Structure: Top View	Critical Dimension	Types of Analytes	References
α-HL			1.4 nm[108]	small molecules, RNA, ssDNA, dsDNA, proteins	25, 28, 33, 37, 38, 41, 45, 52, 128, 129, 134-143, 148-150, 152, 153, 157
OmpG			1.3 nm[118]	small molecules, proteins	43, 44, 125, 126
MspA			1.2 nm[110]	ssDNA, dsDNA	31, 51, 120-122, 154-156
AeL			1.0 nm[111]	ssDNA, proteins	39, 46, 49, 127, 144, 145, 151
Phi29 Motor			3.6 nm[114]	small molecules, ssDNA, dsDNA, proteins	48, 50, 130, 131, 146
ClyA			3.3 nm[116]	dsDNA, proteins	40, 42, 47, 127, 132, 133, 147

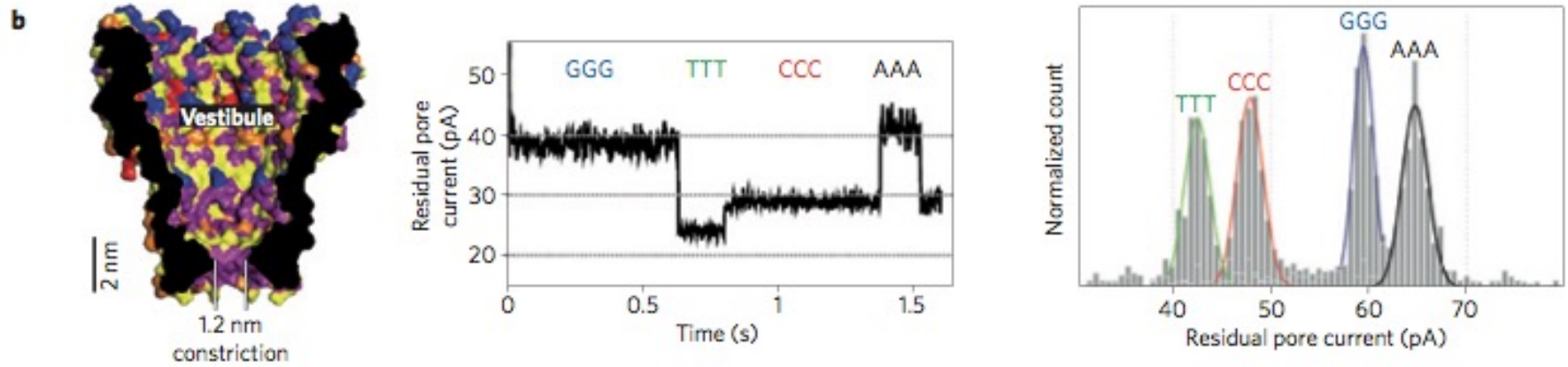
- **α-HL. Staphylococcal alpha hemolysin (α-HL)**
- ***Mycobacterium smegmatis* porin A (MspA)**
- Aerolysin pore from *Aeromonas hydrophila* (AeL)
- The bacteriophage phi29 motor
- Cytolysin A from *Salmonella typhi* (ClyA)
- Outer membrane protein G (OmpG)

Anal. Chem. 2017, 89, 157−188;
Nature Nanotechnology volume 6, pages 615–624 (2011)

Structural cross-section of **Staphylococcal** α-haemolysin



Structural cross-section of ***Mycobacterium smegmatis*** **porin A** MspA.



Nature Nanotechnology volume 6, pages 615–624 (2011)

Nucleic acid sensing

- **Detection**
- **Characterization**
  - **Sequence**
  - **Chemical Composition**
  - **Size**

How can be produced near real-time molecular sensing (sequencing)?

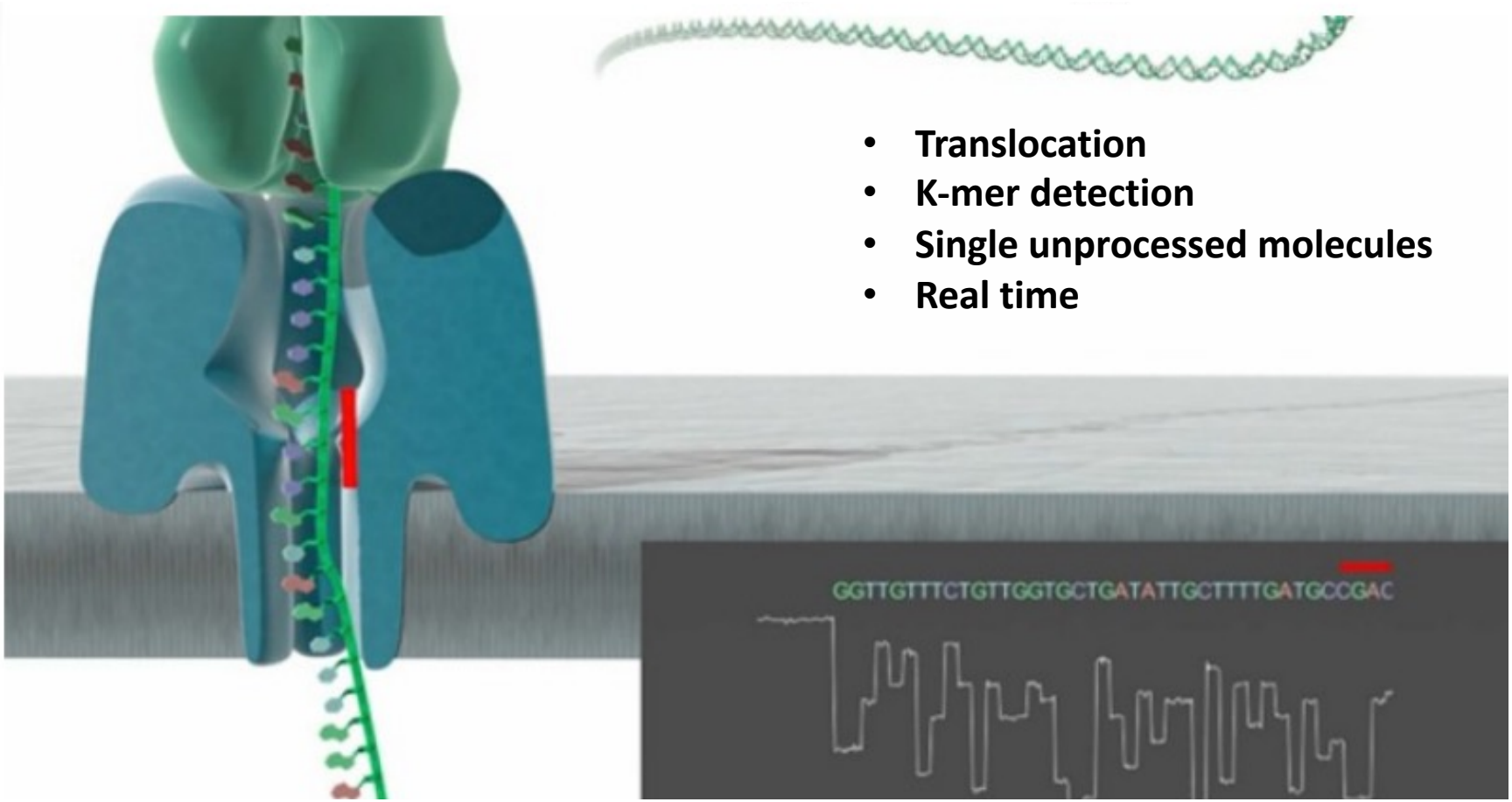
- **Translocation Speed > 400 events (nt) /s**
- **Application-Specific Integrated Circuit (ASIC) real time signal acquisition and processing**
- **Real time data computation**

Types of Nanopore sequencing

- Strand sequencing:
  - Sequencing in real-time as the intact DNA polymer passes through the nanopore.

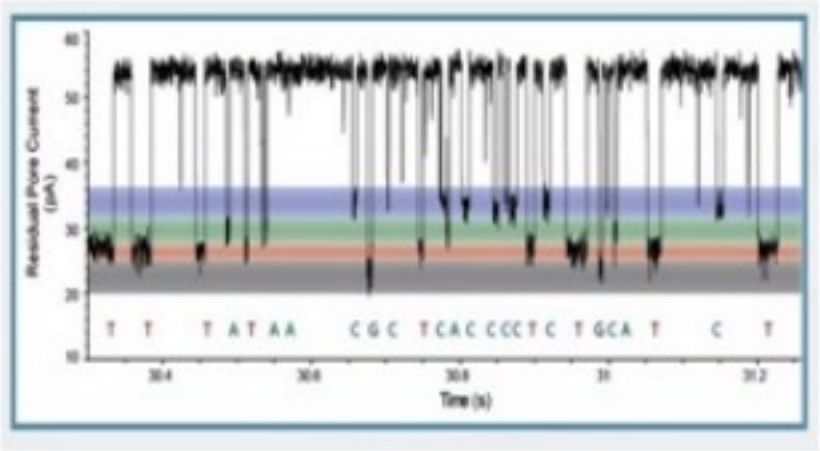
- Exonuclease sequencing:
  - Individual nucleotides pass through the nanopore by the aid of processive exonuclease.

1- Strand sequencing

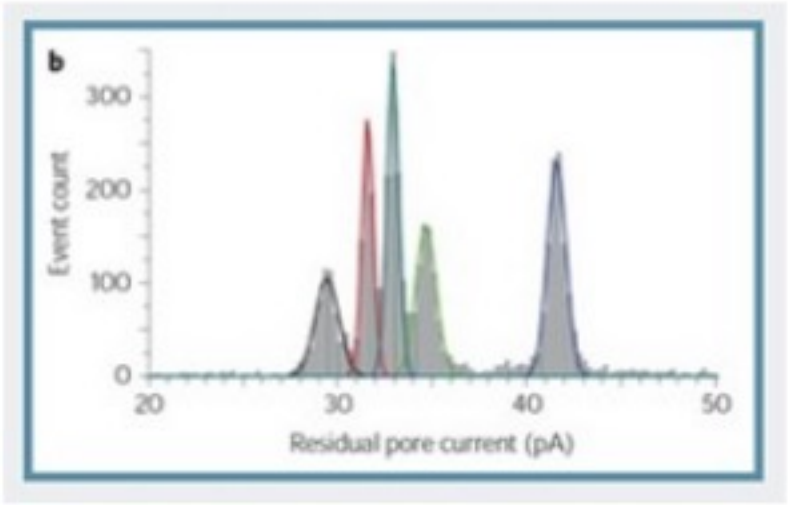


- Translocation
- K-mer detection
- Single unprocessed molecules
- Real time

Detection of non-canonical bases (NCNB) and modifications (MNB)

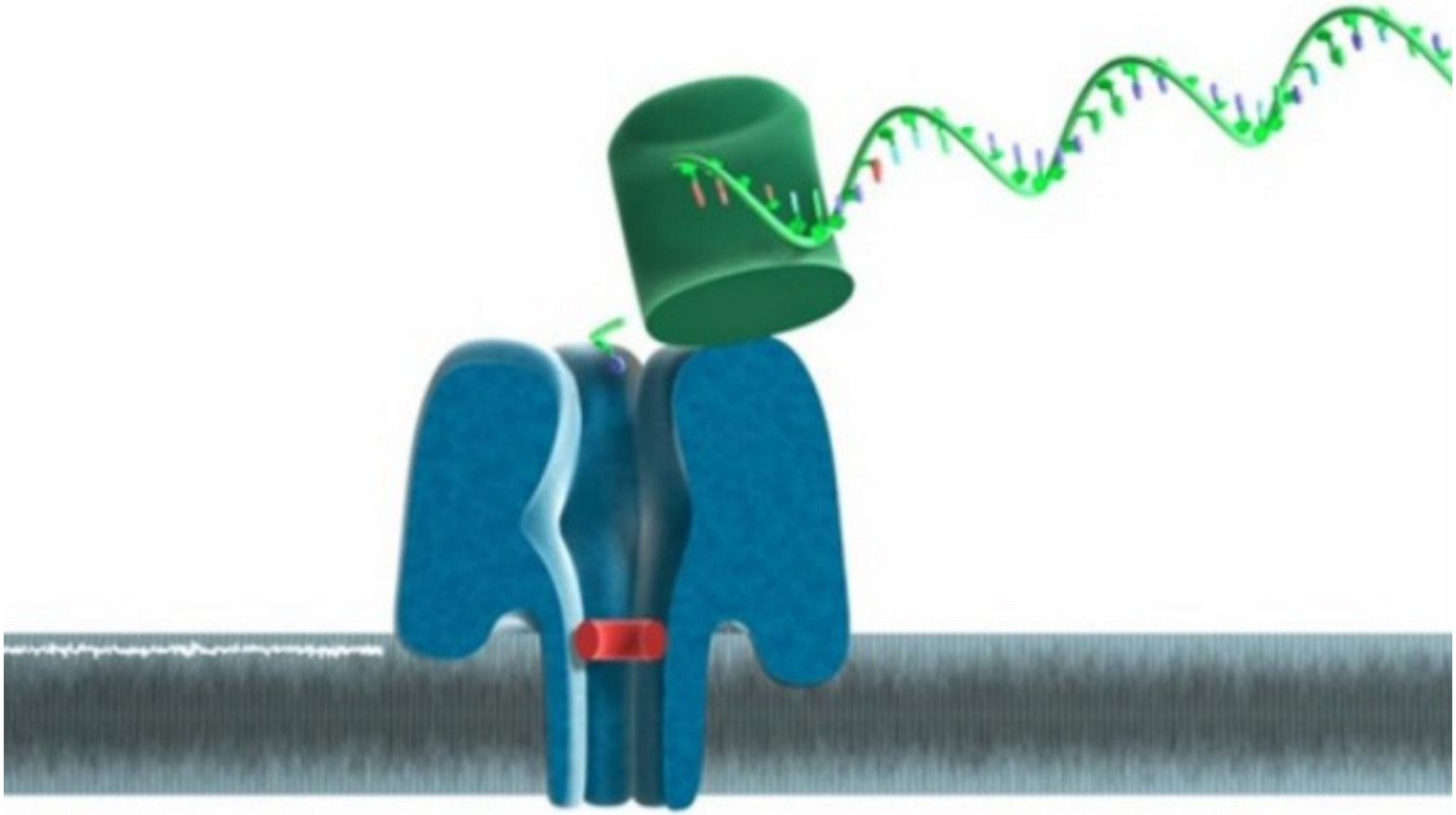


Four different magnitudes of disruption which can be classified as C, G, A or T



Modified base, e.g. methylated cytosine, can be directly distinguished from the four standard bases

2- Exonuclease Sequencing



Snapshot from movie at http://www.nanoporetech.com

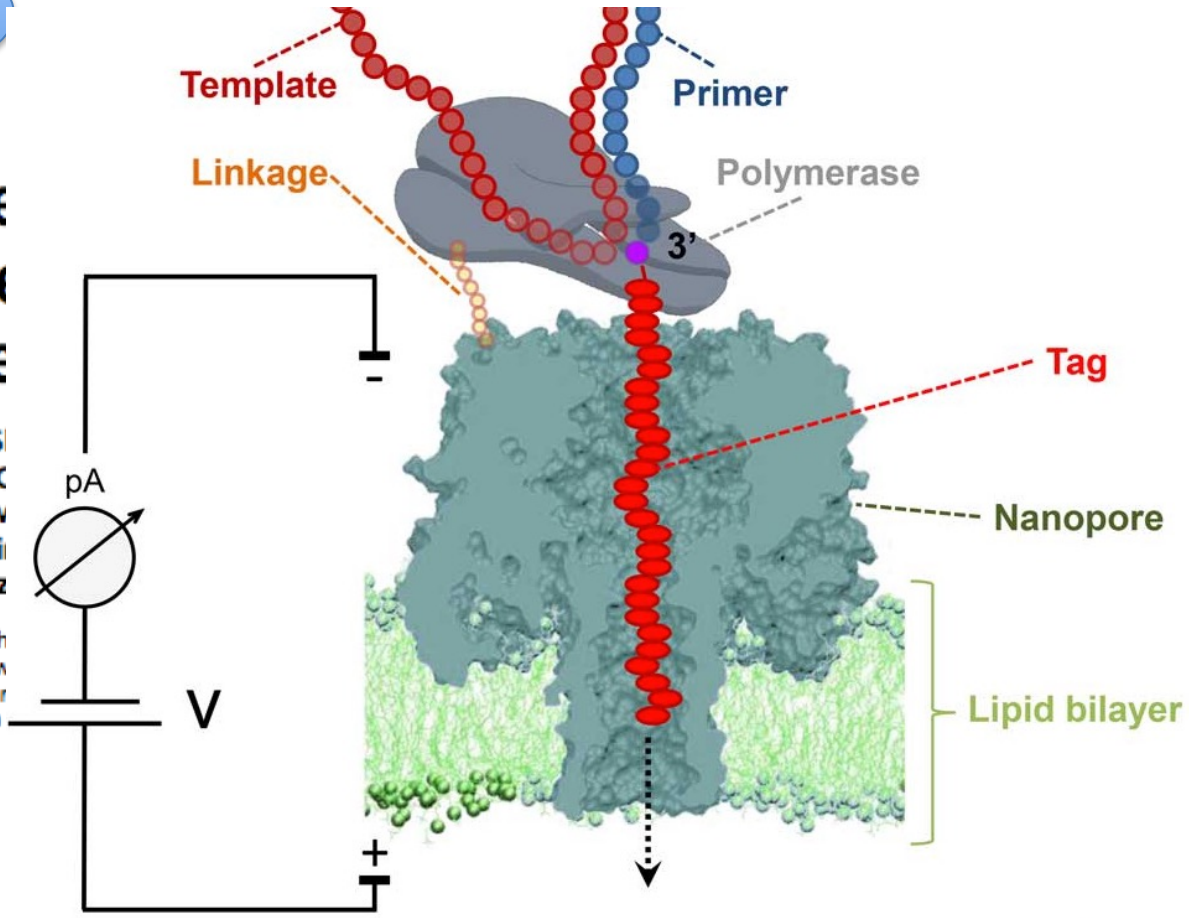
DNA sequencing by synthesis using polymer-tagged nucleotides

In Development TRL 6

Real-time
by synthe
nanopore

Carl W. Fuller[a,b,c], S
Michael Dorwart[c], C
Anne Aguirre[c], Edw
Cleoma Arnold[c], Mi
John J. Kasianowicz

[a]Center for Genome Tech
Columbia University, New
02115; [e]Semiconductor ar
Pharmacology, Columbia



ncing
n a

ew Trans[c],
J. Russo[a,b],

ineering,
, Boston, MA
Department of

3-5238;

Protein Sensing

- Direct
  - Characterization 5D
- Indirect
  - Based on the principles of DNA origami
  - DNA aptamer
  - ag-mab (modified ELISA)

Nanopore sensors enter digital age with DNA barcodes for multiplexed protein identification

MENU

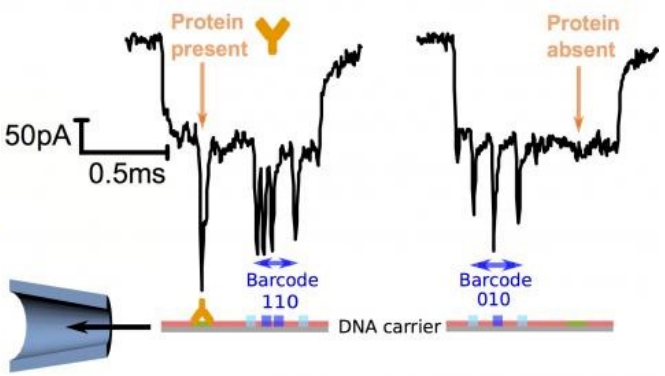
nature nanotechnology

Article | Published: 04 April 2016

Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores

Nicholas A. W. Bell & Ulrich F. Keyser

Nature Nanotechnology **11**, 645–651 (2016) | Download Ci



Nature Nanotechn

Selective single molecule nanopore sensing of proteins using DNA aptamer-functionalised gold nanoparticles

Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

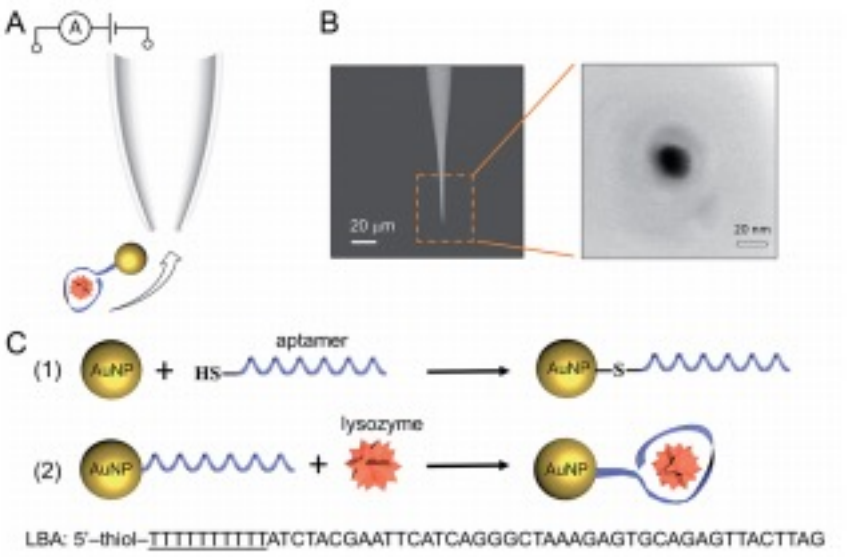
View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2017, 8, 3905

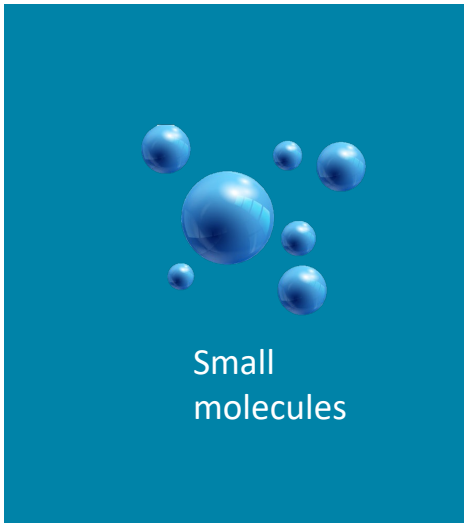
Selective single molecule nanopore sensing of proteins using DNA aptamer-functionalised gold nanoparticles†

Xiaoyan Lin, Aleksandar P

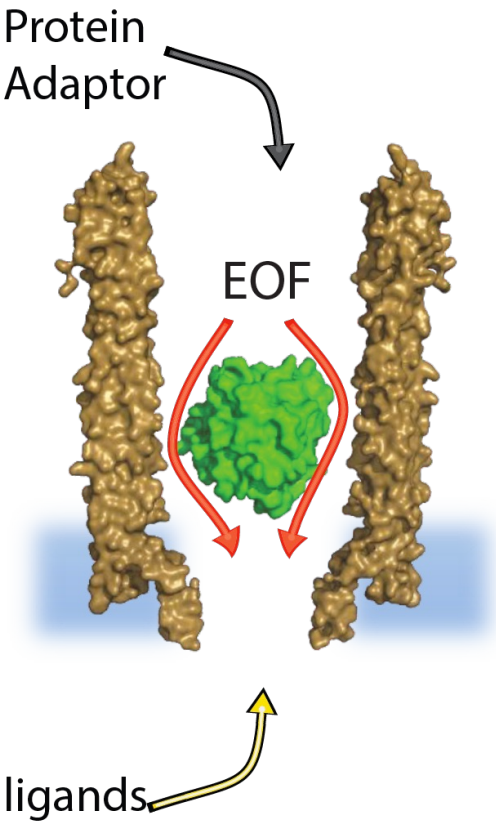


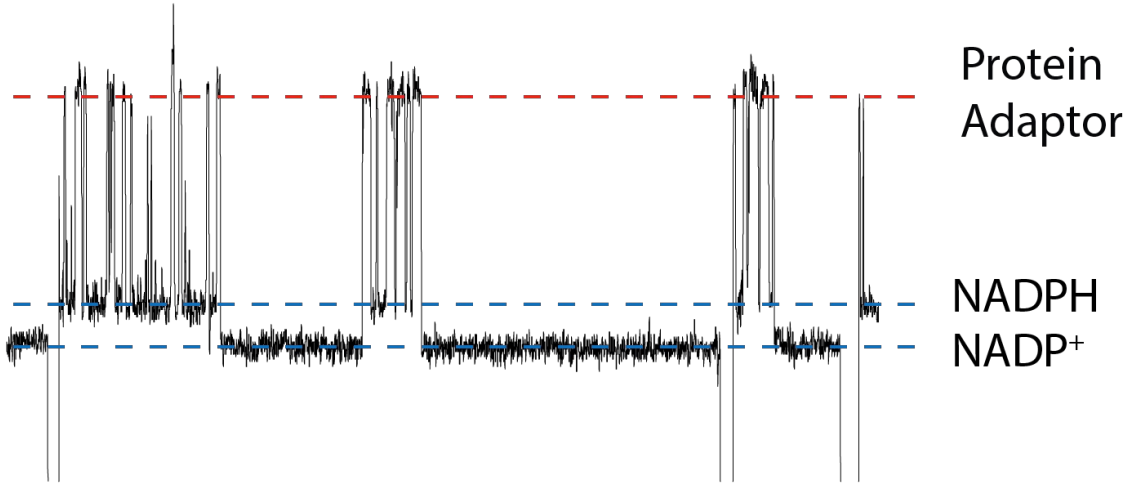
Small Molecule Sensing

- Direct
- Indirect
  - Composed blocage

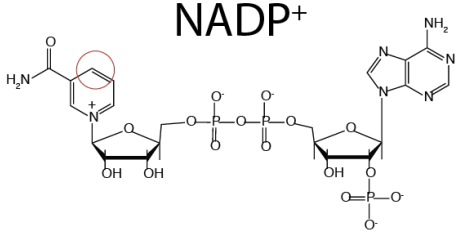


Direct Small Molecule Sensing

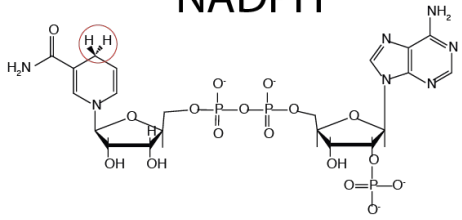




NADP+



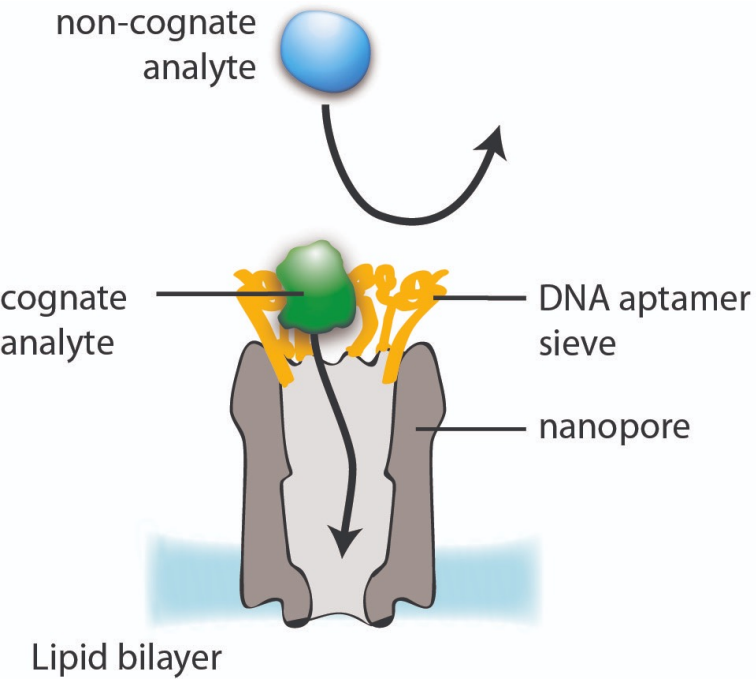
NADPH

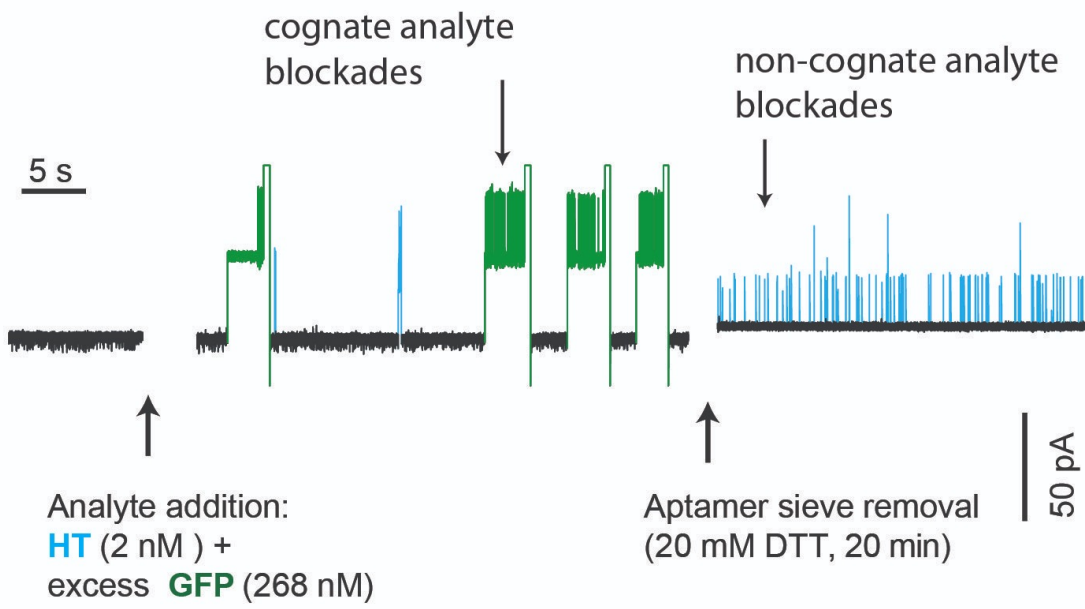


J Am Chem Soc. 6;137(17):5793-7 (2015)

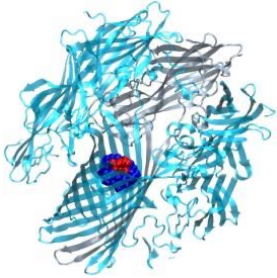
Indirect Small Molecule Sensing

Composed blocage

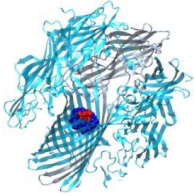




J Am Chem Soc. 6;137(17):5793-7 (2015)



Nanopore as Sequencing Platform



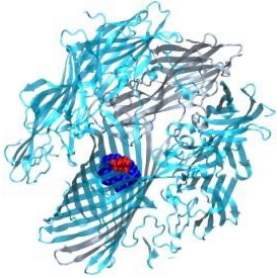
DNA sequencing generations

Then + Now	Now	Now	Now
1st Gen Sanger	2nd Gen -parallised	3rd Gen -single mol or electronic	4th Gen -single mol AND real time
•Low throughput •High cost •Accurate •Broad user base	•Optical •Amplification needed •Highly parallel •Improved cost and Throughput •More centralised users	•Optical •Single-molecule •Highly parallel •Cost similar •New applications •Or electronic, clonal	•Direct electrical (no optics) •Single-molecule, highly parallel •Transformation of workflow •Designed to broaden user base, deliver step change in cost, power •New applications
Sanger	GAII (Solexa/Illumina) SOLiD (Agencourt/LIFE) FLX (454/Roche)	Helicos Pacific Biosciences Ion Torrent (LIFE Starlight)	Nanopores

Estimated cost of a human genome using these technologies

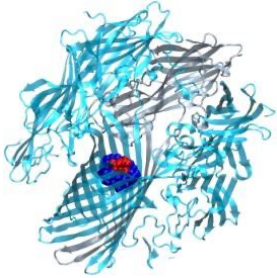
$70M	$200k --- $50k ---- $20k --- 15k---	?$5k - $?

Nanopore molecular Sensing Platforms



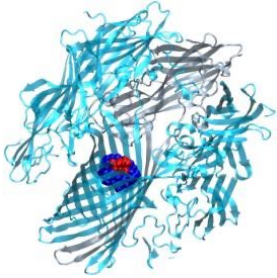
- Allows relatively easy integration in low-cost and portable electronic devices
- Field deployable
- Real time data acquisition
- Unprocessed molecules
- Analytics on DNA and RNA sequence, DNA methylation on the same run
- Access to the raw electronic physical signal

Nanopore molecular Sensing Platforms



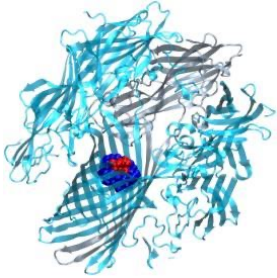
- No fixed run time
  - Can be run one or more nodes for minutes or days.
  - Data analysis takes place in real time.
  - Longer run enables collecting more data points.
- Run until... sufficient data
  - The GridION system enables users to run an experiment until sufficient data has been collected to reach a predetermined experimental endpoint.

Nanopore molecular Sensing Platforms



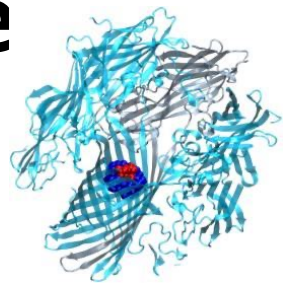
- Scalable
  - SmidgION < 1Gbp; 128 nanopores expected)
  - Flongle (1-2 Gbp; 128 nanopores expected)
  - MinION (5 Gbp; 2000 nanopores; 512 channels)
  - GridION (25-100 Gbp; 10.000 nanopores)

Applications



- Whole genome sequencing
- De novo assembly
- **Scaffolding and finishing**
- Variant analysis: structural variation
- Variant analysis: SNVs, phasing
- Resequencing
- Targeted sequencing
- Panels – amplicons, sequence capture, exome
- Variant analysis: structural variation
- Variant analysis: SNVs, phasing
- 16S rRNA analysis
- RNA sequencing
- Splice variant analysis
- Transcriptome / gene expression
- **Fusion transcript analysis**
- Metagenomics
- Real-time, unbiased analysis of mixed samples
- **Epigenetics**
- **Methylation**
- Histone modification
- **Non-coding RNA activity**

Challenges for producing large scale Knowledge in real-time



- Improvement on Base calling Accuracy

- Demanding data processing - Computational Resources;
- Need of new international nomenclature for nucleic acid representation
  - Non-canonical bases
  - Chemical modifications
- New database concept for nucleic acid information (not only data)
  - Including genetic, structural, and genomic events, genetic synteny …
- New approaches for real time data mining

rpdias@ciencias.ulisboa.pt