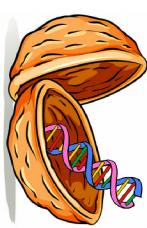


Summary: DNA

- Basics of DNA structure:**
 - Doubt helix hold together by H-bond between G-C, A-T
 - B-form helix is the most stable conformation
 - Other possible conformations: A-form, Z-form, Triplex, Hairpin, Cruciform (Holliday Junction) or unwound DNA

Building with DNA:

- Based on strain hybridization and stably branched DNA
 - Title-based or scaffold based
 - Building blocks: tensegrity triangles, modular DNA (rigid organic vertex) DNA tiles, bundle DNA tiles, DNA origami
 - Dynamic DNA devices:
 - Strand displacement by 'toehold' mechanism (speed related to length + G-C content)
 - DNA tweezers
 - DNA walkers (non-autonomous and autonomous)
 - Computing with DNA (OR and AND gates)
- New articles in Toledo



Summary: Protein Structure

Basics of Protein structure:

- Primary structure held by peptide bonds (peptide plane, angles, Ramachandran plot)
- Tertiary and quaternary structures held by H-bonds (α -helix, beta sheet, beta turn), electrostatic interactions, van der Waals forces, hydrophobic/hydrophilic interactions) AND S-S bonds
- Weak bonds allow dynamic changes
- Protein domains

Protein classification:

- Three main classes: fibrous, globular and membrane proteins
- Fibrous proteins - repeats of amino acids
- Different structures (collagen, alpha keratin, beta keratin)
- Globular protein - hydrophobic core, complex surface with hydrophilic residues
- Membrane protein - peripheral (extrinsic) or integral (intrinsic)
 - Peripherical membrane proteins are bound through ionic/H-bond interactions, hydrophobic alpha helix, hydrophobic loop, association with integral protein
 - Integral proteins are bound through helical transmembrane domains, membrane spanning alpha-sheets or lipid anchors



Learning objectives

Protein production:

- Gene editing
- Cell-based systems
- Case study: insulin

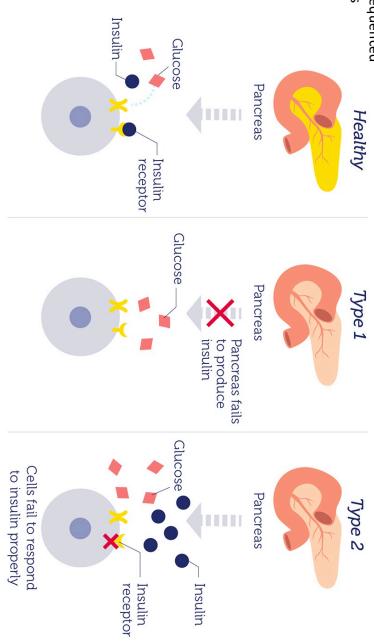
Nanopores

- What are nanopores
- How do nanopores work
- What can nanopores be used for



Diabetes and insulin

- Why insulin?
- First protein to be sequenced
- Several Nobel prizes
- First use in genetic modified organisms



Diabetes and insulin

How is insulin produced?

- Chemically synthesized
- Extracted from bacteria
- Extracted from pancreas from dogs
- Extracted from pancreas from pigs

17

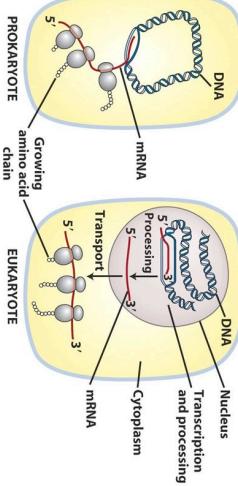
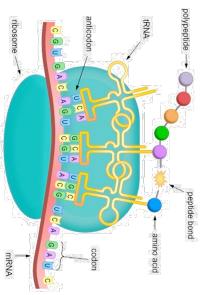
Production of recombinant proteins

- Which organism to use?

- Which plasmid should be chosen?

18

Protein synthesis

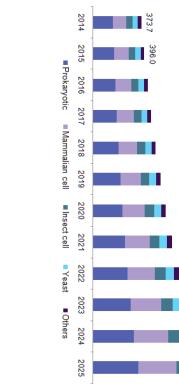


Protein folding and post-translational modifications depend on the system used!

19

Protein production

1. Which organism to use?

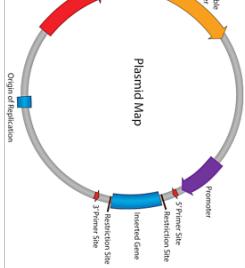


Expression System	Description	Advantages	Disadvantages
Mammalian	Functional assays Structural analysis Enzymatic or metabolic pathway Protein interactions	• Higher-level analysis • Can analyse proteins after maturation or modification • Robust optimisation methods available for most proteins	• Germ and liver needs • Culture conditions difficult to maintain • More expensive
Insect	Functional assays Enzymatic or metabolic pathway Protein interactions	• Similar to mammalian • Functional assays • Enzymatic or metabolic pathway • Protein interactions	• Similar to mammalian • Enzymatic or metabolic pathway • Protein interactions
Yeast	Structural analysis Antibody generation Protein interactions	• Stable • Low cost • Simple culture conditions • Similar to mammalian • Stable media requirements	• Protein stability • Enzymatic or metabolic pathway • May be difficult to express • Some mammalian proteins do not work well in yeast
Bacterial	Stability Antibody generation Protein interactions	• Stable • Low cost • Simple culture conditions • Similar to mammalian • Stable media requirements	• Protein stability • Enzymatic or metabolic pathway • May be difficult to express • Some mammalian proteins do not work well in bacteria
Algal	Stability improvements Plant models Genetic engineering Biosurfactant production	• Genetic modification and enzyme engineering • Biosurfactant production • Stable environment control	• Less developed compared to other organisms • Not much information available • Some organisms are not well understood
Cat-Flow	Toxic proteins or unnatural use of amino acids Fast expression Protein interactions Transporter inhibitor Screening	• Toxic proteins or unnatural use of amino acids • Fast expression • Protein interactions • Transporter inhibitor • Screening	• Some genes may not be correctly expressed • Screening may not be completely effective

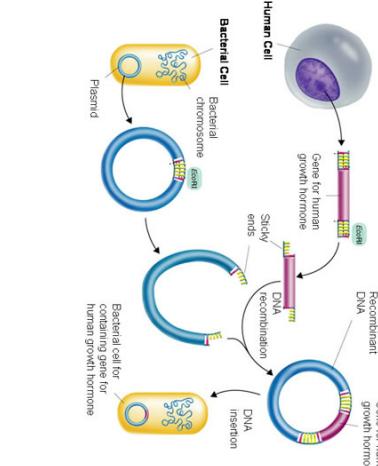
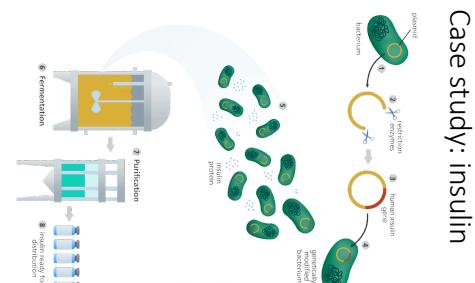
What is a plasmid?

Small circular piece of DNA found in bacterial cells

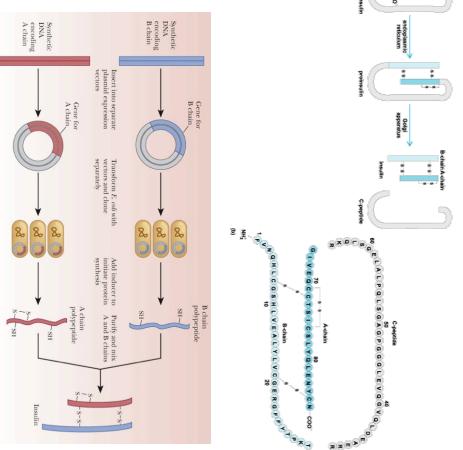
Vector Element	Description
Origin of Replication (ORI)	DNA sequence which allows initiation of replication within a plasmid by recruiting transcription machinery proteins
Antibiotic Resistance Gene	Allows for selection of plasmid-containing bacteria.
Multiple Cloning Site (MCS)	Short segment of DNA which contains several restriction sites allowing for the easy insertion of DNA. In expression plasmids, the MCS is often downstream from a promoter.
Insert	gene, promoter or other DNA fragment cloned into the MCS for gene表达
Primer Binding Site	Primer site for PCR amplification or sequencing
Selectable Marker	A short sequence of DNA sequence used to have selectable markers or use in many plasmids for PCR amplification or sequencing



Case study: insulin



22



23

Recombinant proteins

Steps for protein production

- Get protein or DNA sequence
- Select expression system (prokaryotic, eukaryotic...)
- Insert DNA sequence coding for the protein in plasmid DNA
- Deliver plasmid DNA into expression system
- Let the cells multiply – and synthesize the protein
- Extract and purify your protein

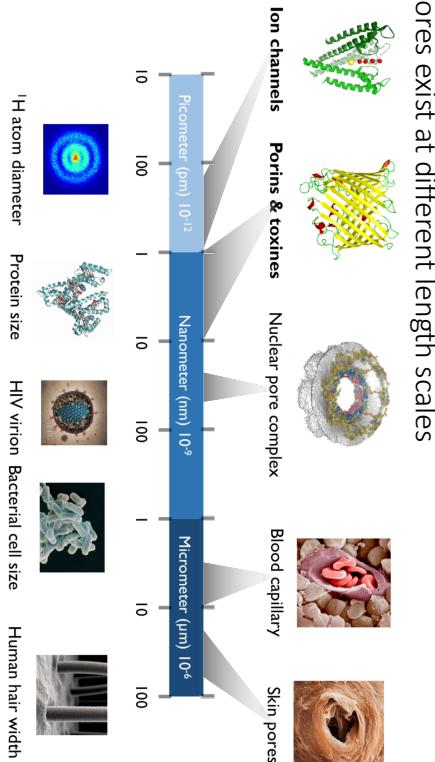
25

What are nanopores?

from Greek πόρος (*poros*)
meaning 'passage'

nanopore

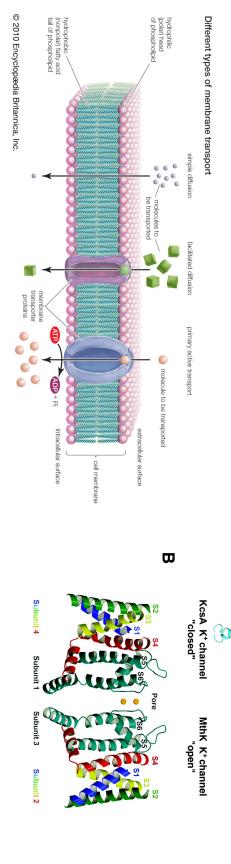
from Greek νάνος (*nanos*)
meaning 'dwarf'



25

Pores exist at different length scales

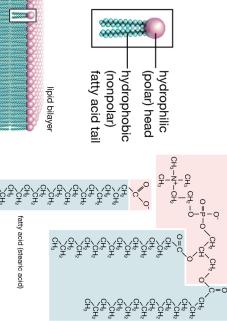
- Transport across membrane is regulated by protein pores
- Channels for ions, water, nutrients and waste products
- Can also be toxins: secreted by bacteria to perforate and kill host cells



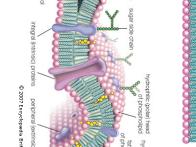
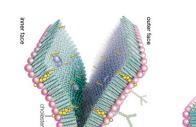
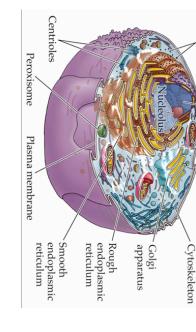
31

In the cell, the lipid bilayer is the 'border wall'...

- Membrane separates the inside from the outside of the cell
- Consists of a lipid bilayer with hydrophobic core
- Impermeable for charged or large molecules
- Used for protection, energy storage or signal transduction



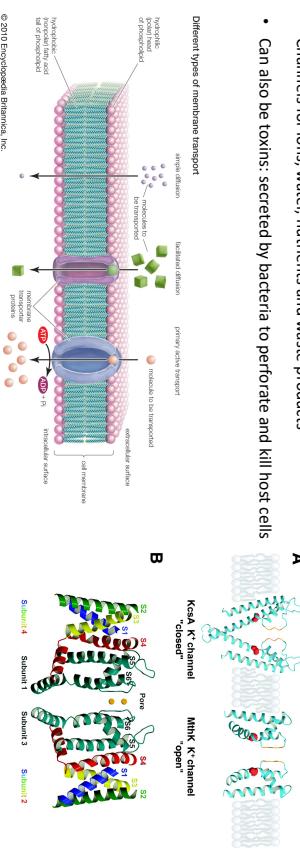
30



30

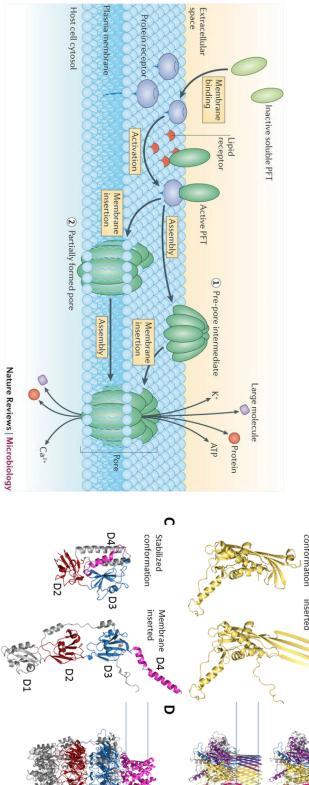
... and nanopores are its 'immigration officers'

- Soluble PFTs are recruited to the host membrane by protein receptors and/or specific interactions with lipids
- Upon membrane binding, the toxins concentrate and start the oligomerization process
- Protein ring stabilizes and punctures the cell membrane



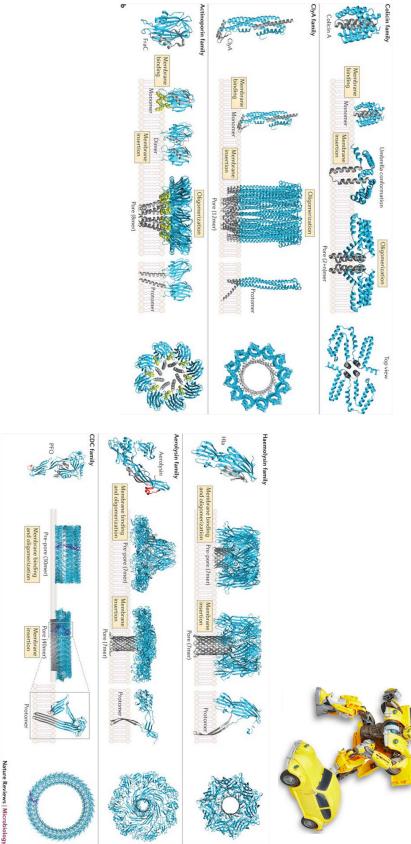
31

Most nanopores in nanotechnology are bacterial cytotoxins



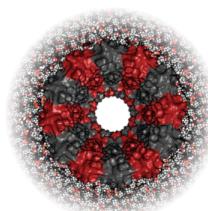
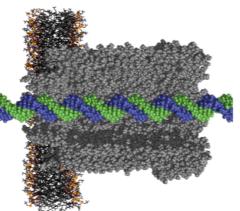
Nature Reviews Microbiology

Structural architectures and pore formation mechanisms of pore-forming toxin families



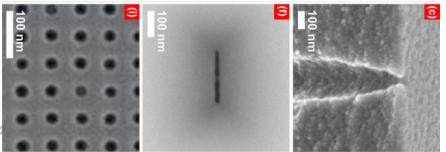
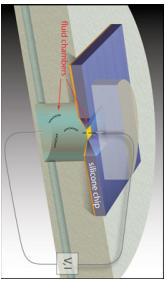
Nanopores come in “biological” flavors ...

- Bottom up fabrication (self-assembly), protein or DNA-based
- Advantages
 - ✓ Atomic level precision
 - ✓ Highly reproducible dimensions
 - ✓ High bio-compatibility
 - ✓ Cheap (protein- or DNA-based)
- Disadvantages
 - ✗ Fragile due to lipid bilayer
 - ✗ Difficult to parallelize (stochastic)

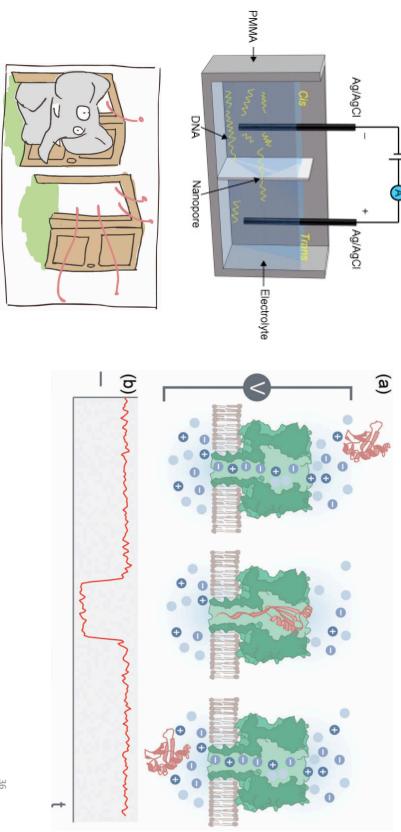


... but also exist in “solid-state” materials

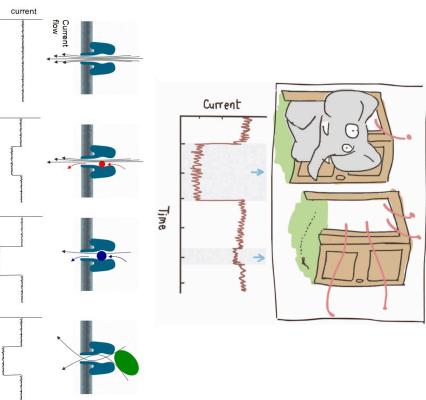
- Top down fabrication (micromachining), silicon-based
- Advantages
 - ✓ Robust due to “solid” materials
 - ✓ Adaptable shape, size and length
 - ✓ Straightforward to create arrays
- Disadvantages
 - ✗ Poor size reproducibility
 - ✗ Atomic level modifications are difficult
 - ✗ Expensive to fabricate



How do nanopores work?

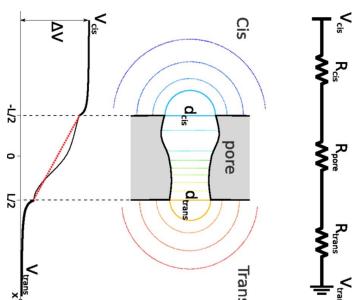
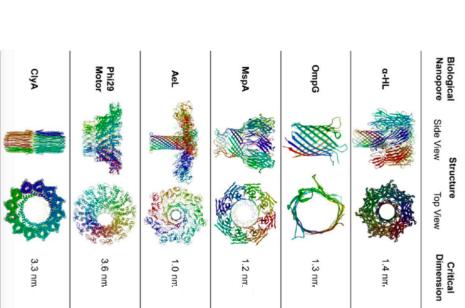


Size matters!



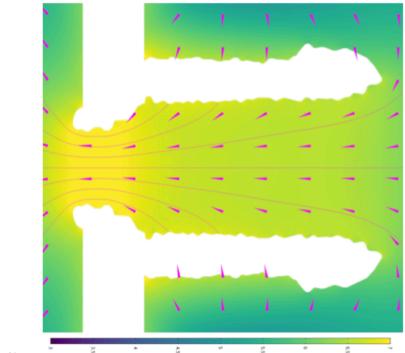
The application of a bias voltage results in a strong electrical field in the nanopore

- Nanopores have very high electrical (ionic) resistances ($\equiv 0.1$ to $1\text{ G}\Omega$)
- Most of the voltage change ($\Delta V/V$) occurs near and within the nanopore, resulting in very high electrical fields (10^6 to 10^8 V/m)
- Example: $\Delta V = 100 \text{ mV}$ over a nanopore with length $l = 10 \text{ nm}$:
 $E_z = \Delta V/l = 10 \text{ mV/nm} = 10^7 \text{ V/m}$

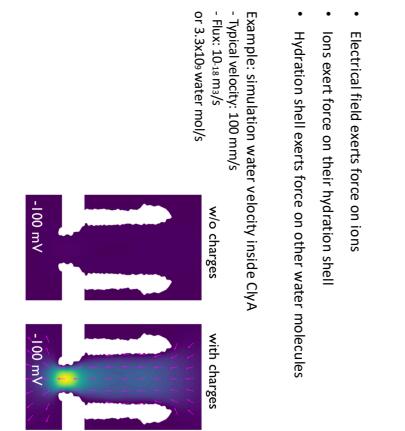


The electrical field drives ions and charged molecules through the pore

- Ions in solution migrate along electrical field lines
 - current depends on mobility of ions (size and charge)
 - highly charged pore is more selective for it
- For highly charged molecules (e.g. DNA), electrophoresis typically dominates
 - effective charge (~ionic strength) determines counter-ion mobility



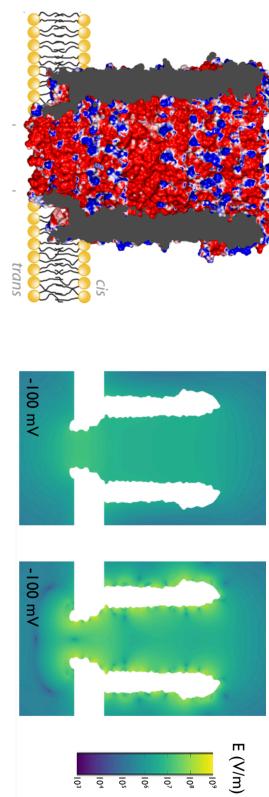
39



40

Electrical field is heavily influenced by fixed charges inside the pore walls.

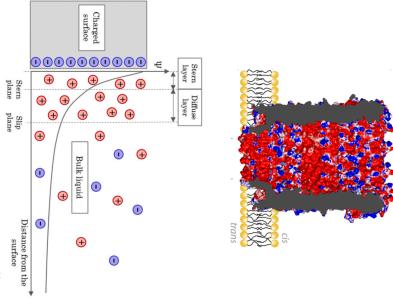
The cytolytic A (CytA) protein nanopore contains a highly negatively charged interior



41

Many nanopores have fixed charges in their walls, resulting in an electrical double layer (EDL)

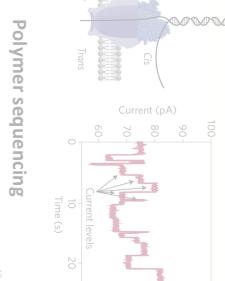
- **Bulk electrolyte**
 - ions are mobile charges
 - locally electrically neutral (+ == -)
- **Electrolyte near charged surface**
 - fixed charges attract layer of oppositely charged counter-ions (= electrical double layer, EDL)
 - EDL hides the presence of the surface charges from the bulk solution ('screening')
- **Electrical double layer**
 - thickness depends on the ionic strength (cfr. Debye length)



42

Exploring molecular interactions

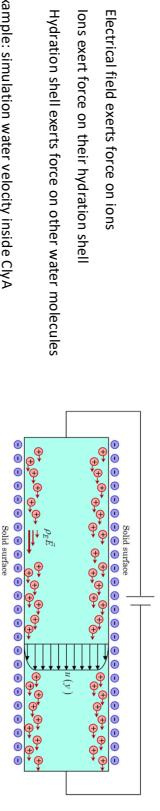
Molecular recognition/binding



43

Applications of protein pores

- Detecting a wide range of substances including organic species, ions and proteins
- Studying biomolecular folding and unfolding
- Investigating covalent and non-covalent interactions
- Probing enzyme kinetics
- Sequencing DNA molecules



44

The drag on the fluid by the movement of the EDL results in a net electro-osmotic flow

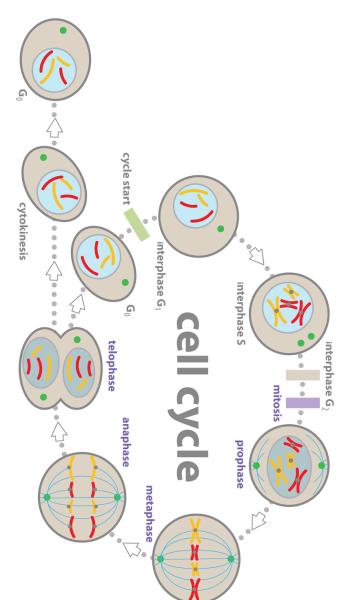
- Electrical field exerts force on ions
 - ions exert force on their hydration shell
 - Hydration shell exerts force on other water molecules
- Example: simulation water velocity inside CytA
 - Typical velocity: 100 nm/s
 - Flux: 10 μ m/s
 - or 3.3x10⁶ water mol/s

45

How much does it cost to sequence the genome of one person today?

DNA replication

- > US\$ 10?
- > US\$ 100?
- > US\$ 1.000?
- > US\$ 10.000?
- > US\$ 100.000?



54

How much would it cost today?



Human Genome Project

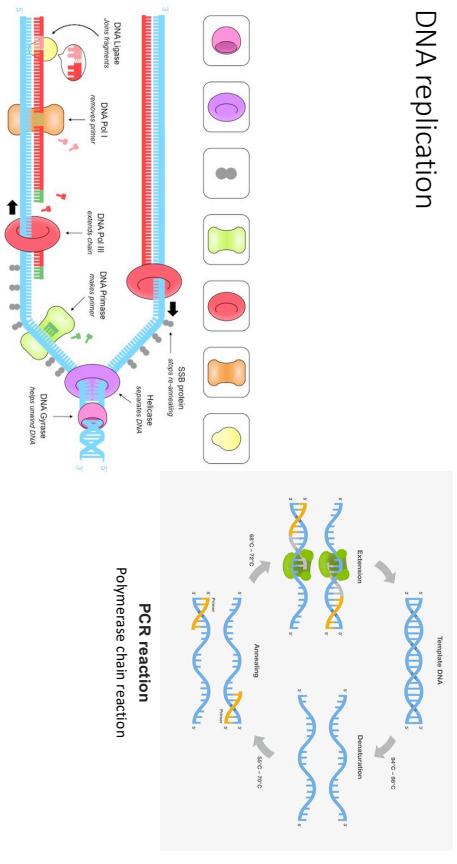
\$2.7 billion
13 years

Veritas
\$600
12 weeks



52

DNA replication



55

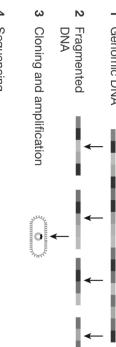
Evolution DNA sequencing techniques



- Gel-based systems
- Automated slab gel
- First-generation capillary sequencing
- Second-generation capillary sequencing
- Microfluidic pyrosequencing
- Massively parallel sequencing
- Single molecule?

53

First generation DNA sequencing



56

53

Polymerase chain reaction with special nucleotides

Next generation methods - Illumina

https://www.youtube.com/watch?time_continue=2&v=Fc6685HRaz8

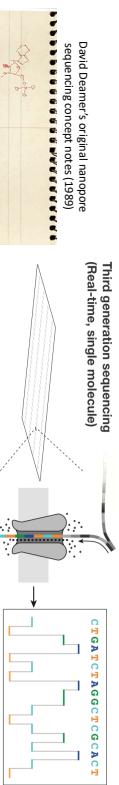
DNA capture

Single stranded DNA is readily captured due to its high negative charge

- DNA is a highly negatively charged polymer ($-1 \text{ e}^- / 0.3 \text{ nm}$)
- High electrical field in and around nanopore automatically captures DNA and translocates it

- Single stranded DNA is balled up in solution, double stranded DNA is more rigid!

Third generation – real time, single molecule



64

Translocation speed

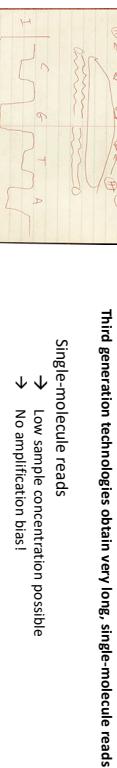
Translocation speed of DNA is generally *too fast*, but it can be controlled in various ways.

Translocation rate

- naked DNA: $> 1 \mu\text{s}/\text{base}$
- ideally: $> 1 \text{ ms}/\text{base}$

Non-enzymatic control

- Modification of pore
- Modification of DNA
- Modification of electrolyte

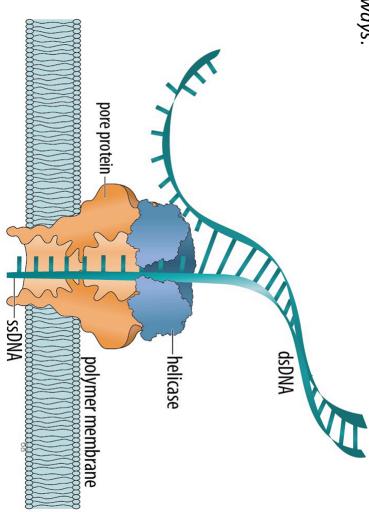


65

The challenges of strand sequencing

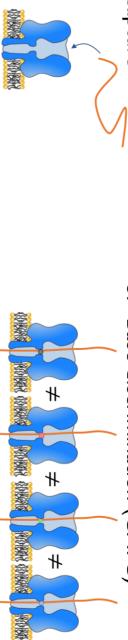
Electrostatic interactions can dramatically reduce DNA translocation speed

- Electrostatic 'brakes'
- additional positive charges
- slowdown of DNA translocation



67

1. DNA capture
2. Translocation speed
3. Base discrimination (GATC)
4. Single base resolution



Charges in the β barrel

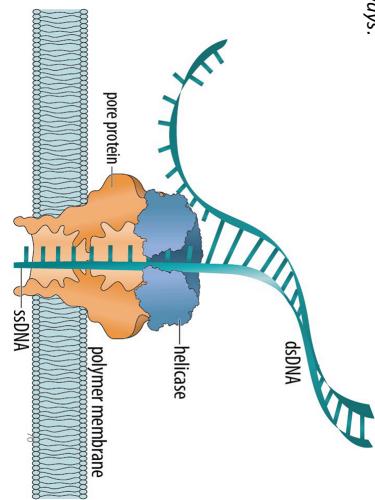
66

69

Translocation speed

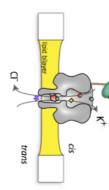
Translocation speed of DNA is generally too fast, but it can be controlled in various ways.

- naked DNA: < 1 ns/base
- ideally: > 1 ns/base
- Non-enzymatic control
 - Modification of pore
 - Modification of DNA
- Enzymatic control
 - DNA polymerase
 - Processive exonuclease
 - Helicase
 - Single strand binding protein



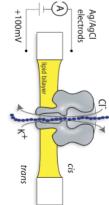
DNA sequencing with (biological) nanopores

Exonuclease sequencing



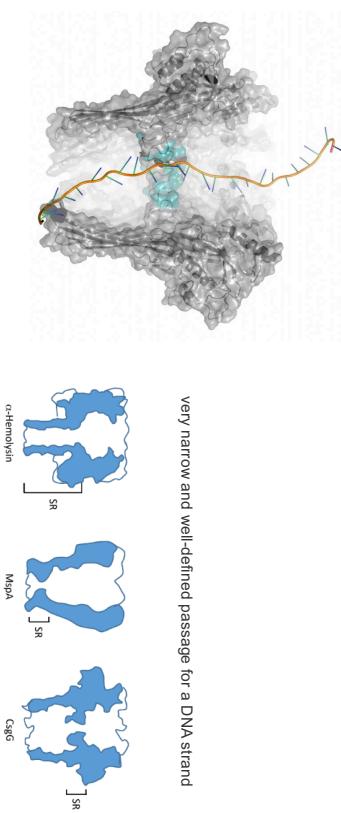
- Exonuclease enzyme cleaves DNA one base at a time
- Each cleaved base is captured by the nanopore and identified by its unique ionic current disruption

Strand sequencing



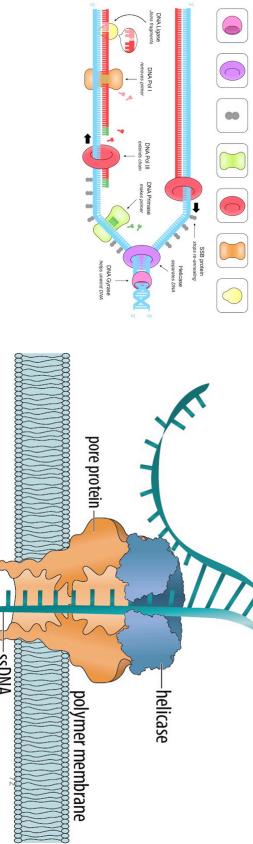
- Entire ssDNA strand is driven through the pore via electrophoresis
- The base at the narrowest location in the pore determines the ionic current disruption

Nanopore used



Enzymatic control using DNA helicase

- Enzymes are used to control the speed of translocation across the nanopore
- Using an helicase, the speed of DNA across the nanopore is modulated by the concentration of ATP
- DNA is recognized during the unzipping of the DNA



Learning objectives 2nd lesson

Protein production:

- Gene editing

- Cell-based systems

- Case study: insulin

Nanopores

- What are nanopores
- How do nanopores work
- What can nanopores be used for (determination chemicals, DNA sequencing)

Commercial devices

Oxford Nanopore technologies

- First commercial nanopore sequencing
- From hand-held to large scale
- Core technologies
 - Strand sequencing with biological pore (CsgG) and enzymatic translocation control
 - 512 microfluidics 'wells' with a polymerized block-copolymer bilayer (stability!)
 - Disposable chamber



MinION

GridION

PromethION