Protein Prediction - Part 1: Structure
Definitions

- Computational Biology
- Bioinformatics
- http://www.rostlab.org/talks/
Definitions

- **Computational Biology**
  - biology replacing experiments by computers (include neurobiology, image processing)

- **Bioinformatics**
  - anything that has to do with storing and using the information about bio-sequences

- [http://www.rostlab.org/talks/]
2013 Nobel Prizes in Chemistry
2013 Nobel Prize in Chemistry

“for the development of multiscale models for complex chemical systems” (nobelprize.org)

**Martin Karplus**  
Harvard Univ, USA & Univ Strasbourg, France

**Michael Levitt**  
Stanford Univ, USA

**Arieh Warshel**  
USC, Los Angeles, USA
Protein structure comparisons

- All-alpha: 3sdh
- All-beta: 1bww
- AlphaBeta: 1xne
Predict protein function
Classical physics vs quantum physics

Figure 1
Figure 1 Multi-copper-oxidase embedded in water.

Figure 3. To understand how lysozyme cleaves a glycoside chain, it is necessary to model only the relevant parts of the system using quantum chemistry, while most of the surrounding may be treated using molecular mechanics or a continuum model. The figure is adapted from A Warshel & M Levitt (1976) JMB 103, 227.

Classical physics vs quantum physics

Figure 4. The detailed structure of a polypeptide chain (top) is simplified by assigning each amino acid residue with an interaction volume (middle) and the resulting string-of-pearls like structure (bottom) is used for the simulation.

Michael Levitt (AP): It’s sort of nice ... to see that computational science, computational biology is being recognized. ... It’s become a very large field and it’s always in some ways been the poor sister, or the ugly sister, to experimental biology.

Arieh Warshel (NobelPrize) [experimental results are] like seeing a watch and wondering how actually it works. ... what we developed is a way which requires a computer to look ... and then to eventually understand how exactly [a protein] does what it does.
Who are we?
Burkhard Rost
Some Professional Positions

- Professor for Computational Biology, TUM, Dept. Computer Science
- Fellow Institute of Advanced Studies, TUM
- Affiliate professor @ Columbia University in the City of New York, College for Physicians and Surgeons, Dept. Biochemistry and Molecular Biophysics
- Member of New York Structural Biology Center, City College New York
Study complex systems

  Master: Neuronal networks
  (Heidelberg, Bochum, Princeton, Harvard, UCSF)
- 1989-1995: PhD/Postdoc: Application of physics to
  Computational Biology, EMBL (European Molecular Biology
  Laboratory, Heidelberg)
- 1995-1996: European Bioinformatics Inst. (EBI), Cambridge, UK
- 1998: LION Biosciences
- 1998-now: Professor at Columbia University, New York
  (tenure since 2005)
- 2009: Alexander von Humboldt professorship
- 2009-now: Professor at TUM, Chair for Comp. Biology
Nothing makes me happier than to be proven wrong.
My passion is my job
Little $, but attention, friends, travel, fun
Little $, but attention, friends, travel, fun

- Seminars on 5 continents
  Dec 2009 - 2011
  (TV in 4: USA, Mali, Uruguay, Germany)
Little $, but attention, friends, travel, fun

- Seminars on 5 continents
  Dec 2009 - 2011
  (TV in 4: USA, Mali, Uruguay, Germany)
- 140 invited talks in 21 countries
Little $, but attention, friends, travel, fun

- Seminars on 5 continents
  Dec 2009 - 2011
  (TV in 4: USA, Mali, Uruguay, Germany)

- 140 invited talks in 21 countries

- 197 publications, 150 peer-reviewed
Little $, but attention, friends, travel, fun

- Seminars on 5 continents
  Dec 2009 - 2011
  (TV in 4: USA, Mali, Uruguay, Germany)
- 140 invited talks in 21 countries
- 197 publications, 150 peer-reviewed
- > 25,000 citations to our work; H-index > 55
Little $, but attention, friends, travel, fun

- Seminars on 5 continents
  Dec 2009 - 2011
  (TV in 4: USA, Mali, Uruguay, Germany)
- 140 invited talks in 21 countries
- 197 publications, 150 peer-reviewed
- > 25,000 citations to our work;
  H-index > 55
- google.com “rost”
Rostlab & friends @ ISMB/ECCB Berlin

www.rostlab.org
Contact: Lothar Richter: teaching@rostlab.org
Tim Karl

Contact: Tim Karl:
karl@rostlab.org

Videos/Sysadmin

How sysadmin gets you started into studying Bioinformatics....
Exercises

PhD students in group:
- Tatyana Goldberg
- Max Hecht
- Lothar Richter

Contact: see www.rostlab.org

http://rostlab.org/cms/teaching/
Contact: Marlena Drabik: assistant@rostlab.org
## Announcements

- **Videos:** YouTube / [www.rostlab.org](http://www.rostlab.org)

  **THANKS:**
  - Tim Karl + Jonas Reeb

- **Special lectures:**
  - Apr 15: Andrea Schafferhans
  - Nov 28: Arthur Dong
  - Dec 03+05: Marco De Vivo/Marco Punta
  - Dec 17+19: Andrea Schafferhans

- **No lecture:**
  - Apr 17/22 Easter
  - May 01 Thu May day
  - May 06 Tue Student assembly
  - May 29 Thu Ascension day
  - Jun 10 Tue Whitsun holidays
  - Jun 19 Thu Corpus Christi

- **LAST lecture:** July 1

- **Examen:** July 3/8
  - Makeup: Oct 21 - morning

---

**CONTACT:** Lothar Richter richter@rostlab.org
Lecture plan (PP1: Structure)-generic

- 01: 2014/04/08 Tue: sorry
- 02: 2014/04/10 Thu: welcome: who we are
- 03: 2014/04/15 Tue: Intro I - acids/structure (Andrea Schafferhans)
- 04: 2014/04/17 Thu: SKIP: Easter vacation
- 05: 2014/04/22 Tue: SKIP: Easter vacation
- 06: 2014/04/24 Thu: Intro II - 3D comparisons
- 07: 2014/04/29 Tue: Alignment 1
- 08: 2014/05/01 Thu: SKIP: “May day” - (NOT to be confused with “m’aidez”)
- 09: 2014/05/06 Tue: SKIP: student assembly (SVV)
- 10: 2014/05/08 Thu: Alignment 2
- 11: 2014/05/13 Tue: Comparative modeling 1
- 12: 2014/05/15 Thu: Comparative modeling 2
- 13: 2014/05/20 Tue: Experimental structure determination
- 14: 2014/05/22 Thu: 3D -> 1D: Secondary structure assignment
- 15: 2014/05/27 Tue: 1D: Secondary structure prediction 1
- 16: 2014/05/29 Thu: SKIP: holiday (Ascension Day)
- 17: 2014/06/03 Tue: Normal mode analysis (Edda Kloppmann)
- 18: 2014/06/05 Thu: 1D: Secondary structure prediction 3
- 19: 2014/06/10 Tue: SKIP: Whitsun holidays
- 20: 2014/06/12 Thu: 1D: Transmembrane helix prediction
- 21: 2014/06/17 Tue: Nobel prize symposium
- 22: 2014/06/19 Thu: SKIP: Corpus Christi (Fronleichnam)
- 23: 2014/06/24 Tue: 1D: Transmembrane strand prediction, solvent accessibility
- 24: 2014/06/26 Thu: 2D prediction
- 25: 2014/07/01 Tue: 3D prediction/wrap up
- 26: 2014/07/03 Thu: examen, no lecture
- 27: 2014/07/10 Thu: no lecture
Exercises

☐ SNP in membrane proteins -> disease?
☐ predict protein expression Jonas, Tanya, et al
☐ another idea
☐ anything else?
Exercises - ECTS/score/admin

- final score: 60% exercises - 40% examen
Foundations
What is this?

© Wikipedia
What is similar?
What is similar?
What is similar?

- Silicone
- Glass

© Wikipedia

© http://www.lionsons.com
What is this?

Diamond

Graphite/coal

5mm
1/4"

© Wikipedia

© Burkhard Rost (TU Munich)
What is this?

Diamond

Graphite/coal

carbon

diamond lattice

© Wikipedia
What is this?

Diamond

Graphite/coal

carbon
diamond lattice

© Wikipedia
Periodic table

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>1</td>
<td>H</td>
<td>Li</td>
<td>Be</td>
<td>B</td>
<td>C</td>
<td>N</td>
<td>O</td>
<td>F</td>
<td>Ne</td>
<td>Na</td>
<td>Mg</td>
<td>Al</td>
<td>Si</td>
<td>P</td>
<td>S</td>
<td>Cl</td>
<td>Ar</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>K</td>
<td>Ca</td>
<td>Sc</td>
<td>Ti</td>
<td>V</td>
<td>Cr</td>
<td>Mn</td>
<td>Fe</td>
<td>Co</td>
<td>Ni</td>
<td>Cu</td>
<td>Zn</td>
<td>Ga</td>
<td>Ge</td>
<td>As</td>
<td>Se</td>
<td>Br</td>
<td>Kr</td>
</tr>
<tr>
<td>3</td>
<td>Rb</td>
<td>Sr</td>
<td>Y</td>
<td>Zr</td>
<td>Nb</td>
<td>Mo</td>
<td>Tc</td>
<td>Ru</td>
<td>Rh</td>
<td>Pd</td>
<td>Ag</td>
<td>Cd</td>
<td>In</td>
<td>Sn</td>
<td>Sb</td>
<td>Te</td>
<td>I</td>
<td>Xe</td>
</tr>
<tr>
<td>4</td>
<td>Cs</td>
<td>Ba</td>
<td>Hf</td>
<td>Ta</td>
<td>W</td>
<td>Re</td>
<td>Os</td>
<td>Ir</td>
<td>Pt</td>
<td>Au</td>
<td>Hg</td>
<td>Tl</td>
<td>Pb</td>
<td>Bi</td>
<td>Po</td>
<td>At</td>
<td>Rn</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Fr</td>
<td>Ra</td>
<td>Rf</td>
<td>Db</td>
<td>Sg</td>
<td>Bh</td>
<td>Hs</td>
<td>Mt</td>
<td>Ds</td>
<td>Og</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Ac</td>
<td>Th</td>
<td>Pa</td>
<td>U</td>
<td>Np</td>
<td>Pu</td>
<td>Am</td>
<td>Cm</td>
<td>Bk</td>
<td>Cf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© Wikipedia
What elements make most of our body?
What elements make most of our body?

H2O - water

how much?
What elements make most of our body?

H₂O - water

about 50-65% in adults
Which ones/how many make up life?

© Wikipedia

© Burkhard Rost (TU Munich)
Which ones make up life?

- Oxygen - O: 65.00%
- Hydrogen - H: 9.70%
- Body weight: 96.5%

© Wikipedia

Wednesday April 9, 2014
Which ones make up life?

<table>
<thead>
<tr>
<th>Element</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>oxygen - O</td>
<td>65.00</td>
</tr>
<tr>
<td>carbon - C</td>
<td>18.60</td>
</tr>
<tr>
<td>hydrogen - H</td>
<td>9.70</td>
</tr>
<tr>
<td>nitrogen - N</td>
<td>3.20</td>
</tr>
<tr>
<td>calcium - Ca</td>
<td>1.80</td>
</tr>
<tr>
<td>phosphorus - P</td>
<td>1.00</td>
</tr>
<tr>
<td>body weight</td>
<td>96.5</td>
</tr>
</tbody>
</table>

© Wikipedia

Wednesday April 9, 2014
### Which ones make up life?

<table>
<thead>
<tr>
<th>Element</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>oxygen - O</td>
<td>65.00</td>
</tr>
<tr>
<td>carbon - C</td>
<td>18.60</td>
</tr>
<tr>
<td>hydrogen - H</td>
<td>9.70</td>
</tr>
<tr>
<td>nitrogen - N</td>
<td>3.20</td>
</tr>
<tr>
<td>calcium - Ca</td>
<td>1.80</td>
</tr>
<tr>
<td>phosphorus - P</td>
<td>1.00</td>
</tr>
<tr>
<td>potassium - K</td>
<td>0.40</td>
</tr>
<tr>
<td>sodium - Na</td>
<td>0.20</td>
</tr>
<tr>
<td>chlorine - Cl</td>
<td>0.20</td>
</tr>
<tr>
<td>magnesium - Mg</td>
<td>0.10</td>
</tr>
<tr>
<td>sulfur - S</td>
<td>0.05</td>
</tr>
<tr>
<td>iron - Fe</td>
<td>0.03</td>
</tr>
<tr>
<td>iodine - I</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>body weight</strong></td>
<td><strong>96.5</strong></td>
</tr>
</tbody>
</table>

© Wikipedia
What is this?

© Wikipedia
What is this?

Escherichia coli
What is this?

Escherichia coli

© Wikipedia

© Wikipedia
What is this?

What is this?

Escherichia coli /E coli

What is this?

Streptococcus

bigbackground.com
What is this?

© Wikipedia
What is this?

Salmonella bongori © Wikipedia
What is this?

What is this?

Avian or bird flu (virus)

What is this?


Wednesday April 9, 2014
What is this?

HIV virus

What is this?
Velcro - invented by George de Mestral

George de Mestral (1907-1990)

• Swiss electrical engineer
• Invented velcro 1949

http://www.engr.sjsu.edu/WofMatE/polymers.htm

© Wikipedia
Velcro copied from plants

George de Mestral (1907-1990)

- Swiss electrical engineer
- Invented velcro 1949

Velcro copied from plants

George de Mestral

(1907-1990)

- Swiss electrical engineer
- Invented velcro 1949

© Wikipedia

Wednesday April 9, 2014
Bacteria differ in shape

© Wikipedia

© Burkhard Rost (TU Munich)
What do they have in common?

Cocci
- coccus
- diplococci
- diplococci
- encapsulated
- Pneumococcus
- Staphylococci
- streptococci
- sarcina
- tetrad

Others
- enlarged rod
- Fusobacterium

Bacilli
- coccobacillus
- bacillus
- diplobacilli
- palisades
- Streptobacilli

Budding and appendaged bacteria
- hypha
- stalk

© Wikipedia
Common to bacteria: single cells

Single cell

© Wikipedia
Cells

Eukaryote
- Membrane-enclosed nucleus
- Nucleolus
- Mitochondrion
- Ribosomes
- Cell Membrane

Prokaryote
- Nucleoid
- Capsule (some prokaryotes)
- Flagellum
- Cell Wall (in some eukaryotes)

© Wikipedia
Do we carry bacteria?

*Leonardo Da Vinci (1452-1519)*

*Vitruve Luc Viatour (~1492)*
we carry bacteria when we are sick
Our bodies are teeming with bacteria

Life is diverse

Trilobite Bergeroniellus spinosus
Lena River Gorge, Siberia
http://www.emory.edu/COLLEGE/ENVS/research/ichnology
Trilobite Bergeroniellus spinosus
Lena River Gorge, Siberia
http://www.emory.edu/COLLEGE/ENVS/research/ichnology

Common?
DNA/Gene

www.chemistryexplained.com

© Wikipedia
Where do we have genes?
Where do we have genes?
Genes in nucleus and mitochondria

green: cell (actin)  red: mitochondria  blue: nucleus

© Wikipedia
Genes in nucleus and mitochondria

green: cell (actin)  
red: mitochondria  
blue: nucleus

© Wikipedia
Genes are it: what are proteins?
Central dogma of molecular biology

DNA → RNA → Protein

- DNA: information, code, library, manual
- RNA: intermediate step
- Protein: machinery of life

© Laszlo Kajan, TUM
Central dogma of molecular biology

DNA → RNA → Protein

information, code, library, manual

intermediate step

machinery of life

© Laszlo Kajan, TUM

© Burkhard Rost (TU Munich)
Protein synthesis

(a) Transcription
(b) Post-transcription
(c) Translation
(d) Post-translation

RNARNA
NuclearMembrane
Ribosome
mRNA
Polypeptide
Protein Folds Upon Itself
Effector Molecule
Active Protein

© Wikipedia

© Burkhard Rost (TU Munich)
Protein synthesis

(a) Transcription
(b) Post-transcription
(c) Translation
(d) Post-translation

© Laszlo Kajan, TUM

© Wikipedia
Quiz

1. what is the smallest building block of life that can replicate?
   • protein - quarks - cells - organelles
2. different cells in typical human?
   • 20 - 200 - 400 - 1000
3. what are the parts of cells called?
   • cellular parts - organelles - organs - celluoplasts
4. which part of cells is called the “powerhouse”?
   • nucleus - Golgi apparatus - mitochondria - Endoplasmic reticulum
5. what part of a plant cell is involved with photosynthesis?
   • mitochondria - nucleus - smooth reticulum - chloroplast
6. what is mitosis?
   • cell death - cellular respiration - cell division - cellular communication
7. who first used the term cell?
   • Aristotle - Captain Hooke - Robert Hooke - James Watson
8. how many elements are found in amounts larger than trace amounts (0.01%) in our bodies?
   • 92 - 48 - 13 - 11
9. when communities of living things interact with non-living things they are called a ...
   • population - community - biosphere - ecosystem
10. the most common molecule in the human body is?

© Ask a Biologist (askabiologist.asu.edu)
Quiz

1. **what is the smallest building block of life that can replicate?**
   - protein - quarks - cells - organelles

2. **different cells in typical human?**
   - 20 - 200 - 400 - 1000

3. **what are the parts of cells called?**
   - cellular parts - organelles - organs - celluoplasts

4. **which part of cells is called the “powerhouse”?**
   - nucleus - Golgi apparatus - mitochondria - Endoplasmic reticulum

5. **what part of a plant cell is involved with photosynthesis?**
   - mitochondria - nucleus - smooth reticulum - chloroplast

6. **what is mitosis?**
   - cell death - cellular respiration - cell division - cellular communication

7. **who first used the term cell?**
   - Aristotle - Captain Hooke - Robert Hooke - James Watson

8. **how many elements are found in amounts larger than trace amounts (0.01%) in our bodies?**
   - 92 - 48 - 13 - 11

9. **when communities of living things interact with non-living things they are called a ...?**
   - population - community - biosphere - ecosystem

10. **the most common molecule in the human body is?**

© Ask a Biologist (askabiologist.asu.edu)