



# Computer skills for wetlab scientists

- **Molecular Biology Software**

Cloning: **SnapGene**, **PlasmaDNA\***, **SerialCloner<sup>+</sup>**

Sequence assembly: **Staden<sup>+</sup>**

Sequence analysis: **EMBOSS<sup>+</sup>**,

Molecular Visualization: **PyMol<sup>+</sup>**,

Alignments and Phylogeny: **ClustalX/W/W2/Ω<sup>+</sup>**, **T-Coffee<sup>+</sup>**

- **General software**

Pixel graphics: **Photoshop**, **GIMP<sup>+</sup>**

3D: **Blender<sup>+</sup>**

Vector graphics: **CorelDraw**, **Illustrator**, **Inkscape<sup>+</sup>**

Text Editors: **MS Office**, **LibreOffice<sup>+</sup>**

DTP: **InDesign**, **Scribus<sup>+</sup>**, **LaTeX<sup>+</sup>**

- **Image analysis**

**ImageJ (NIH)<sup>+</sup>**, **Icy<sup>+</sup>**, **Jmicrovision<sup>+</sup>**, **Image Studio Lite**

- **Reference management**

**EndNote**, **RefWorks**, **Papers**, **Zotero<sup>+</sup>**, **Mendeley<sup>+</sup>**

- **Plotting, graph drawing, statistics**

**GraphPad Prism**, **R**, **JMP**

Blue:

Free or open source software

Red:

Licensed by Helsinki University

Green:

Commerical, no HU site license

\*developed at Helsinki University

<sup>+</sup>true cross-platform  
(works on Windows,  
MacOS and Linux)





# SnapGene

- Download link:  
[https://jeltsch.org/snapgene\\_download\\_for\\_2014\\_course\\_participants](https://jeltsch.org/snapgene_download_for_2014_course_participants)
- Register with  
Group Name: Helsinki PMB  
Registration Code: available from the course organizer
- Tutorials:  
[http://www.snapgene.com/support/tutorial\\_videos/introduction\\_to\\_snapgene](http://www.snapgene.com/support/tutorial_videos/introduction_to_snapgene)
- The free SnapGene Viewer might be enough for some people and does much more than just viewing: design primers, create maps, display restriction sites
- If you consider buying a license, enter “[michael@jeltsch.org](mailto:michael@jeltsch.org)” into the discount field to get 10% off the price.



# Serial Cloner

- Download link:  
[http://serialbasics.free.fr/Serial\\_Cloner.html](http://serialbasics.free.fr/Serial_Cloner.html)
- Tutorials: <https://www.youtube.com/user/serialbasics>
- Like many other programs, Serial Cloner does not import GeneConstructionKit, SnapGene, Clone Manager and many other file types (SnapGene imports almost anything)
- Does not do Gibson Assembly or In-Fusion
- Similar free programs to consider: [ApE](#), [pDRAW32](#), [CLC Sequence Viewer](#), [Cloning Help \(online tool\)](#)



# The Staden Package

- For large scale (and small) sequencing projects
- Was designed to run without human intervention on mainframe computers (→ UI is not the most intuitive)
- For automatic assembly and verification of (Sanger) sequencing of cloning products (the same is possible with SnapGene, but is less powerfull and not possible to automate)
- Gap4 (Sanger) vs. Gap5 (NGS)
- You may need go to a course to learn it...





# Bioinformatics Courses

- CSC (Center for Scientific Computing)
- <http://web.csc.fi/csc/kurssit>



# EMBOSS

- EMBOSS = European Molecular Biology Open Software Suite
- Old, command line tool
- Local (on your computer), also as web service with a GUI (e.g. [emboss.bioinformatics.nl](http://emboss.bioinformatics.nl))
- Examples
  - pepstats (display properties of a protein)
  - needle (global alignment of two proteins)

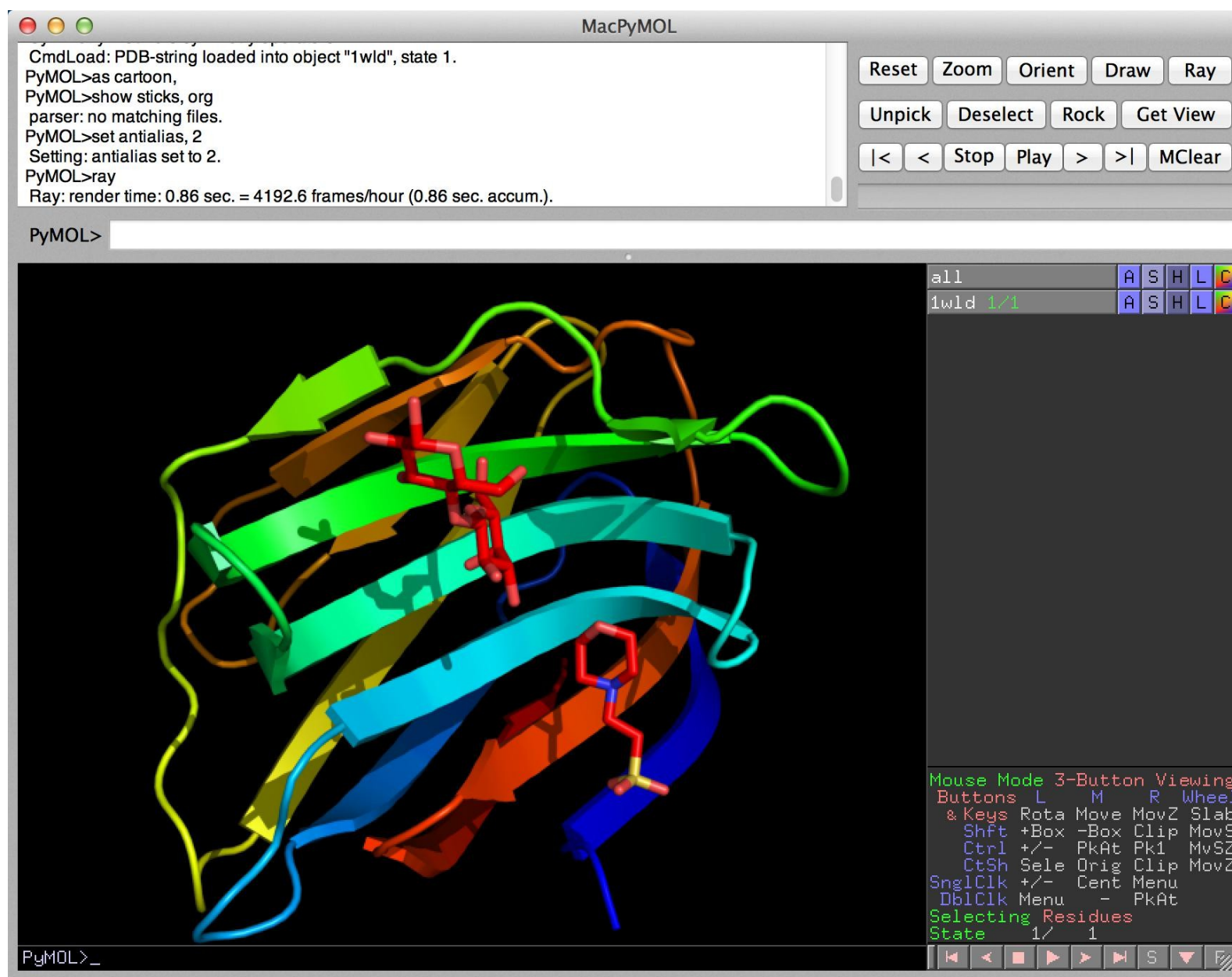


# EMBOSS

```
lm2-bmh-dr54:Lecture_07 mjeltsch$ pepstats hVEGF-C.fasta
Protein statistics
Pepstats program output file [cwt.pepstats]:
lm2-bmh-dr54:Lecture_07 mjeltsch$ more cwt.pepstats
PEPSTATS of Cwt from 1 to 125
Molecular weight = 14095.19                      Residues = 125
Isoelectric Point = 8.3248
[...]
PEPSTATS of C-H6 from 1 to 131
Molecular weight = 14918.04                      Residues = 131
Isoelectric Point = 8.3395
[...]
PEPSTATS of C-strep3 from 1 to 153
Molecular weight = 16950.18                      Residues = 153
Isoelectric Point = 8.3288
[...]
lm2-bmh-dr54:Lecture_07 mjeltsch$
```



# Molecular Visualization Software: PyMol





# Molecular Visualization Software

Many others...

MolMol: <http://sourceforge.net/projects/molmol/>

Chimera: <http://www.cgl.ucsf.edu/chimera/>

CN3D: <http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>



*The multiple sequence alignment result as produced by T-coffee.*

\*

|                 |     |                                                                  |     |
|-----------------|-----|------------------------------------------------------------------|-----|
| Homo_sapiens    | 126 | WRKTQCMPREVCIDVGKEFGVATNTFFKPPCVSVYRCGGCCNSEGLQCMNTSTSYLSKTLFEI  | 188 |
| Danio_rerio     | 114 | WRKTLCMPRQVCLDVGKEFGA-TNTFFYKPPCVSVYRCGGCCNSEELQCRNISTSYISKTLFEI | 175 |
| Rattus_norvegic | 122 | WRKTQCMPREVCIDVGKEFGAATNTFFKPPCVSVYRCGGCCNSEGLQCMNTSTGYLSKTLFEI  | 184 |
| Coturnix_coturn | 125 | WRKTQCMPREVCVDVGKEFGATTNTFFKPPCVSIYRCGGCCNSEGLQCMNISTNYISKTLFEI  | 187 |
| Mus_musculus    | 122 | WRKTQCMPREVCIDVGKEFGAATNTFFKPPCVSVYRCGGCCNSEGLQCMNTSTGYLSKTLFEI  | 184 |



# ClustalX/W/W2/Ω T-Coffee

Multiple sequence alignments are a very good tool to support decision making when designing proteins!

e.g. where to cut when you make chimeric proteins  
or  
deletion mutants



# More than Molecular Biology Software

- Image Editors: Go 3D with Blender!
- Text Editors vs. Desktop Publishing Software:  
If you are in charge of the layout, maybe consider NOT using a text editor.
- Bibliography Management: No need for EndNote if you have the free choice (which you mostly don't have...)



# Question

What do all of the following five images have in common?



Image by Bertrand Benoit



Image by Pratik Solanki



Image from the Sintel movie

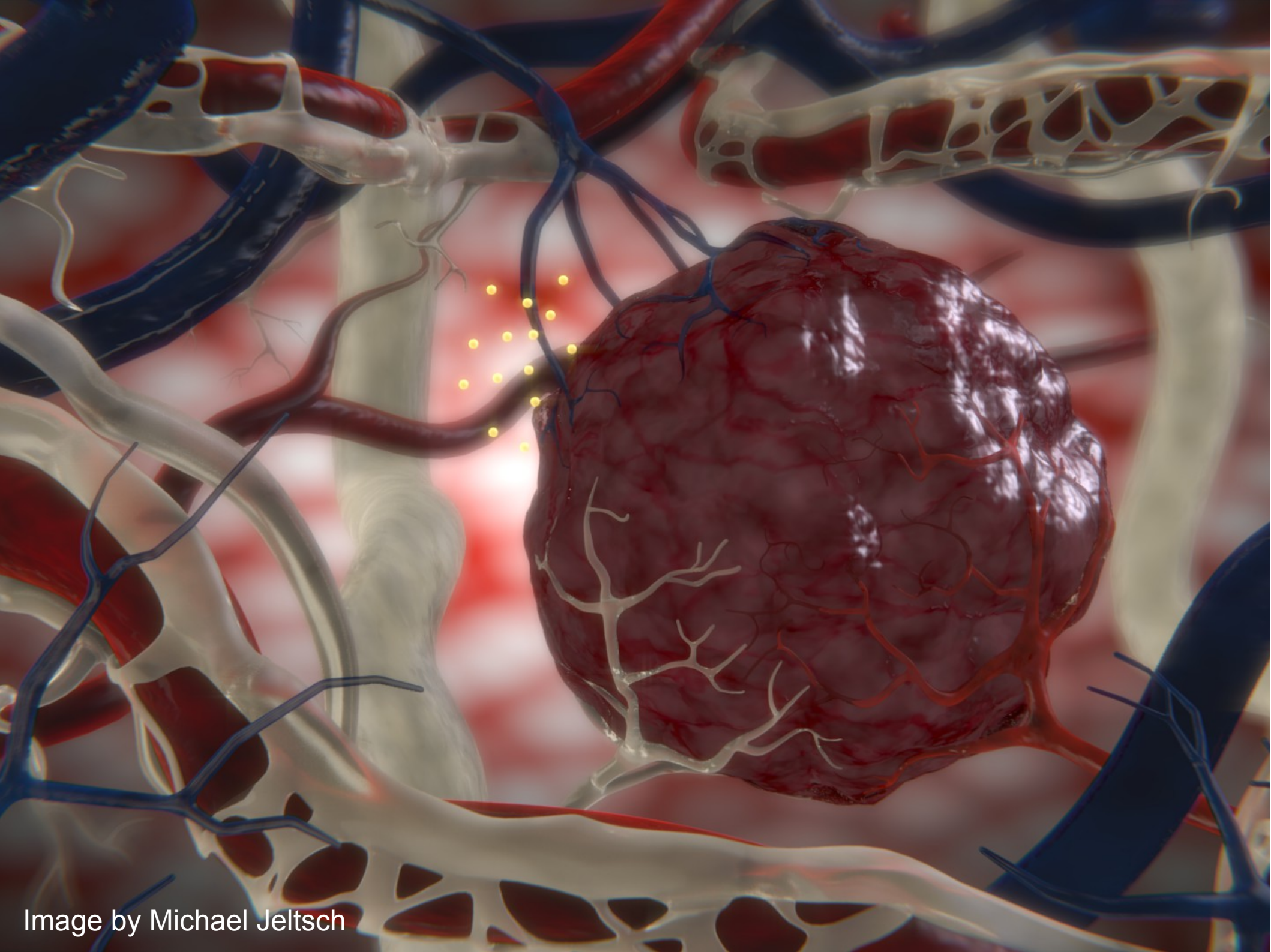
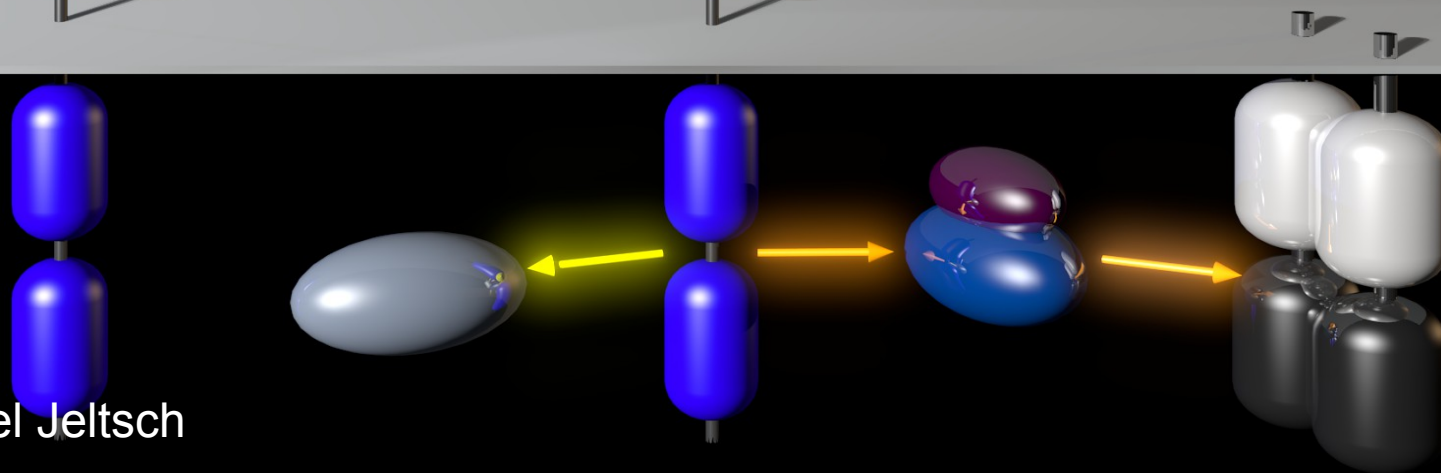
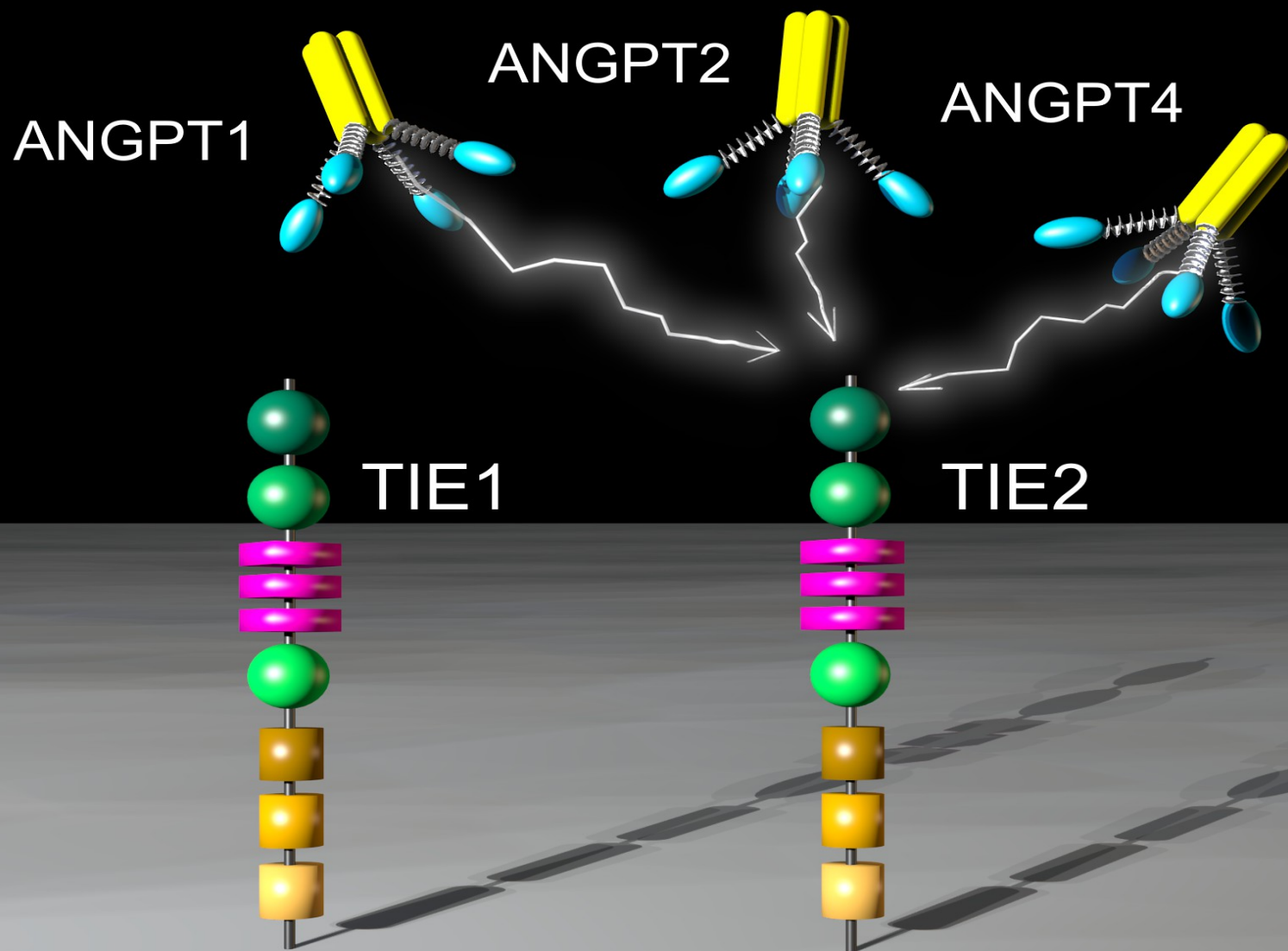


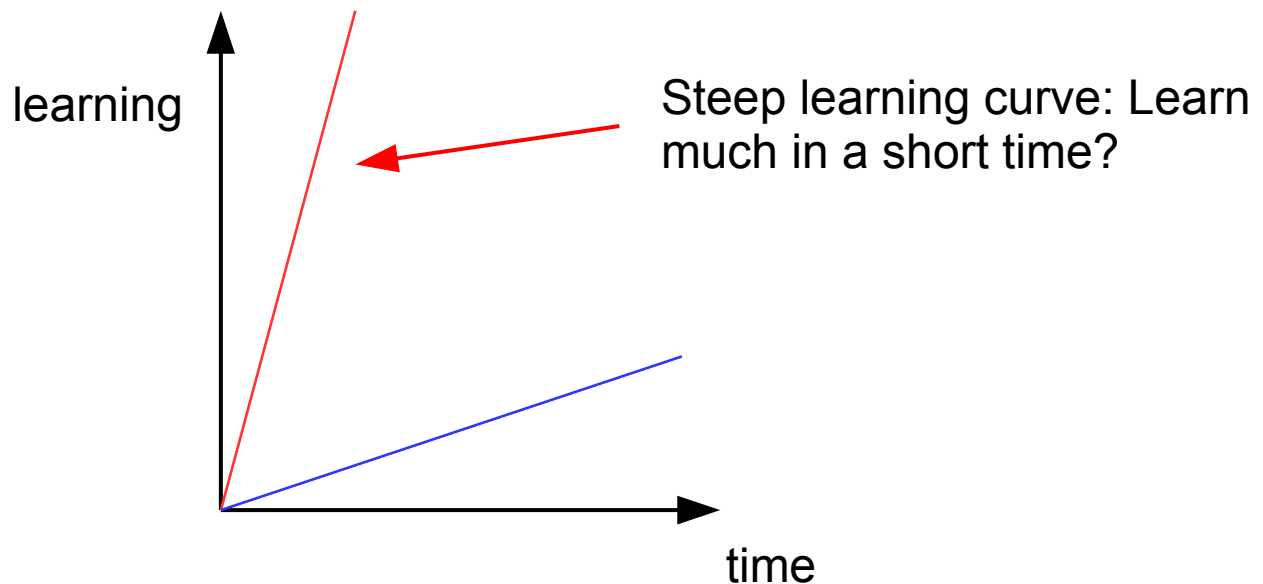
Image by Michael Jeltsch





# Answer

- All previous 5 images were made with Blender
- 3D drawing and animation (including movies)
- <http://www.blender.org/>
- Steep learning curve, but lots of free learning resources available





# Text Editors vs. Desktop Publishing Software

- Have you ever wondered why all publishers want your text in MS-Word format but your Figures and Tables separately?
- Standard for PhD theses in Physics and Mathematics ([ethesis.helsinki.fi](http://ethesis.helsinki.fi)): **TeX** in all its variations: **LaTeX**, **LyX**, etc.
- In Biology and Medicine, the use of TeX is conspicuously absent...
- The solution: DTP programs (**In-Design**, **Scribus**)?
- I used In-Design to prepare **my PhD thesis** (Scribus did not yet exist in 2002)



# Reference Management Systems

- EndNote or RefWorks licensed by Helsinki University
- For compatibility reasons you are probably forced to what your PI is used to
- Very good free alternative: [Zotero](#), implemented as a Firefox browser plugin with cloud storage, that communicates with MS-Word or LibreOffice.
- Stand-alone program for Safari, Chrome & Opera users
- [Mendeley](#) has in addition to reference management social media features; not anymore “open” after acquisition by Elsevier
- [Papers](#) is another good (commercial) alternative to EndNote





# Databases or where to find stuff

| What?            | Name                          | URL                                                                                                                                                                                                      | Example/Remarks                                                                       |
|------------------|-------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Proteins         | Uniprot                       | <a href="http://www.uniprot.org">www.uniprot.org</a>                                                                                                                                                     | VEGF-C                                                                                |
| DNA<br>Sequences | EMBL/<br>GenBank/<br>DDBJ     | <a href="http://www.ebi.ac.uk/ena">www.ebi.ac.uk/ena</a><br><a href="http://www.ncbi.nlm.nih.gov/genbank">www.ncbi.nlm.nih.gov/genbank</a><br><a href="http://www.ddbj.nig.ac.jp">www.ddbj.nig.ac.jp</a> | VEGF-C<br>VEGF-C                                                                      |
| Genomes          | Ensemble                      | <a href="http://www.ensembl.org/">www.ensembl.org/</a>                                                                                                                                                   | VEGFC                                                                                 |
| cDNA             | MGC                           | <a href="http://mgc.nci.nih.gov">mgc.nci.nih.gov</a>                                                                                                                                                     |                                                                                       |
| Literature       | PubMed<br>Ovid Medline        | <a href="http://www.ncbi.nlm.nih.gov/pubmed">www.ncbi.nlm.nih.gov/pubmed</a><br><a href="http://www.terkko.helsinki.fi/medline">www.terkko.helsinki.fi/medline</a>                                       | requires subscription<br>(available from within UH)                                   |
|                  | GoogleScholar<br>ResearchGate | <a href="http://scholar.google.com">scholar.google.com</a><br><a href="http://www.researchgate.net">www.researchgate.net</a>                                                                             | “LinkedIn” for researchers;<br>ask authors for articles,<br>mail corresponding author |
| Images           | CCSearch                      | <a href="http://search.creativecommons.org">search.creativecommons.org</a>                                                                                                                               | mouse                                                                                 |
| Patents          | US PatentDB<br>EPO            | <a href="http://patft.uspto.gov/netahtml/PTO">patft.uspto.gov/netahtml/PTO</a><br><a href="http://worldwide.espacenet.com">worldwide.espacenet.com</a>                                                   | Alitalo, Kari<br>Alitalo, Kari                                                        |



# Online tools

- **Allmost all of the programs from the previous slides have online equivalents!**