# Instructions

Allowed tools: Paper, pen and dictionary.

Write name and personal number on here:

……………………………………………….

Write the secret number given to you by Christoph/Mirco on **EVERY** page and here:

……………………………………………….

If you want your result sent to you be email write your email address here:

………………………………………………

Maximum number of points on the exam is 100. Another 100 points is given at the online exam. The grades will be based 50% of the online and 50% of the written exam. To pass the course you need at least 40 points at each of the exams.

Question 1 at the exam can be skipped if the paper presentation is finished in time. This is true for all 8 that presented the project on Tuesday Sep 24.

Grades are accordingly to the following

* A 92-100%
* B 80-91%
* C 68-79%
* D 56-67%
* E 50-55%
* Fx 45-49%
* F 0-44%

Each question should be answered within the provided space, if you need more space write on the back of the paper.

You can answer in either English or Swedish

Arne can be reached at 070-6951045 for questions, and will pass by at a few times.

Short questions (10 p per question):

1a. Place the following amino acids on the plot below: Arginine (R), Leucine (L), Glutamine (Q), Glycine (G) and Phenylalanine (F). (5 p)



**1b.** In the figure below (which shows the integral membrane protein aquaporin), there are residues colored in green and in blue. What types of residues are these? (5p)



2. Describe how the membrane-protein topology prediction method TOPPRED works based on the two figures below (10 p).





3. Draw an artificial neural network that can solve the XOR problem, i.e. a network that given the following given inputs produce this output (10p):

|  |  |  |
| --- | --- | --- |
| Input 1 | Input 2 | Output |
| 0 | 0 | 0 |
| 1 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 1 | 0 |

4a) The uniprot database consists of two parts, Swiss-Prot and TrEMBL how do they differ (5p)?

4b) Below you find a part of an entry in uniprot. What does it tell about the protein?

ID AQY1\_YEASX Reviewed; 327 AA.

AC P0CD92; O74680; P53386;

DT 02-MAR-2010, integrated into UniProtKB/Swiss-Prot.

DT 02-MAR-2010, sequence version 1.

DT 29-MAY-2013, entry version 22.

DE RecName: Full=Aquaporin-1;

GN Name=AQY1; Synonyms=AQY1-1;

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Dikarya; Ascomycota; Saccharomycotina;

OC Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TaxID=4932;

…..

PE 2: Evidence at transcript level;

KW Cell membrane; Endoplasmic reticulum; Membrane; Repeat; Transmembrane; KW Transmembrane helix; Transport.

FT CHAIN 1 327 Aquaporin-1.

FT /FTId=PRO\_0000391661.

FT TOPO\_DOM 1 48 Cytoplasmic (By similarity).

FT TRANSMEM 49 69 Helical; Name=1; (Potential).

FT TOPO\_DOM 70 91 Extracellular (By similarity).

FT TRANSMEM 92 112 Helical; Name=2; (Potential).

FT TOPO\_DOM 113 136 Cytoplasmic (By similarity).

FT TRANSMEM 137 157 Helical; Name=3; (Potential).

FT TOPO\_DOM 158 176 Extracellular (By similarity).

FT TRANSMEM 177 197 Helical; Name=4; (Potential).

FT TOPO\_DOM 198 203 Cytoplasmic (By similarity).

FT TRANSMEM 204 224 Helical; Name=5; (Potential).

FT TOPO\_DOM 225 248 Extracellular (By similarity).

FT TRANSMEM 249 269 Helical; Name=6; (Potential).

FT TOPO\_DOM 270 327 Cytoplasmic (By similarity).

FT MOTIF 118 120 NPA 1.

FT MOTIF 230 232 NPA 2.

5. It was recently reported by Svante Pääbo and co-workers that it was likely that Neanderthals had contributed to genomes of all non-African humans. Describe how this conclusion was obtained (5p) and how and when he proposed that this was most likely to have happened (5p).

6. Describe the process of a homology modeling method (5p). What is in general the most important feature for the final quality of a model when you do homology modeling (5p)?

7. Dynamic programming can be used to obtain optimal alignments between two genes or proteins. Describe how the algorithm works (5p) and what is the difference between global (Needleman-Wunsch) and local (Smith-Waterman) alignments (5p).

8. Describe the psi-blast algorithm (5p). What features was it that made it one of the most used methods in bioinformatics (5p).

*Long question (20p).*

9. Today the best secondary structure prediction methods use artificial neural networks and multiple sequence alignments. You have been given the task to design a novel predictor of secondary structure predictions based on a hidden Markov model (HMM)

1. Draw a simple scheme how current state of the art secondary structure prediction methods work.
2. Draw how a secondary structure predictor using a hidden Markov model could look like.
3. In comparison with the best what information do you not use in a simple HMM model.
4. Can you propose some method to include this information?