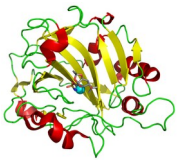
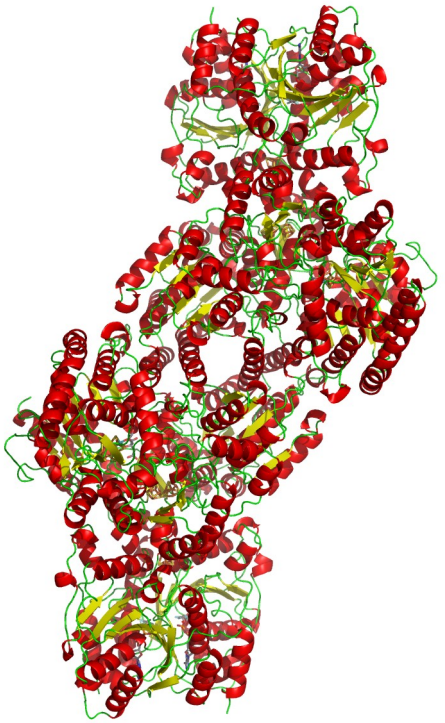
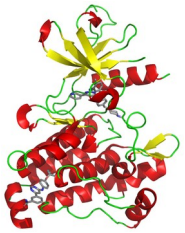
Protein databank (*1971)
http://www.pdb.org



karbonáthydrolyasa

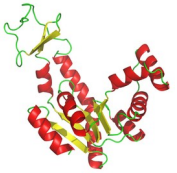




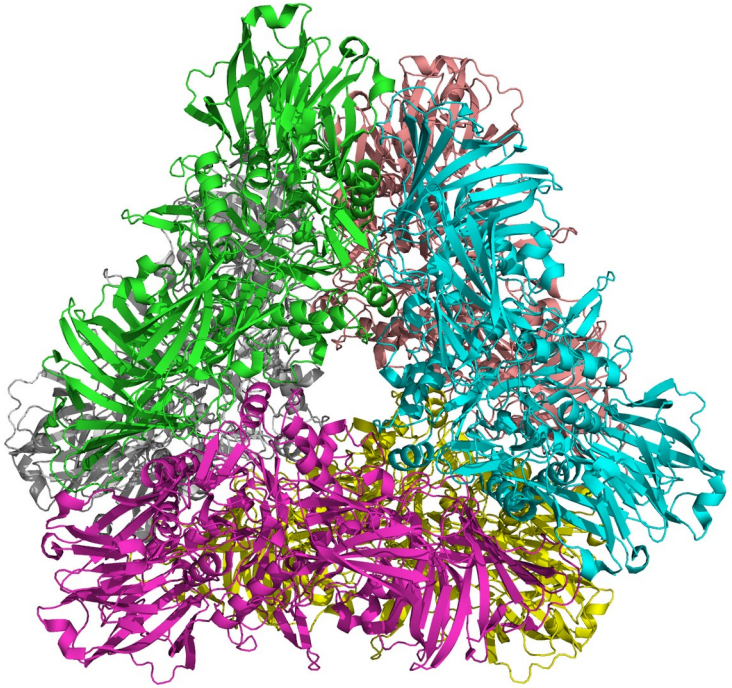
Abl

alkoholdehydrogenasa

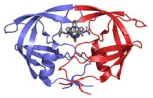
nitrogenasa



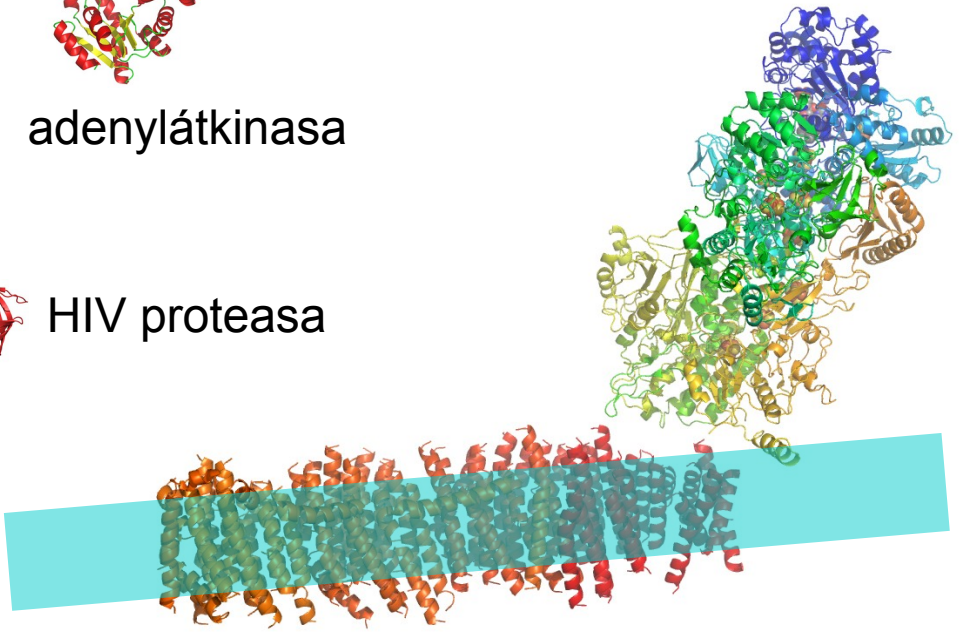
adenylátkinasa



β-galaktosidasa

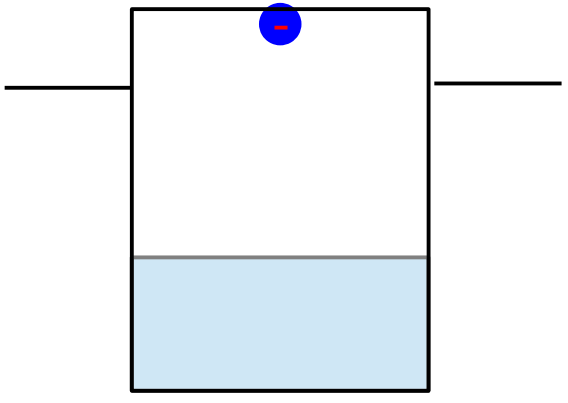


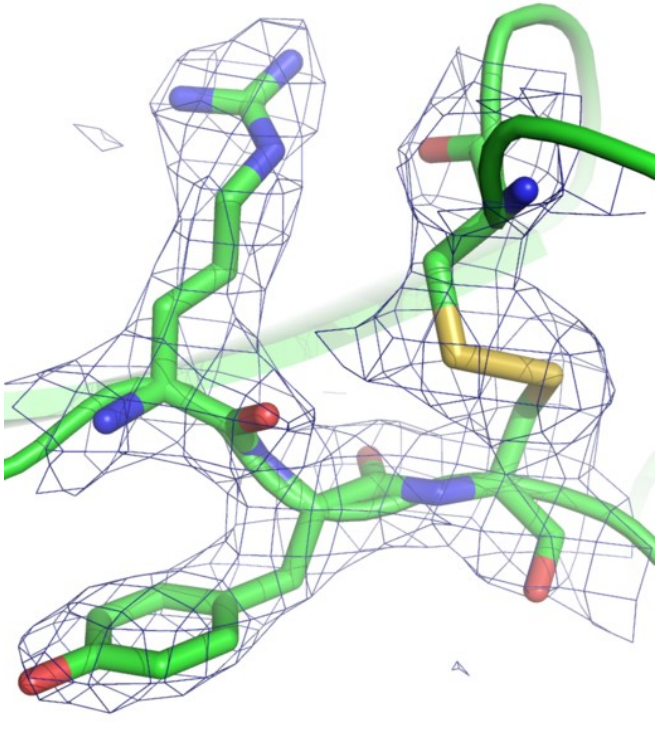
HIV proteasa

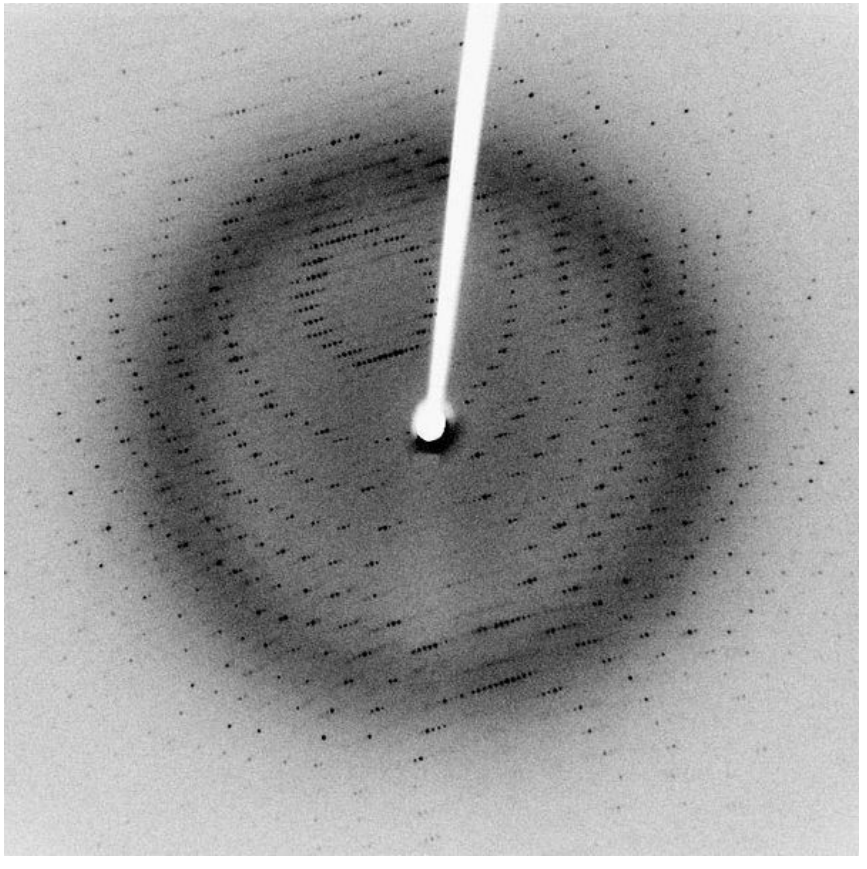


ubichinonreduktasa (kotvený komplex I)

Struktura proteinů:
Rentgenová strukturní krystalorafie

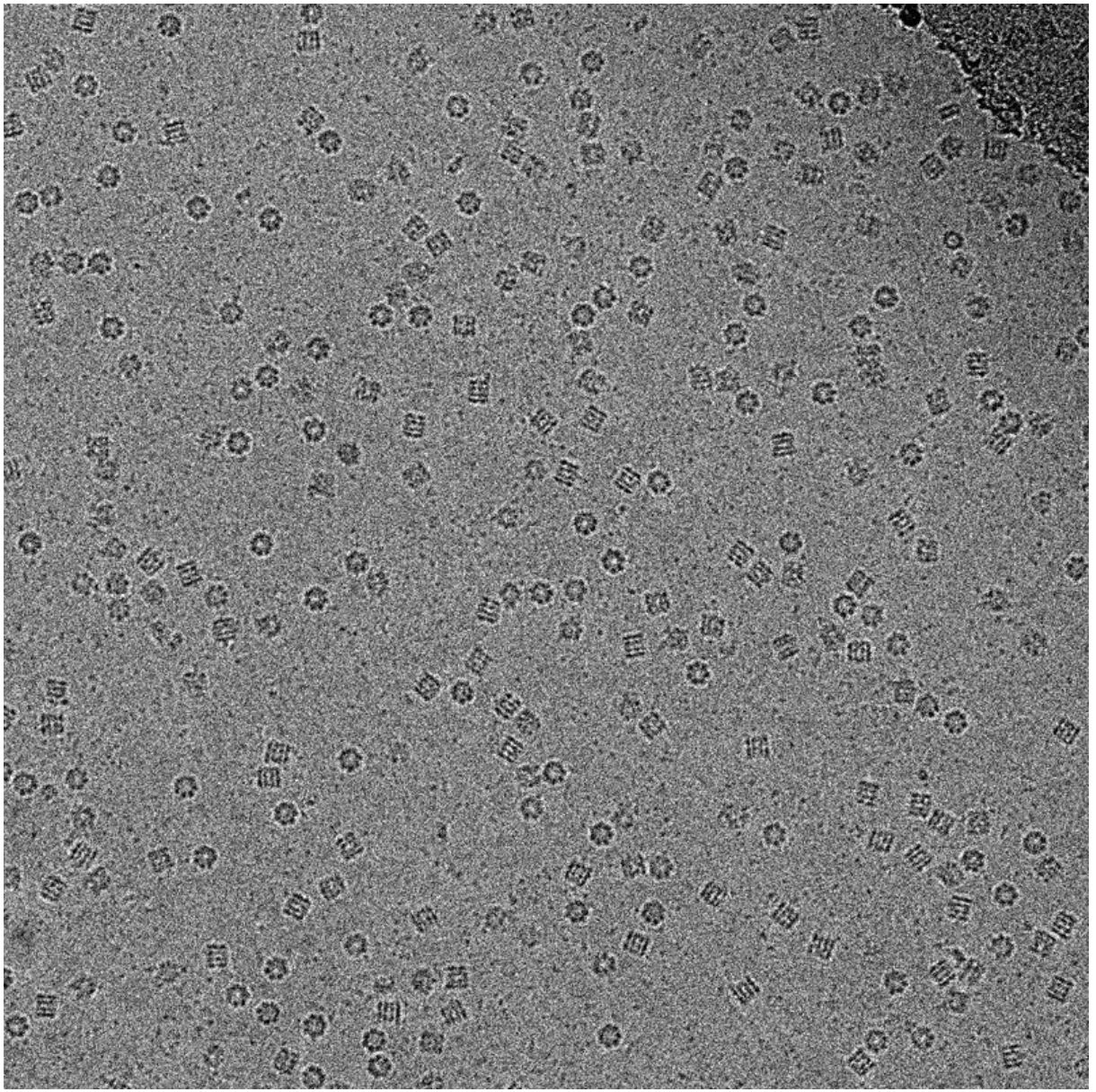






Struktura proteinů:
NMR

Struktura proteinů:
CryoEM



Předpověď struktur proteinů:

Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu

Fold recognition, threading
- proteiny mohou mít podobnou strukturu a nemusí mít (moc) podobnou sekvenci

Ab initio, de novo
- nativní struktura má určité vlastnosti

Strojové učení
- je možné naučit se „jazyk“ proteinových sekvencí a překládat je do 3D struktur

Simulace sbalování
- protein se dokáže sbalit do nativní struktury

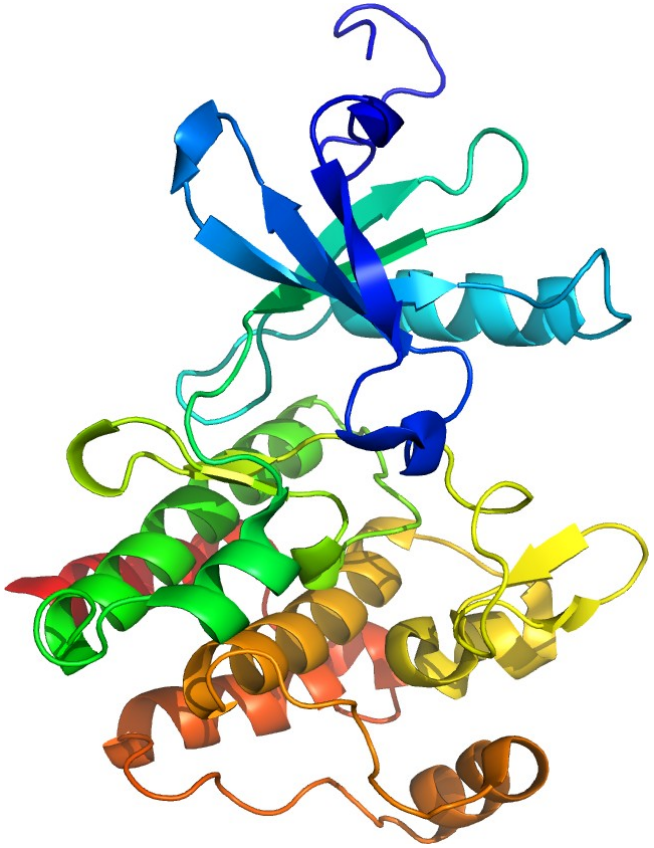
Předpověď struktur proteinů:
Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu

Postup:
1. nalezneme protein(y) se známou prostorovou strukturou
   a s podobnou sekvencí našemu proteinu

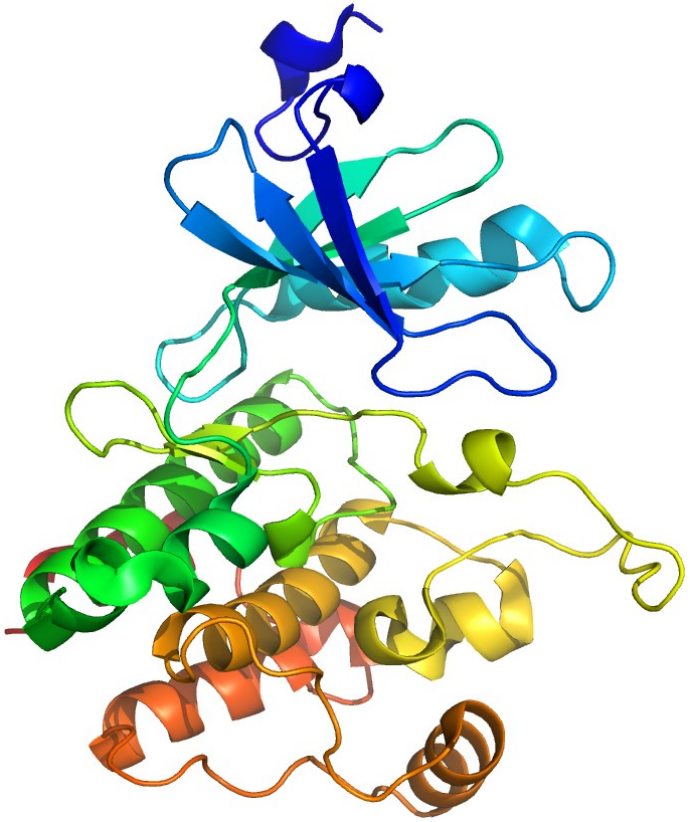
2. vytvoříme zarovnání sekvencí

3. vytvoříme model struktury našeho proteinu

Předpověď struktur proteinů:
Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu



Abl (1IEP)



Lck (2PL0)

Předpověď struktur proteinů:
Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu

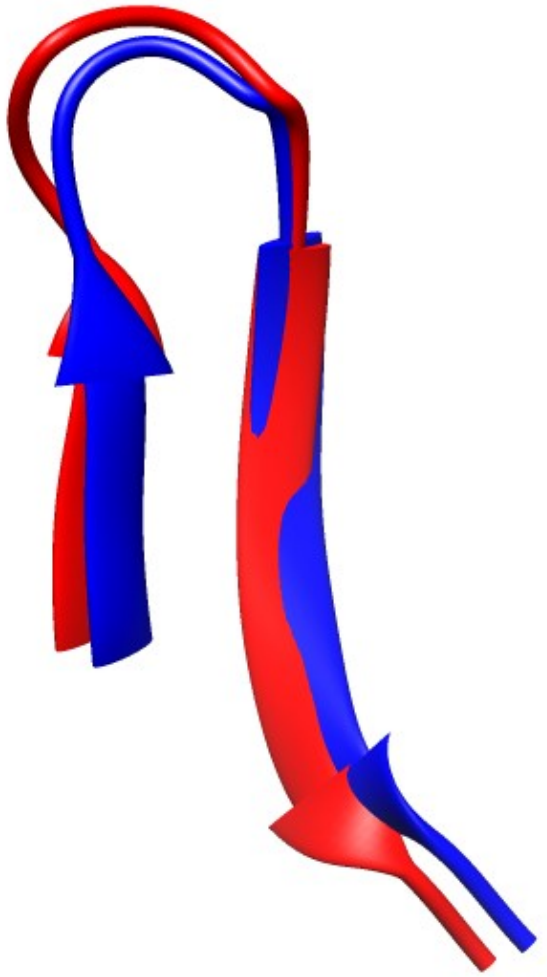
```
>Lck
GSHMQTQKPQKPWWEDEWEVPRETLKLVERLGAGQFGEVWMGYYNGHTKVAVKSLKQGSMSPDAFLAEANLMKQLQHQRL
VRLYAVVTQEPIYIITEYMENGSLVDFLKTPSGIKLTINKLLDMAAQIAEGMAFIEERNYIHRDLRAANILVSDTLSCKI
ADFGLARLIEDNEYTAREGAKFPIKWTAPEAINYGTFTIKSDVWSFGILLTEIVTHGRIPYPGMTNPEVIQNLERGYRMV
RPDNCPEELYQLMRLCWKERPEDRPTFDYLRSVLEDFFTATEGQYQPQP
```

Identita 48 %

```
Query  16   DEWEVPRETLKLVERLGAGQFGEVWMGYYNGHT-KVAVKSLKQGSMSPDAFLAEANLMKQ  74
            D+WE+ R  + +  +LG GQ+GEV+ G +  ++  VAVK+LK+ +M  + FL EA +MK+
Sbjct  6    DKWEMERTDITMKHKLGGGQYGEVYEGVWKKYSLTVAVKTLKEDTMEVEEFLKEAAVMKE  65

Query  75   LQHQRLVRLYAVVTQEP-IYIITEYMENGSLVDFLKTPSGIKLTINKLLDMAAQIAEGMA  133
            ++H  LV+L  V T+EP  YIITE+M  G+L+D+L+  +  ++    LL MA QI+  M
Sbjct  66   IKHPNLVQLLGVCTREPPFYIITEFMTYGNLLDYLRECNRQEVNAVVLLYMATQISSAME  125

Query  134  FIEERNYIHRDLRAANILVSDTLSCKIADFGLARLIEDNEYTAREGAKFPIKWTAPEAIN  193
            ++E++N+IHRDL A N LV +    K+ADFGL+RL+  + YTA  GAKFPIKWTAPE++
Sbjct  126  YLEKKNFIHRDLAARNCLVGENHLVKVADFGLSRLMTGDTYTAHAGAKFPIKWTAPESLA  185

Query  194  YGTFTIKSDVWSFGILLTEIVTHGRIPYPGMTNPEVIQNLERGYRMVRPDNCPEELYQLM  253
            Y  F+IKSDVW+FG+LL EI T+G  PYPG+   +V + LE+ YRM RP+ CPE++Y+LM
Sbjct  186  YNKFSIKSDVWAFGVLLWEIATYGMSPYPGIDLSQVYELLEKDYRMERPEGCPEKVYELM  245

Query  254  RLCWKERPEDRPTFDYLRSVLEDFF  278
            R CW+  P DRP+F  +    E  F
Sbjct  246  RACWQWNPSDRPSFAEIHQAFETMF  270
```

Předpověď struktur proteinů:
Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu

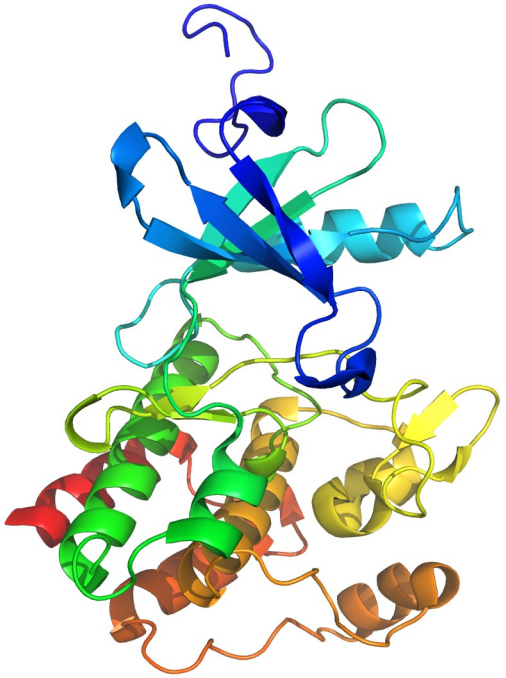
```
Lck   AVVTQEP-IYIITEY
       V T+EP  YIITE+
Abl   GVCTREPPFYIITEF
```



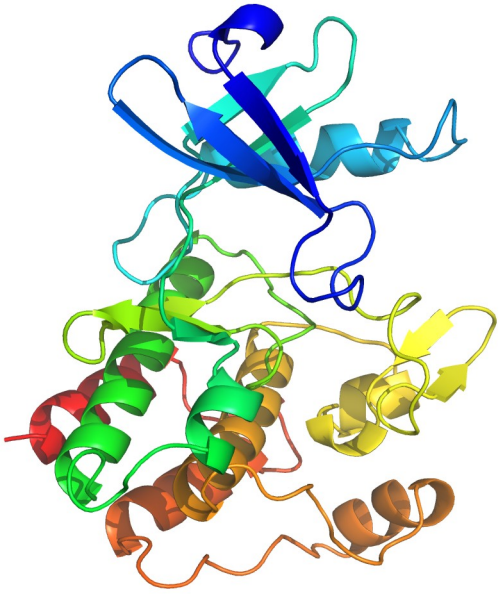
Předpověď struktur proteinů:
Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu

```
C; A sample alignment in the PIR format; used in tutorial

>P1;1IEP
structureX:1IEP:233  :A:498  :A::::
---------------DKWEMERTDITMKHKLGGGQYGEVYEGVWKKYSLTVAVKTLKEDT
MEVEEFLKEAAVMKEIKHPNLVQLLGVCTREPPFYIITEFMTYGNLLDYLRECNRQEVSA
VVLLYMATQISSAMEYLEKKNFIHRDLAARNCLVGENHLVKVADFGLSRLMTGDTYTAHA
GAKFPIKWTAPESLAYNKFSIKSDVWAFGVLLWEIATYGMSPYPGIDLSQVYELLEKDYR
MERPEGCPEKVYELMRACWQWNPSDRPSFAEIHQAFETMFQ*

>P1;lck
sequence:lck:1    : :@   : ::::
---------------DEWEVPRETLKLVERLGAGQFGEVWMGYYNGHT-KVAVKSLKQGS
MSPDAFLAEANLMKQLQHQRLVRLYAVVTQEP-IYIITEYMENGSLVDFLKTPSGIKLTI
NKLLDMAAQIAEGMAFIEERNYIHRDLRAANILVSDTLSCKIADFGLARLIEDNEYTARE
GAKFPIKWTAPEAINYGTFTIKSDVWSFGILLTEIVTHGRIPYPGMTNPEVIQNLERGYR
MVRPDNCPEELYQLMRLCWKERPEDRPTFDYLRSVLEDFFT*
```

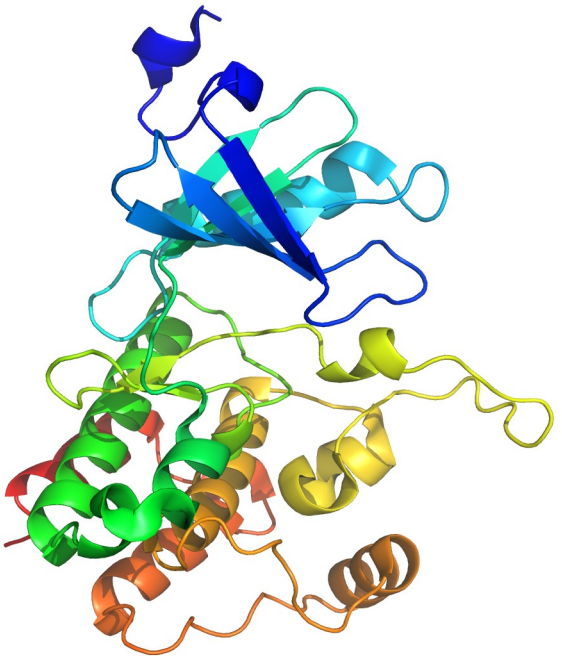
Předpověď struktur proteinů:
Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu



Abl (1IEP)

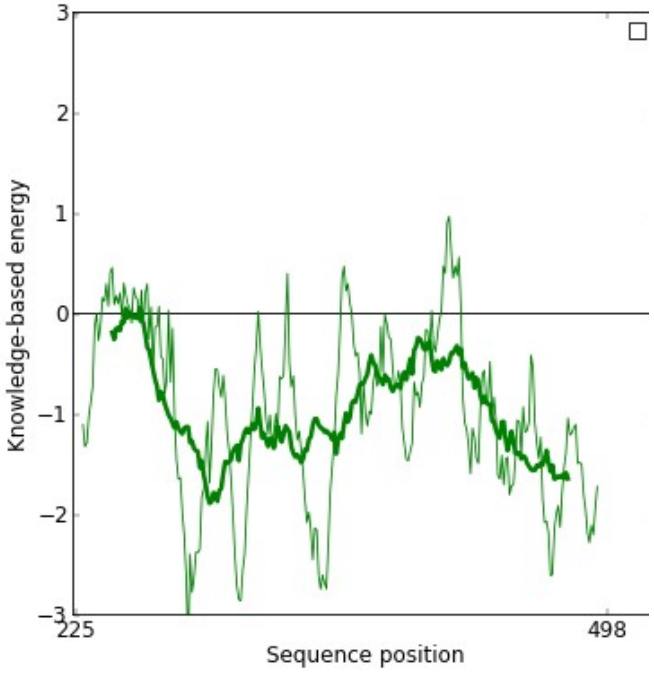


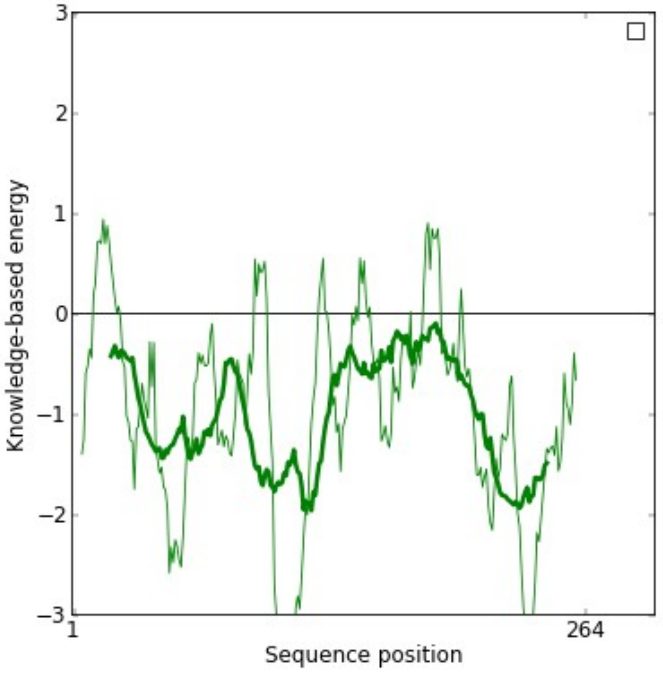
Lck (model)

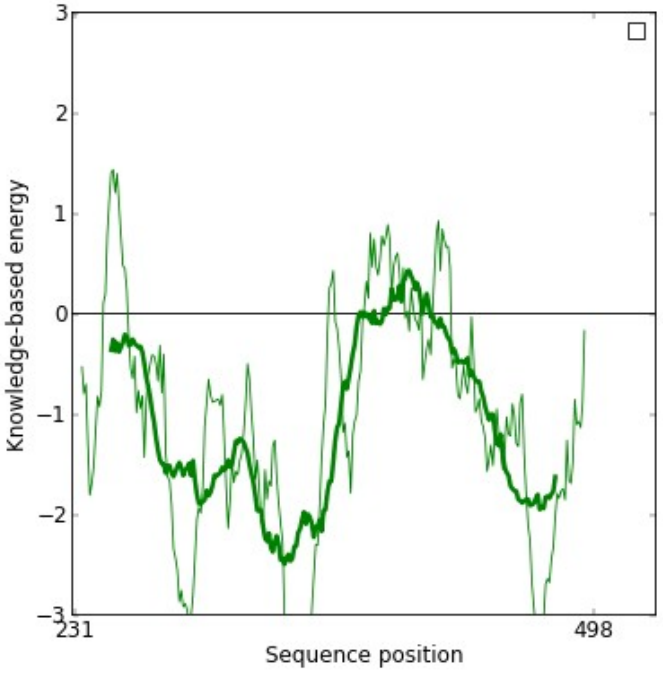


Lck (2PL0)

Předpověď struktur proteinů:
Homologní modelování
- proteiny s podobnou sekvencí mají podobnou strukturu
ProsaII





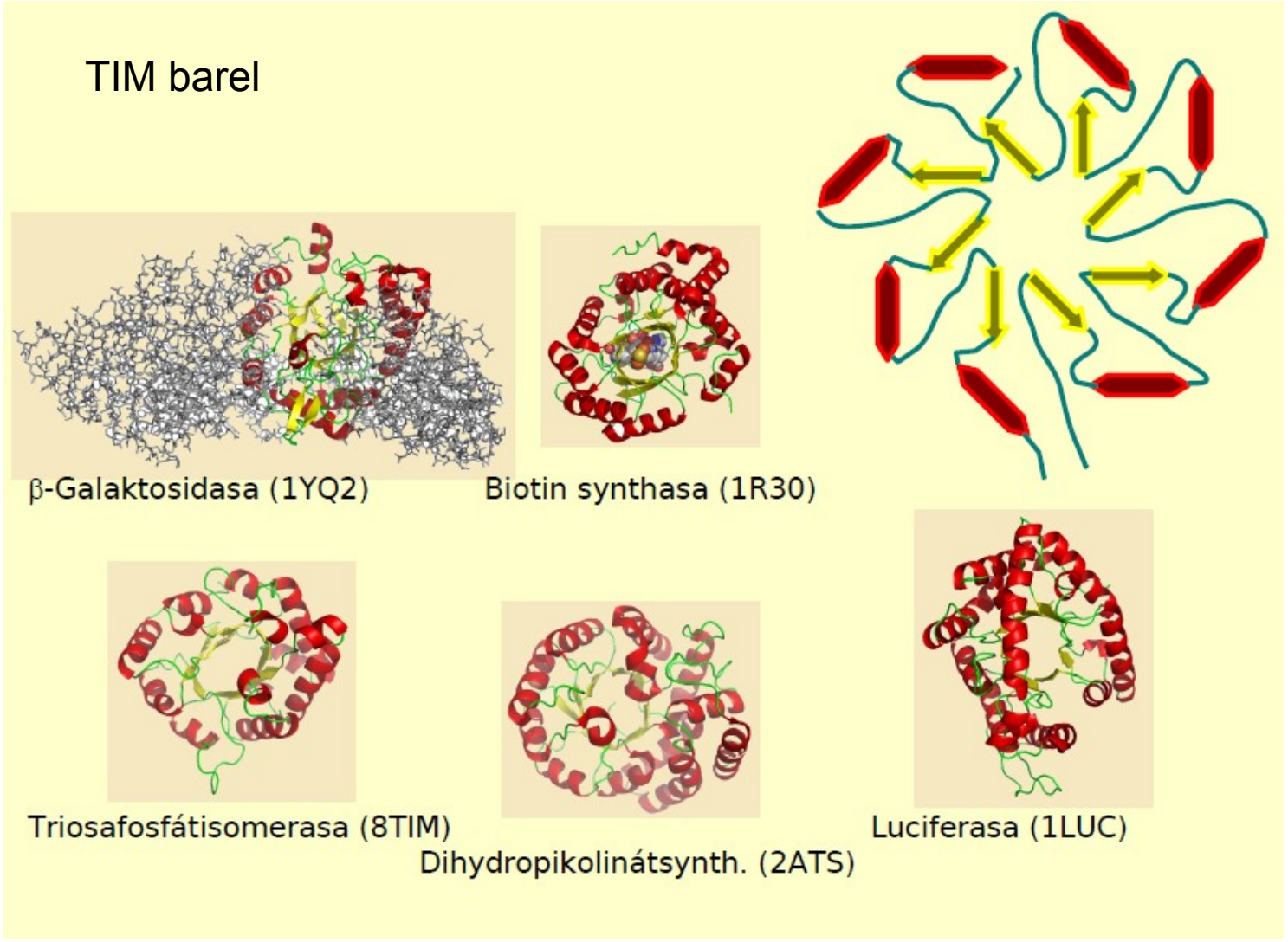


Abl (1IEP)

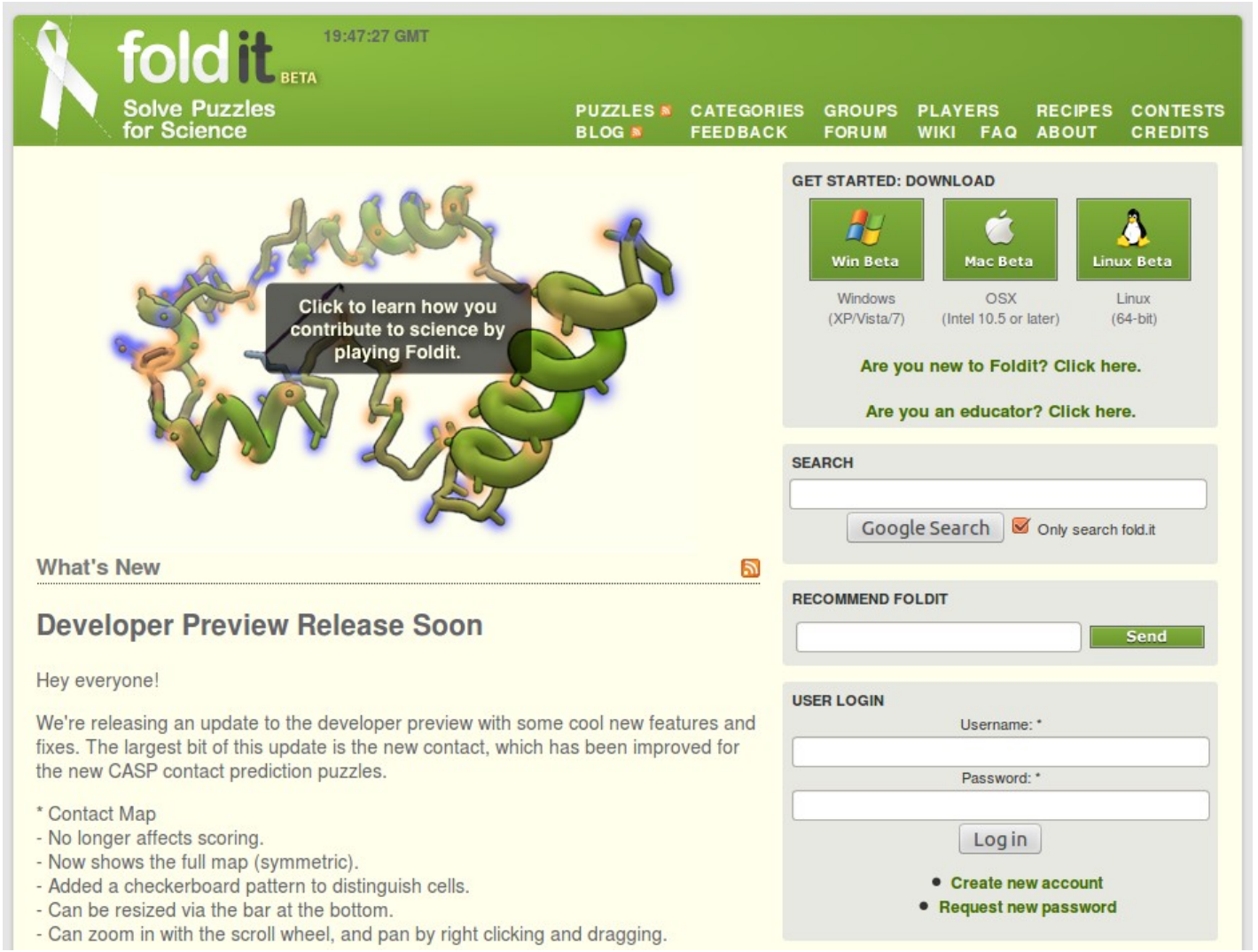
Lck (model)

Lck (2PL0)

Předpověď struktur proteinů:
Fold recognition, threading
- proteiny mohou mít podobnou strukturu a nemusí mít (moc) podobnou sekvenci



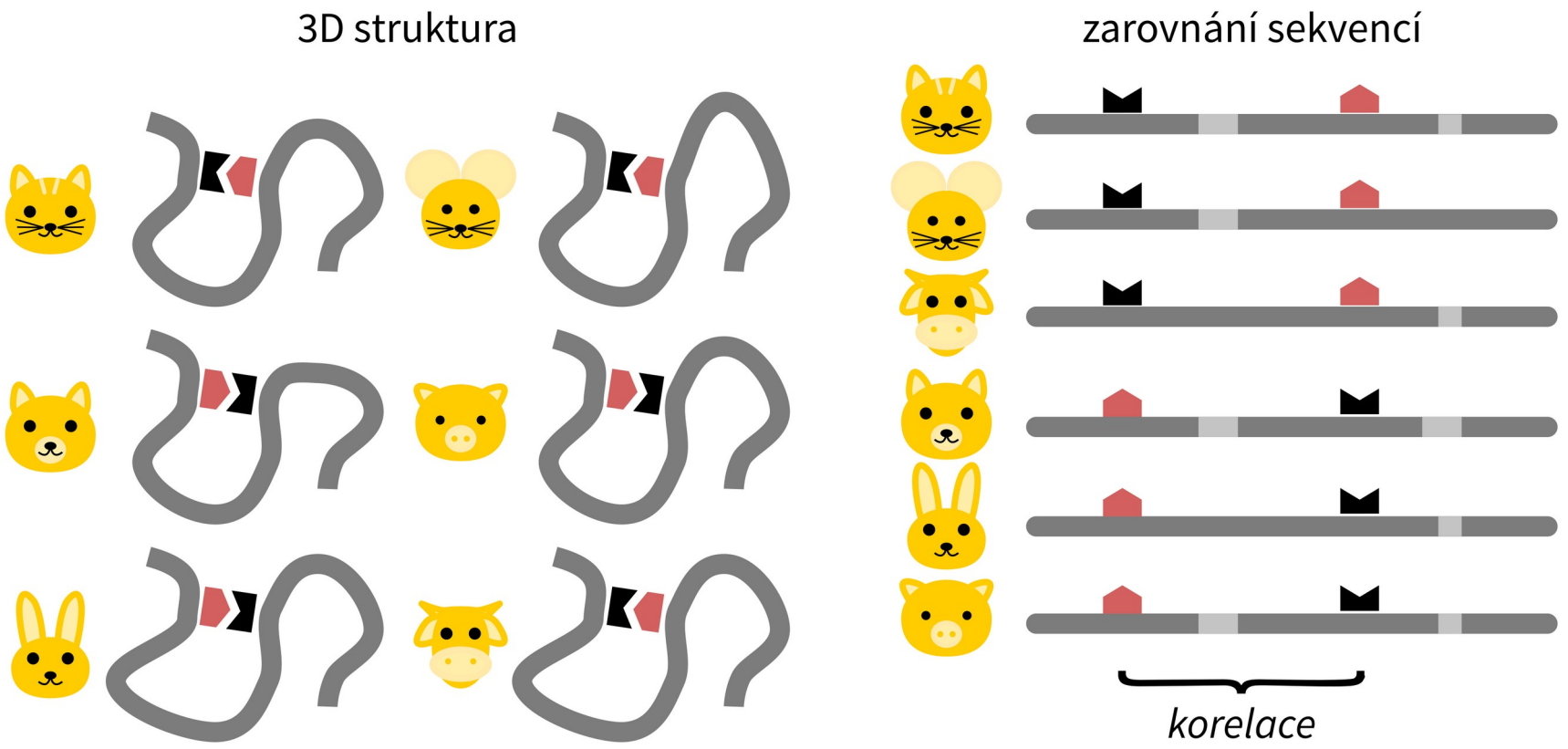
Předpověď struktur proteinů:
Ab initio, de novo
- nativní struktura má určité vlastnosti



http://fold.it

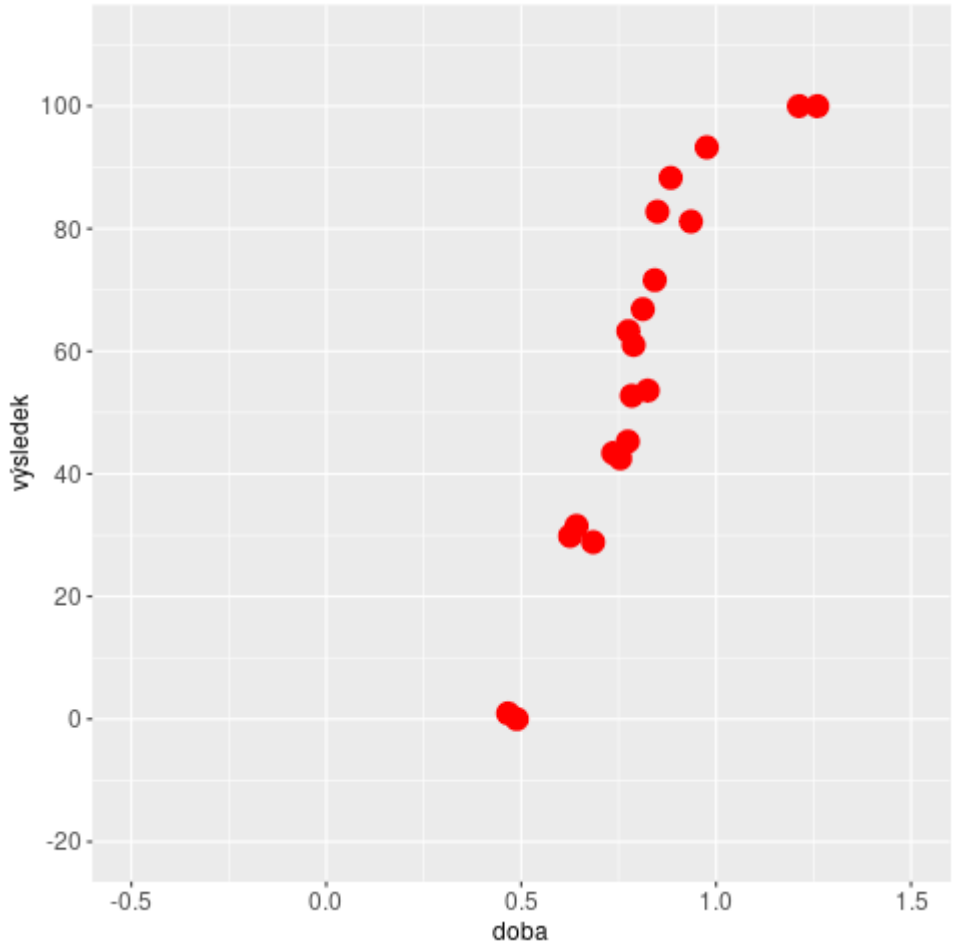
Alphafold:

- koevoluce



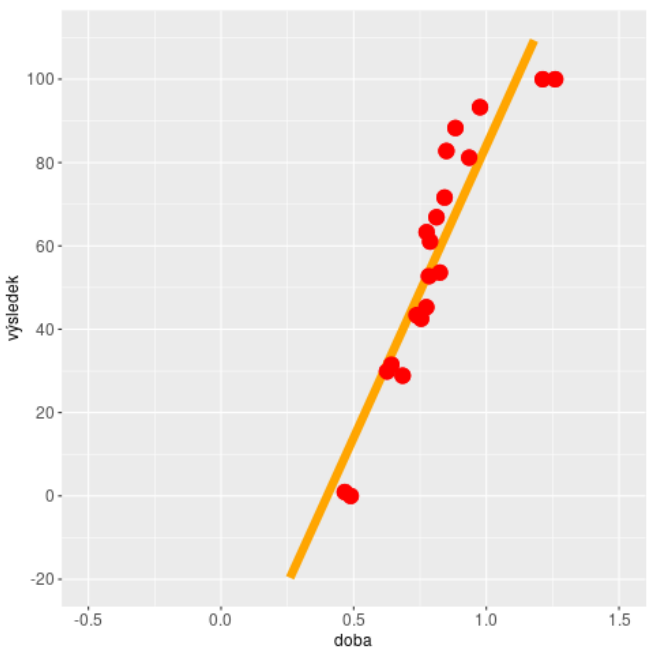
Alphafold:

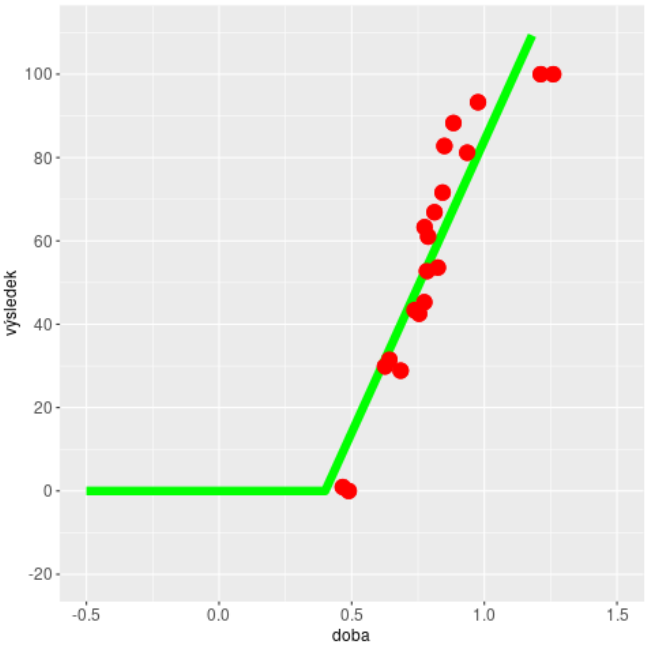
- neuronové sítě

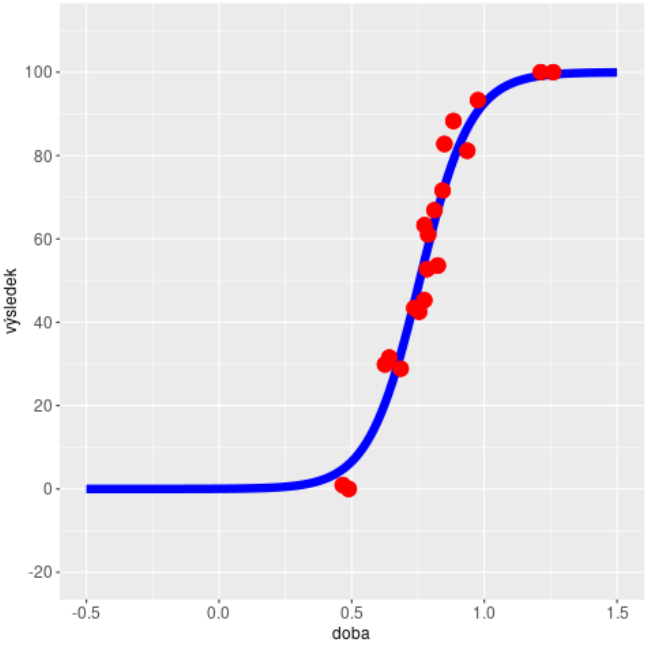


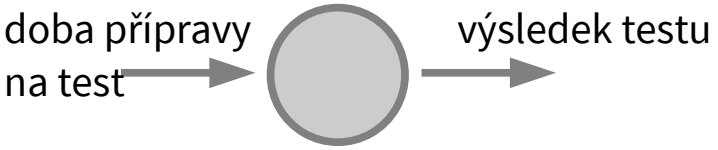
Alphafold:

- neuronové sítě



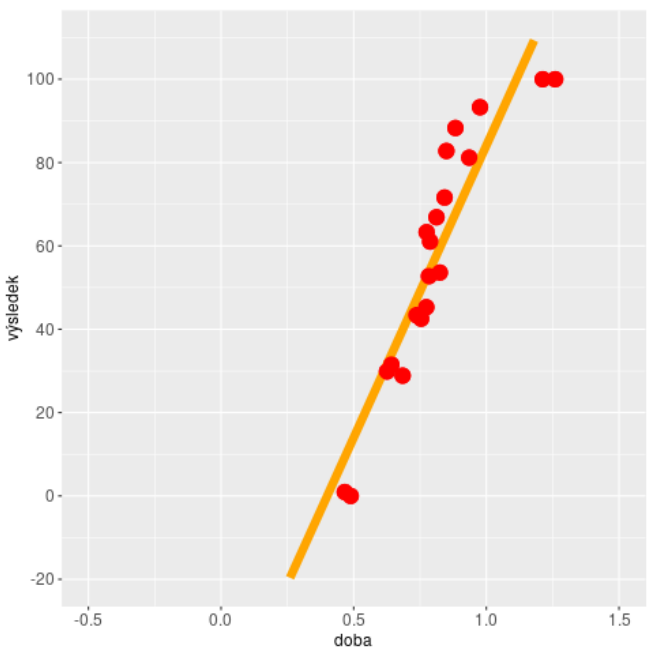


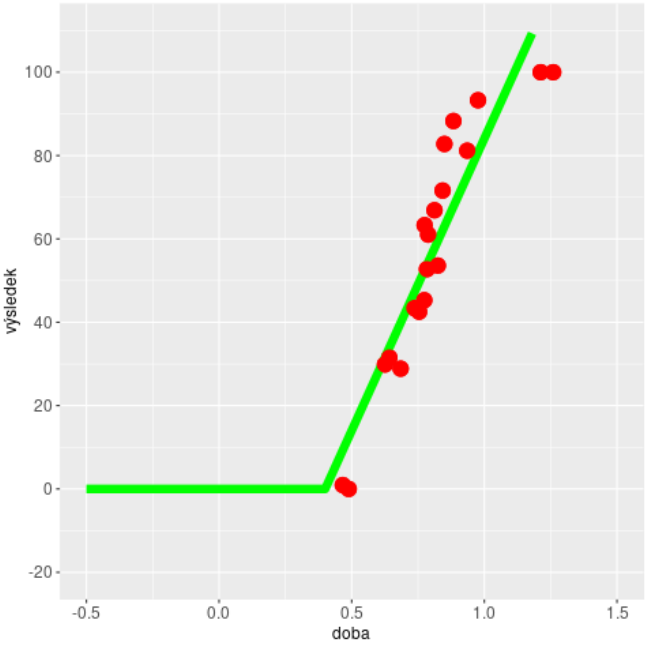


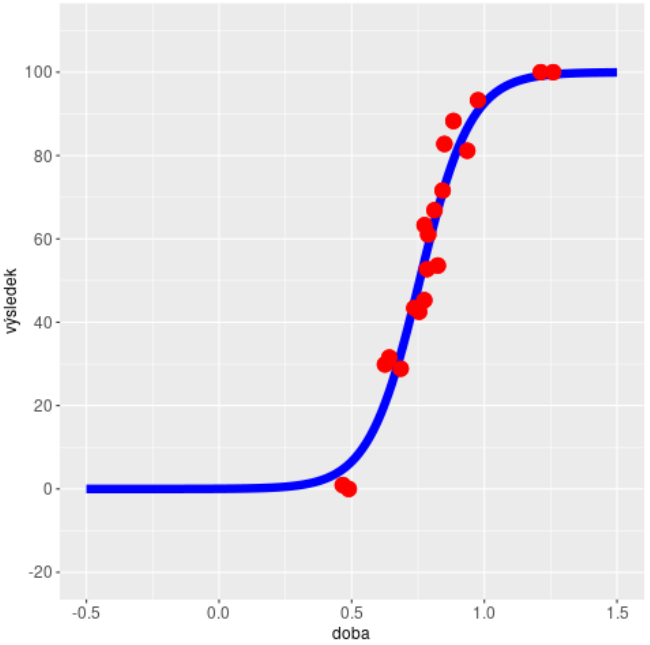


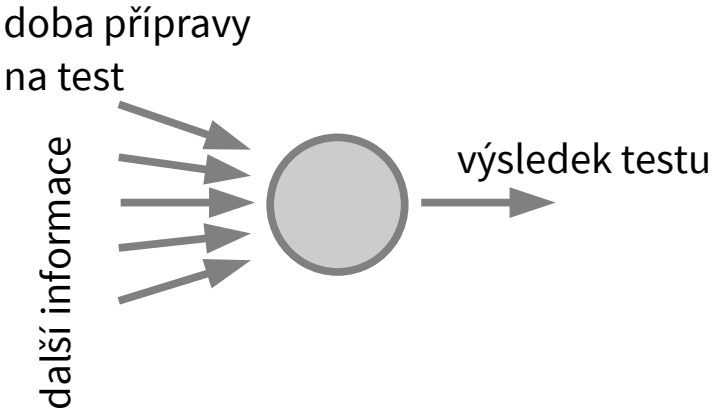
Alphafold:

- neuronové sítě



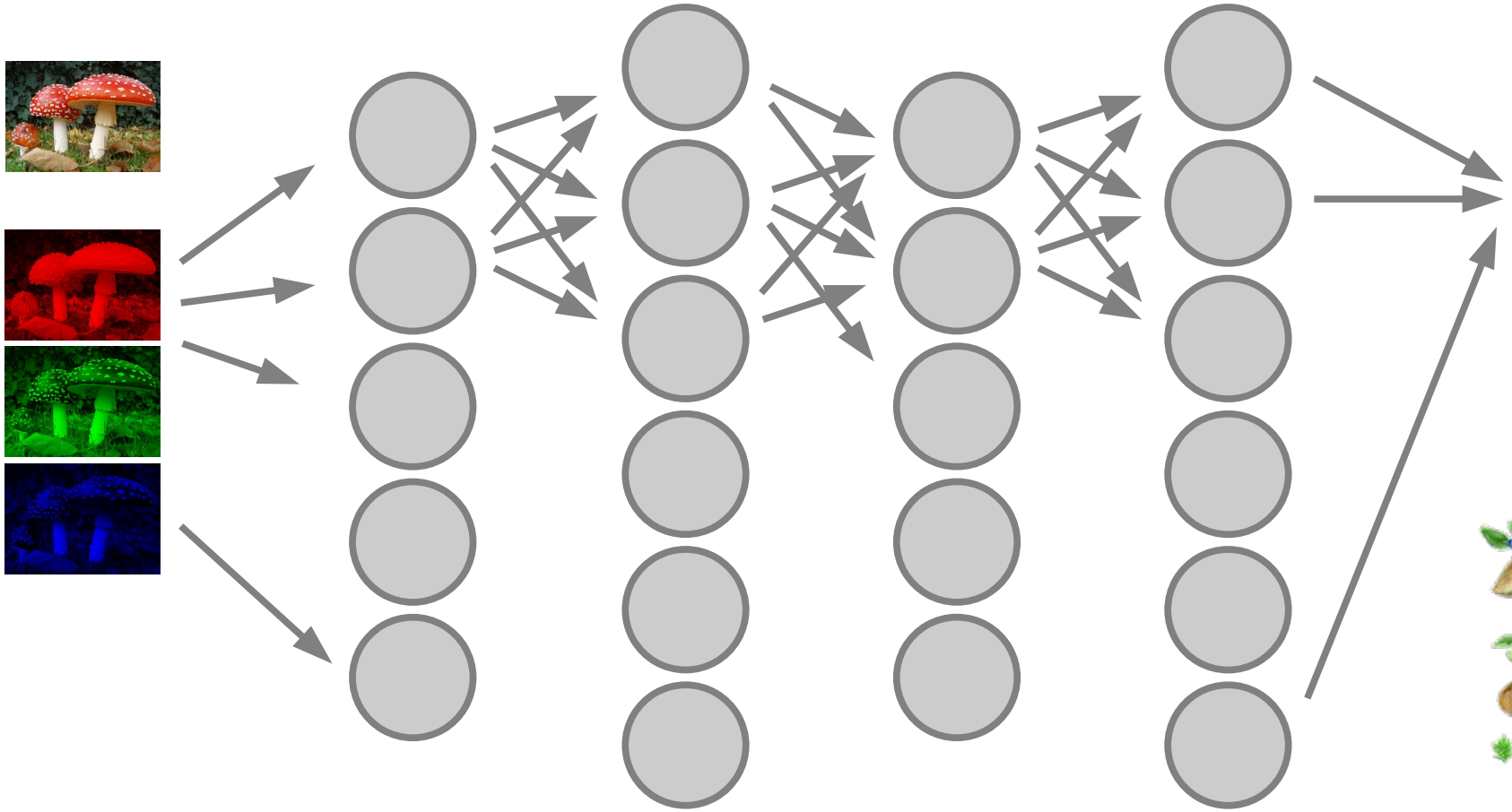






Alphafold:

- neuronové sítě



***Amanita muscaria* 99%**

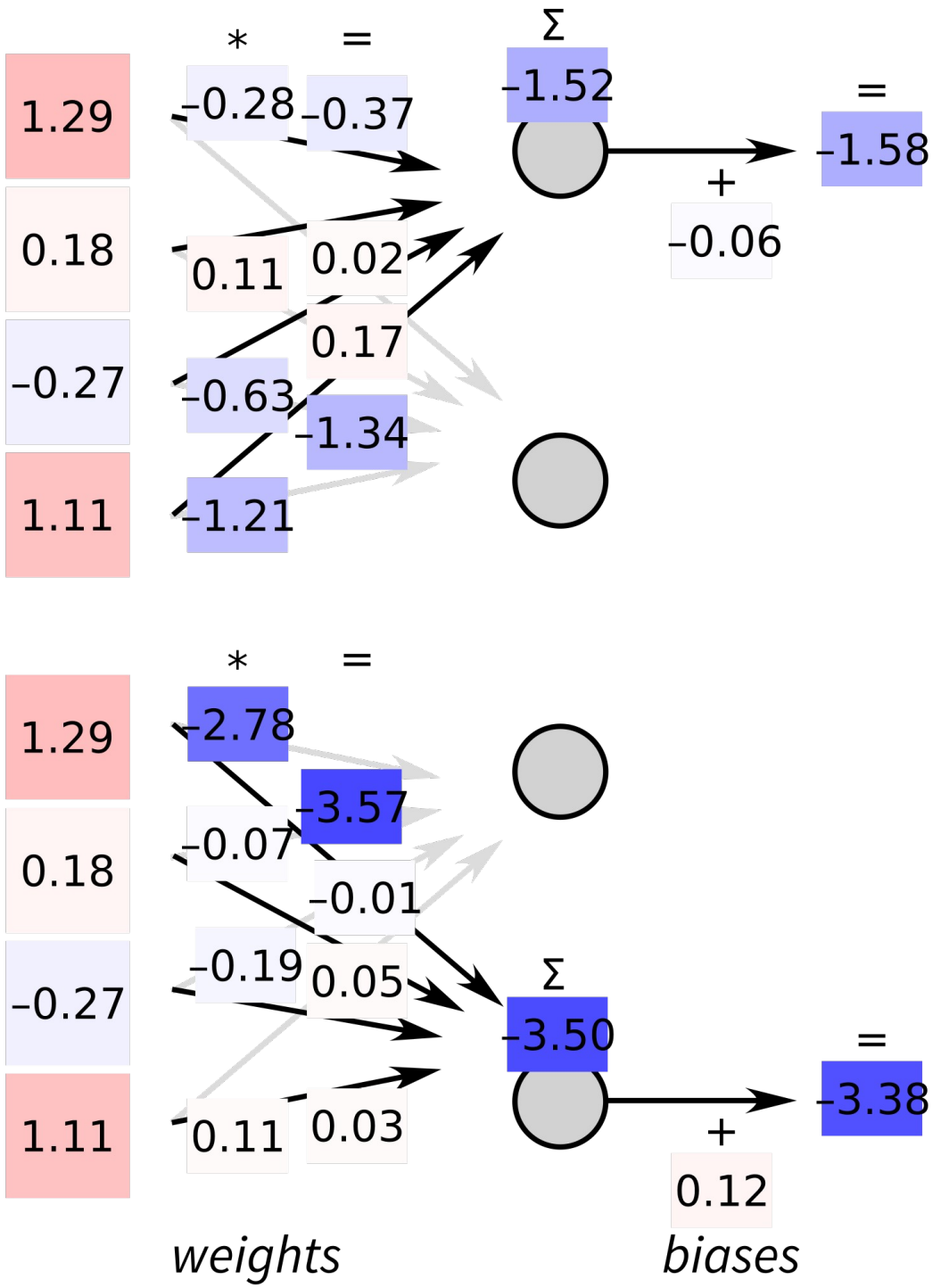
Boletus edulis 0%

Agaricus campestris 0%

https://play.google.com/store/apps/details?id=bazinac.aplikacenahouby

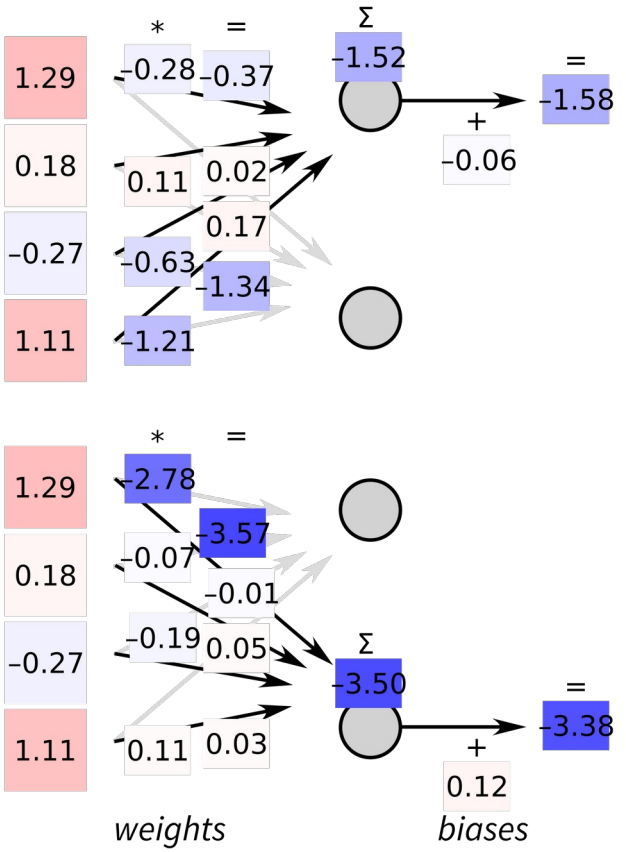
Alphafold:

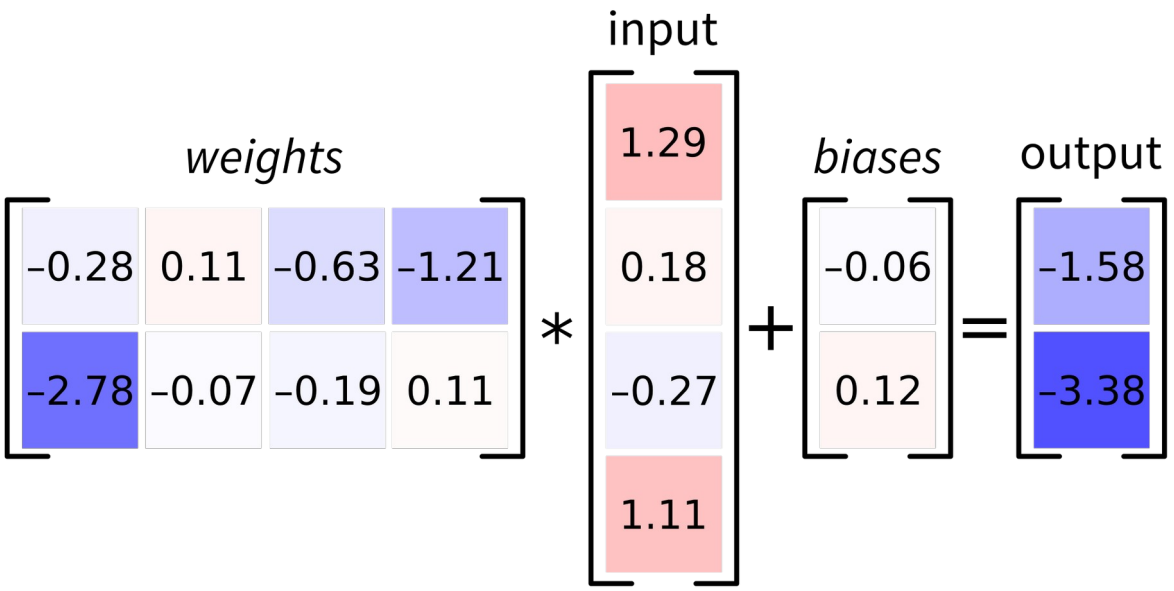
- neuronové sítě



Alphafold:

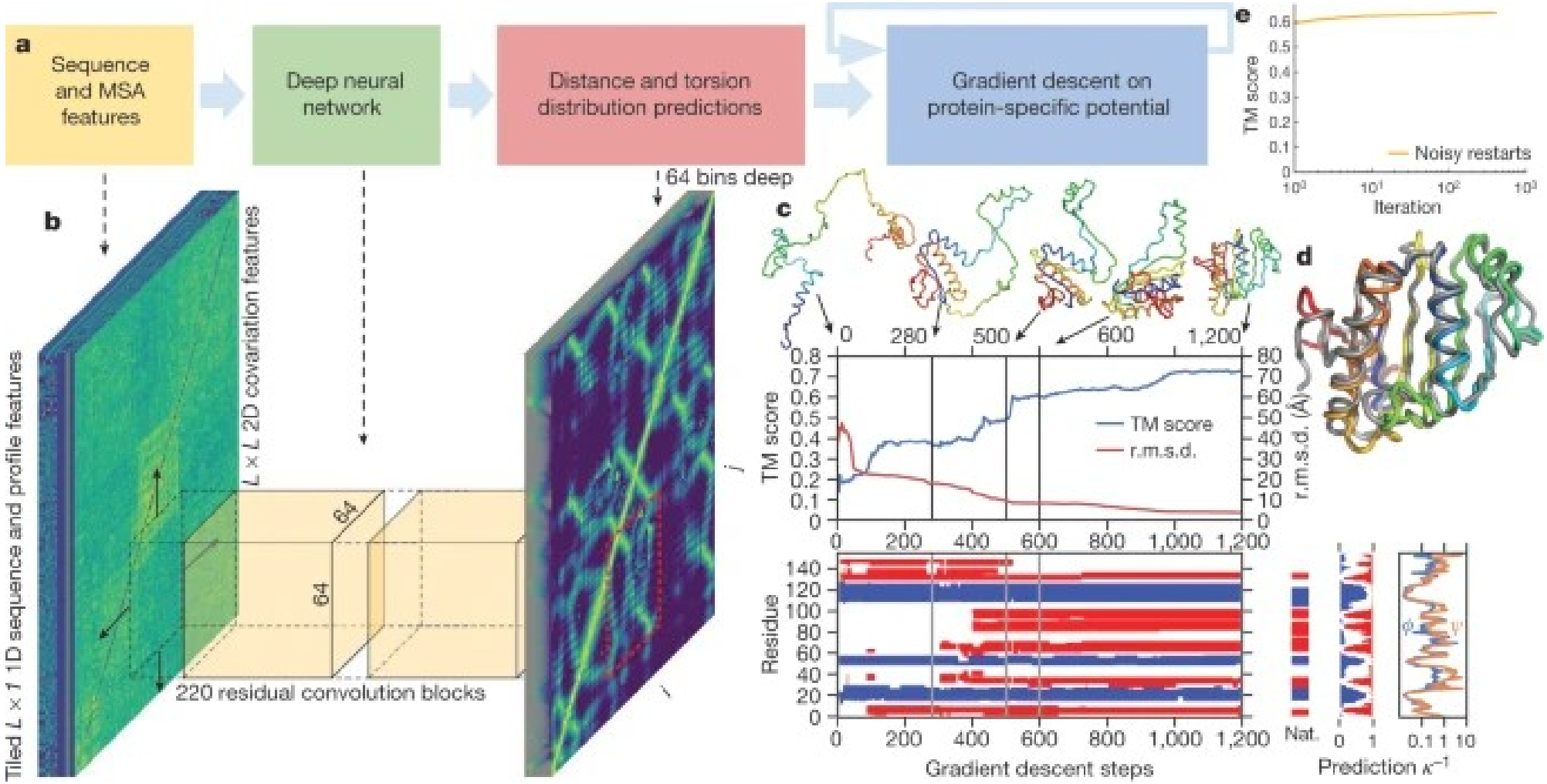
- neuronové sítě





Alphafold:

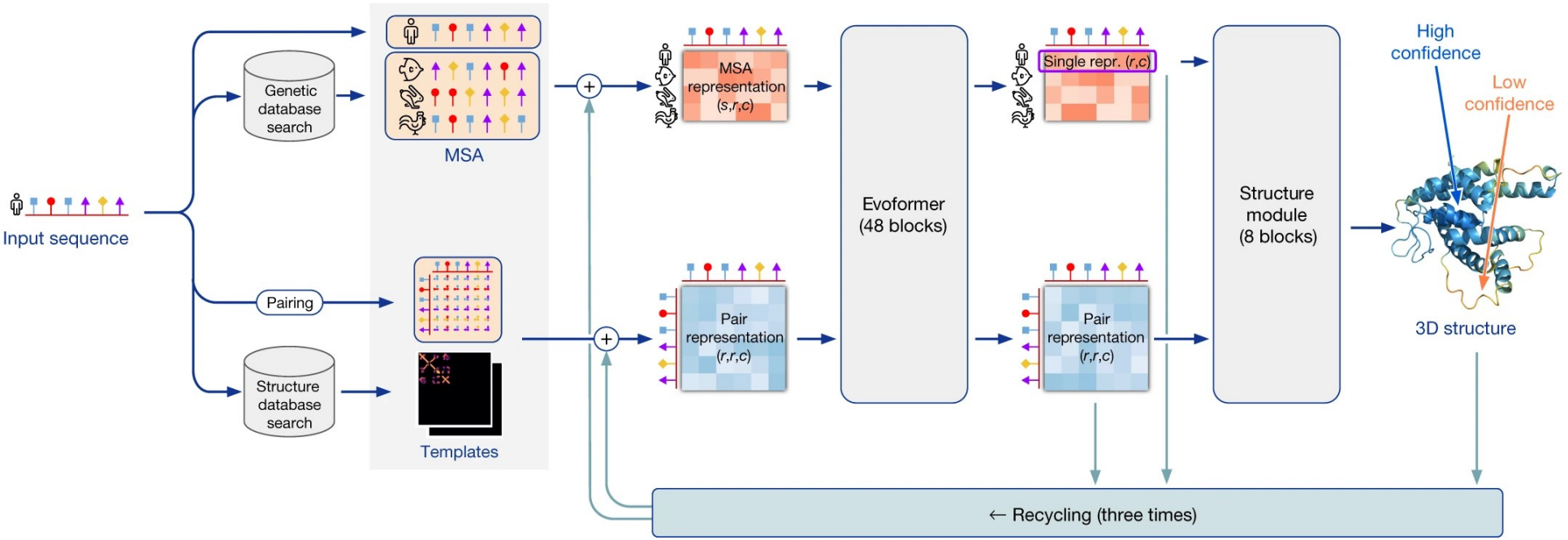
- AF1



Senior, A.W., Evans, R., Jumper, J. et al. Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020). https://doi.org/10.1038/s41586-019-1923-7

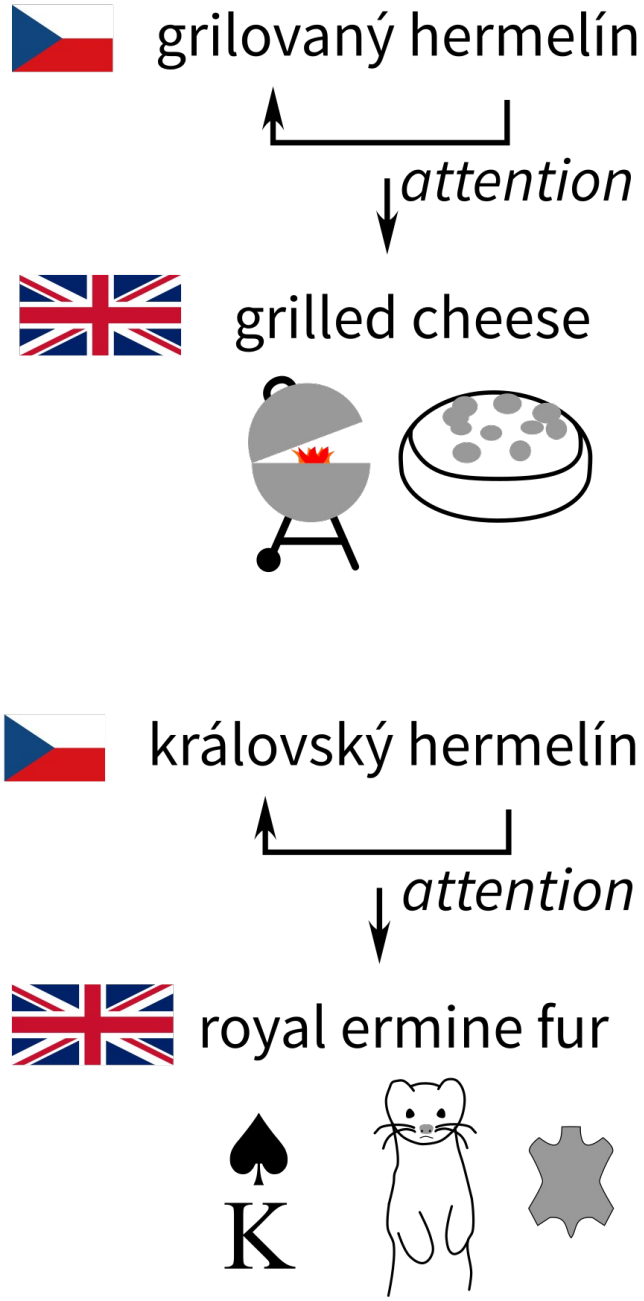
Alphafold:

- AF2

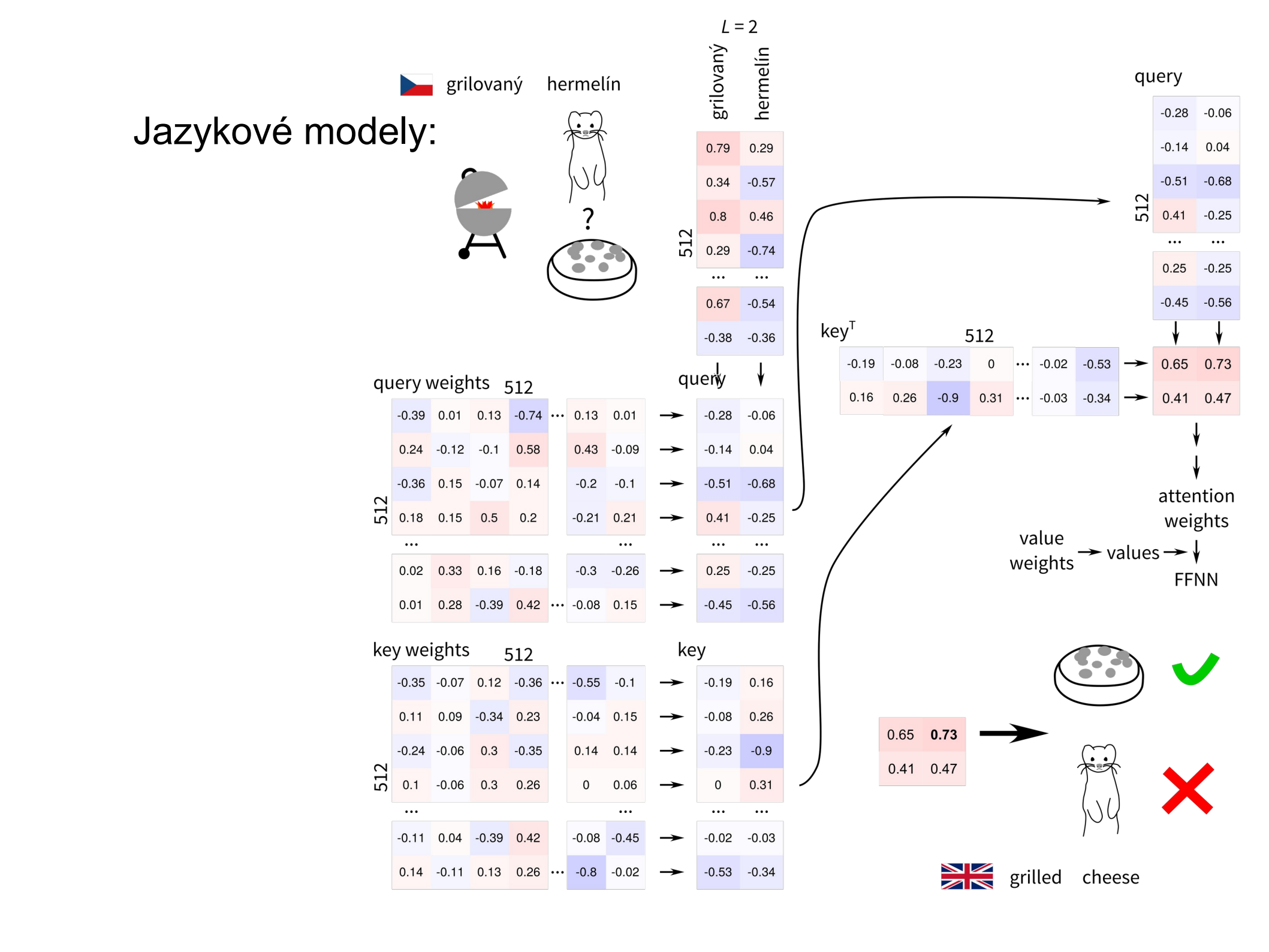


Jumper, J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2

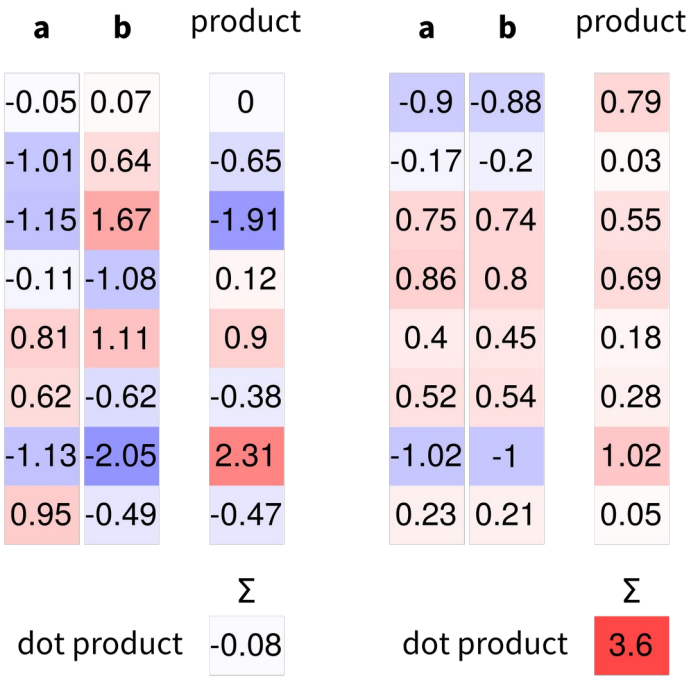
Jazykové modely:



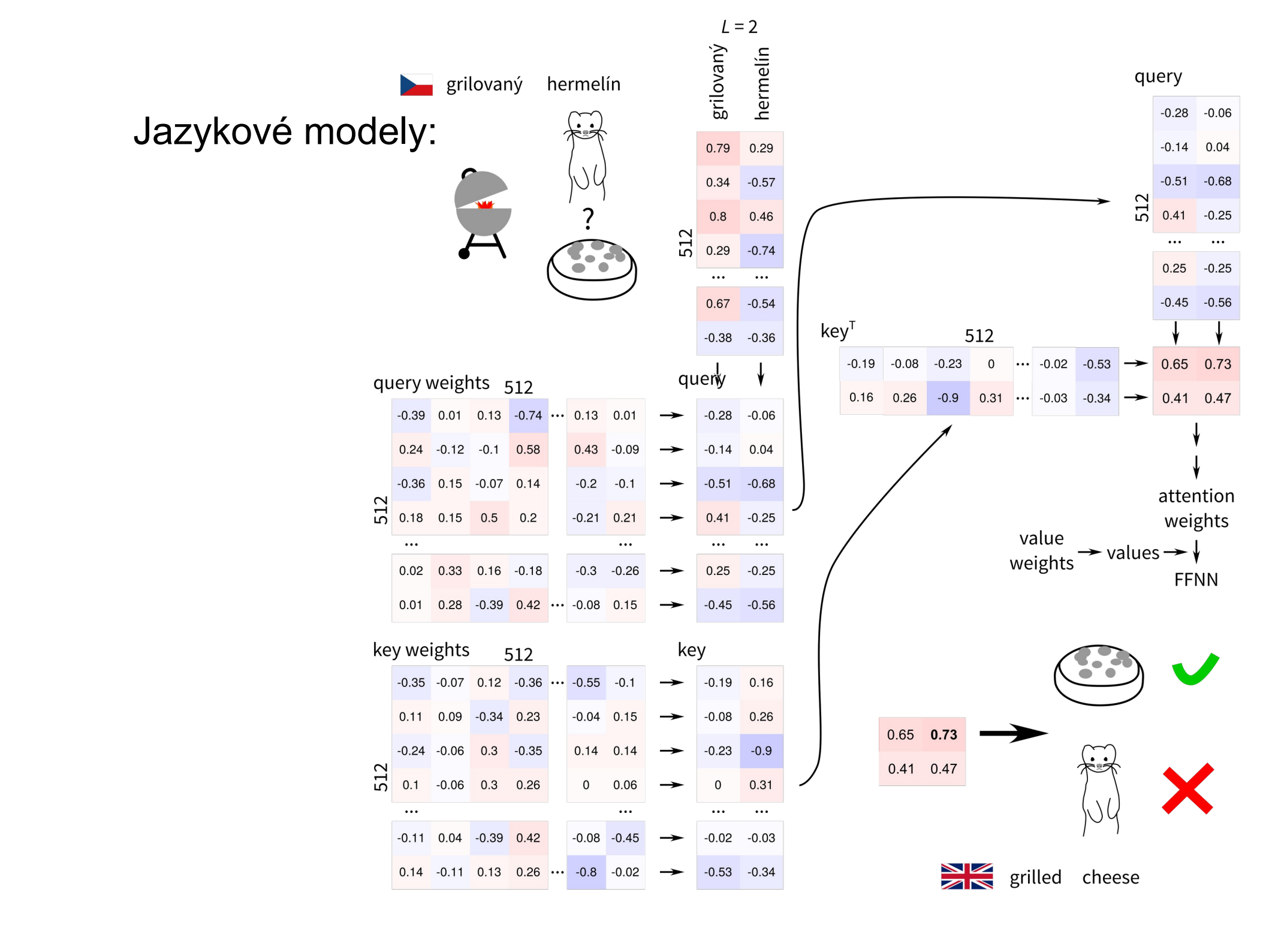
Jazykové modely:



Jazykové modely:

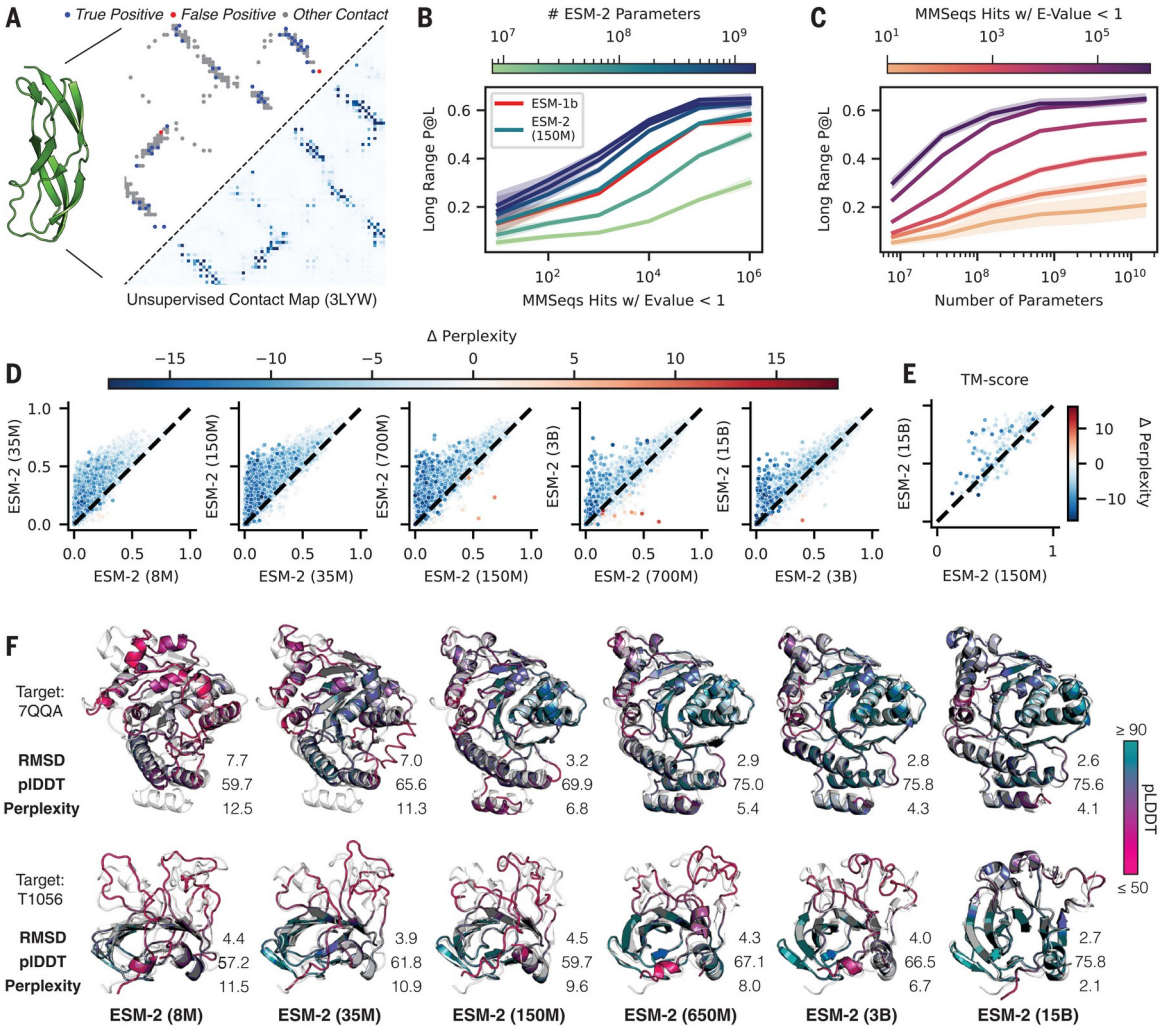


Jazykové modely:



Jazykové modely:

- ESMfold



Zeming Lin et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123-1130 (2023). https://doi.org/10.1126/science.ade2574

Předpověď struktur proteinů:
Critical Assessment of protein Structure Prediction (CASP)

Protein Structure Prediction Center

Menu

- Home
- FORCASP Forum
- PC Login
- PC Registration
- CASP Experiments
  - CASP ROLL
  - CASP11 (2014)
    - Home
    - Registration
    - My CASP11 profile
    - Targets
      - Target List
      - Target Submission
  - CASP10 (2012)
  - CASP9 (2010)
  - CASP8 (2008)
  - CASP7 (2006)
  - CASP6 (2004)
  - CASP5 (2002)
  - CASP4 (2000)
  - CASP3 (1998)
  - CASP2 (1996)
  - CASP1 (1994)
- Initiatives
- Data Archive
- Local Services
- Proceedings
- Feedback
- Assessors

Welcome to the Protein Structure Prediction Center!

Our goal is to help advance the methods of identifying protein structure from sequence. The Center has been organized to provide the means of objective testing of these methods via the process of blind prediction. The Critical Assessment of protein Structure Prediction (CASP) experiments aim at establishing the current state of the art in protein structure prediction, identifying what progress has been made, and highlighting where future effort may be most productively focused.

There have been ten previous CASP experiments. The eleventh experiment will start in May 2014. Description of these experiments and the full data (targets, predictions, interactive tables with numerical evaluation results, dynamic graphs and prediction visualization tools) can be accessed following the links:

CASP1 (1994) | **CASP2 (1996)** | **CASP3 (1998)** | **CASP4 (2000)** | **CASP5 (2002)** | **CASP6 (2004)** | **CASP7 (2006)** | **CASP8 (2008)** | **CASP9 (2010)** | **CASP10 (2012)** | **CASP11 (2014)**

Raw data for the experiments held so far are archived and stored in our data archive.

In November 2011 we have opened a new rolling CASP experiment for all-year-round testing of ab initio modeling methods:

CASP ROLL

Details of the experiments have been published in a scientific journal *Proteins: Structure, Function and Bioinformatics*. **CASP proceedings** include papers describing the structure and conduct of the experiments, the numerical evaluation measures, reports from the assessment teams highlighting state of the art in different prediction categories, methods from some of the most successful prediction teams, and progress in various aspects of the modeling.

Prediction methods are assessed on the basis of the analysis of a large number of blind predictions of protein structure. Summary of numerical evaluation of the methods tested in the latest CASP experiment can be found on this web page. The main numerical measures used in evaluations are described in the papers [1] , [2] . The latter paper also contains explanations of data handling procedures and guidelines for navigating the data presented on this website.

Some of the best performing methods are implemented as fully automated servers and therefore can be used by public for protein structure modeling.

To proceed to the pages related to the latest CASP experiments click on the logo below:

Message Board

CASP11 dry run for servers

Dear CASP participants, Registration for CASP11 is currently under way and more than 60 research groups have already registered for the experiment. We are starting checking connectivity and correct ...

CASP11 registration opens March 31

Dear CASP Participants, Exiting news: new CASP experiment is just around the corner! We hope that you are full of enthusiasm and anxiety (as we are) and have your computers greased and warmed up. ...

Resuming CASP ROLL

Dear CASPers, Best regards for all of you in the New Year! Hoping that you had good rest after the CASP10 experiment and meeting, we are resuming CASP ROLL with two new targets later this week. ...

http://predictioncenter.org/