

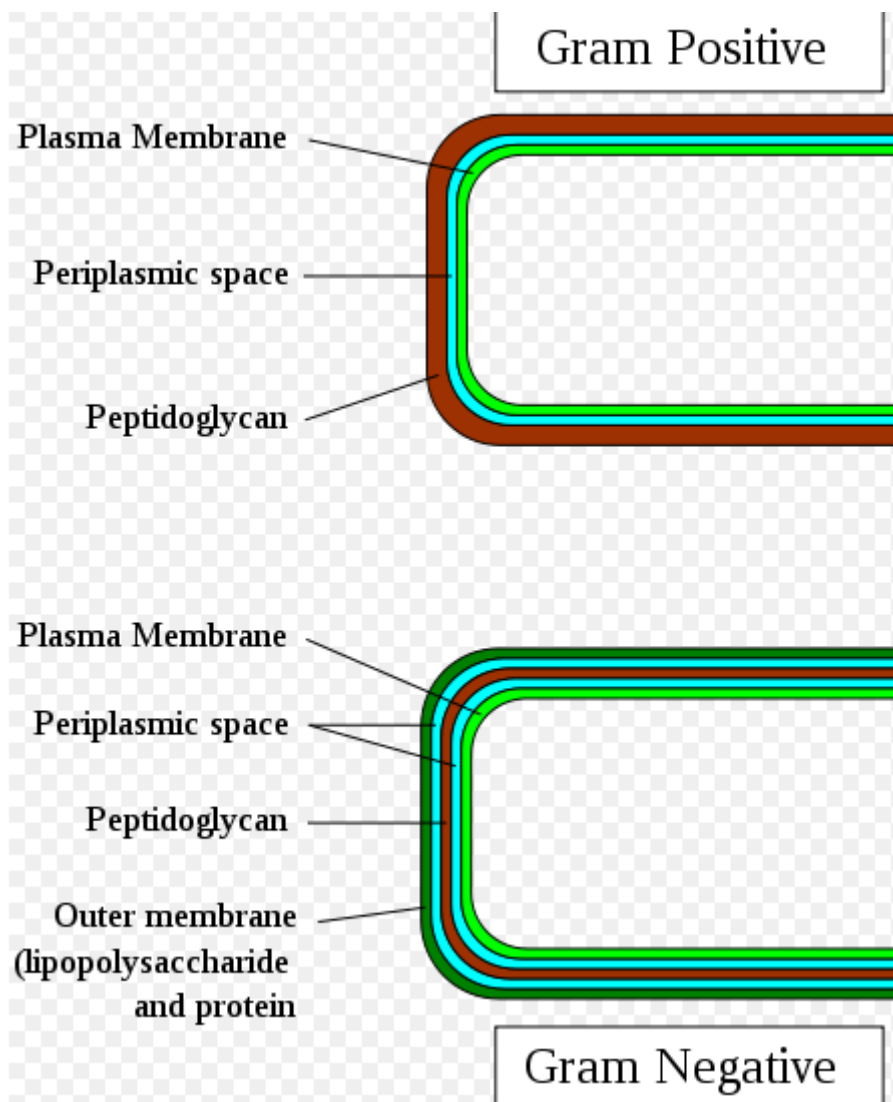
BIOINFORMATIKA V PRAXI – CVIČENÍ 8

PREDIKCE VLASTNOSTÍ PROTEINŮ

Další vlastnosti proteinů

PREDIKCE LOKALIZACE PROTEINU V BUŇCE

Predikce lokalizace proteinu v buňce může být cenným nástrojem pro určení FUNKCE proteinu. Z výskytu proteinů v buňce (cytoplasmatické, membránové, extracelulární, v buněčné stěně...) můžeme například u patogenních mikroorganismů odhadnout, zda se jedná o protein důležitý pro virulenci.



ÚKOL 1

Pomocí programu **PSORT** (<http://www.psорт.org/>) určete lokalizaci následujícího proteinu v buňce:

Protein z bakterie *Micrococcus luteus*

```
MEINGLTWGLTIALIVGLLAFDYFAHVRKAHTPTIKEAAVWSGVYVGLALVFGLVFFAFGDTQHAVEYYTGYLE
KALSVDNLFVFLVIMASFRVPREYQQKVLVLFGITFALISRTLFI LLGAAVIAAWSDFYLFGLFLLI IAGGQLKG
EMSGDAEGAQDEADNVMVRLVKRFLPASDQFDGERLFTTVDGKRLMTPMLLVMI AIGATDILFAFDSIPAI FGVT
QEAYIVFTATAFSLMGLRQLYFLIDGLLDRLVYLAYGLSALLAFIGVKLILHALHENNL PFI NGGENVPVAE IPT
NLSLVVVVVILAITVLVSLYSPKQALRALQNAEKYSYRYSKLAEDADPAERERAAALMDRWTARAEALDQRWRD
QLLEHKDAWSAIIRTAHETRFADPRDDDARGVSEQIVRQDGP TV
```

ÚKOL 2

Pomocí programu **PSORT** určete lokalizaci následujícího proteinu v buňce:

Protein z *Aspergillus fumigatus*

```
MKPQTAAFLLSLLGSTLAAPIQHADKGSVDKPTPTGRGAPGGFFTGFPSPGVPSGLPSGFPGGPVPGGFGGDGPNG
PIPSGPVPTGAAPSGFSPFGTGPAPSGAPQGEESSSFGGQGVQARSPQDFEDSGAAPSGAIPSGAIP TGAVPSG
APNGFGGFGQGGHGGPGGPGEEGSGPSPTGAVPSGAI PSGAAPSGAVGGFDGFGQASENSGNAQFQSSGTS PFGA
SHSGSASGHQGGRRHGGDHRGQHNGNSGAI PSGAAPSGAAPSGAAGGFPGFGQGGEGSGPSPTGAVPSGAAGFGGQ
GHGQQGGSFPTGVAPSDVPSAQPTA
```

ÚKOL 3

Určete lokalizaci proteinu z *Aspergillus fumigatus* v buňce rovněž pomocí **TargetP** (<http://www.cbs.dtu.dk/services/TargetP/>). Výsledky porovnejte.

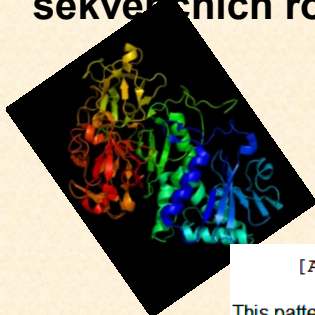
Sekundární databáze proteinů

- **Databáze mohou obsahovat:**

Proteinové **DOMÉNY** odvozené ze známých struktur

Proteinové sekvence seřazené do **SEKVENČNÍCH RODIN**

CHARAKTERISTICKÉ MOTIVY odvozené z těchto sekvencí rodin.



```
LPPNTAFKAIIFYANAADRQDLKLFIDD
IPPNTDFRAIFFANAAEQOHIKLFIGD
LPPHIKFGVTALTHAANDQTIDIYIDD
LPPNIAFGVTALVNSSAPOTIEVFVDD
```

[AC]-x-v-x(4)-{ED}.

This pattern is translated as: [Ala or Cys]-any-Val-any-any-any-any-{any but Glu or Asp}

SEKUNDÁRNÍ DATABÁZE PROTEINŮ

Sekundární databáze obsahují informace odvozené z primárních databází ve formě charakteristických vzorů sekvencí, tj. funkčních nebo strukturních motivů získaných srovnáním primárních dat (sekvencí). Sekundární databáze jsou vhodné pro vyhledávání „vzoru“ charakteristického pro určitou skupinu proteinů, tj. pro určení funkce proteinů.

Sekundární databáze proteinů (**PROSITE**, **Pfam**, **PRINTS**, **ProDom**, **SMART**, **TIGRFAMS**) jsou v současné době sdruženy do integrované klasifikační databáze proteinů **InterPro**.

ÚKOL 4

Jeden z předcházejících proteinů analyzujte pomocí databáze **InterPro** (<http://www.ebi.ac.uk/interpro/>). Jaké domény/motivy protein obsahuje a jaká je předpokládaná funkce proteinu? Popište.

PREDIKCE SEKUNDÁRNÍ STRUKTURY PROTEINU

ÚKOL 5

Určete, zda následující protein obsahuje transmembránové helixy. Využijte nástroj **TMHMM** (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>).

```
MEINGLTWGLTIALIVGLLAFDYFAHVRKAHTPTIKEAAVWSGVYVGLALVFGLVFFAFGDTQHAVEYYTGYLE
KALSVDNLFVFLVIMASFRVPREYQQKVLVLFGITFALISRTLFI LLGAAVIAAWSDFYLFGLFLLI IAGGQLKG
EMSGDAEGAQDEADNVMVRLVKRFLPASDQFDGERLFTTVDGKRLMTPMLLVMI AIGATDILFAFDSIPAI FGVT
QEAYIVFTATAFSLMGLRQLYFLIDGLLDRLVYLAYGLSALLAFIGVKLILHALHENNLPFINGGENVPVAE IPT
NLSLVVVVILAITVLVSLYSPKGQALRALQNAEKYSYRYSKLAEDADPAERERAAALMDRWTARAEALDQRWRD
QLLEHKDAWSAII RTAHE TRFADPRDDDARGVSEQIVRQDGPTV
```

ÚKOL 6

Predikujte sekundární strukturu následujícího proteinu. Využijte libovolné dva programy, které najdete na <http://www.expasy.org/>. Výsledky porovnejte a komentujte.

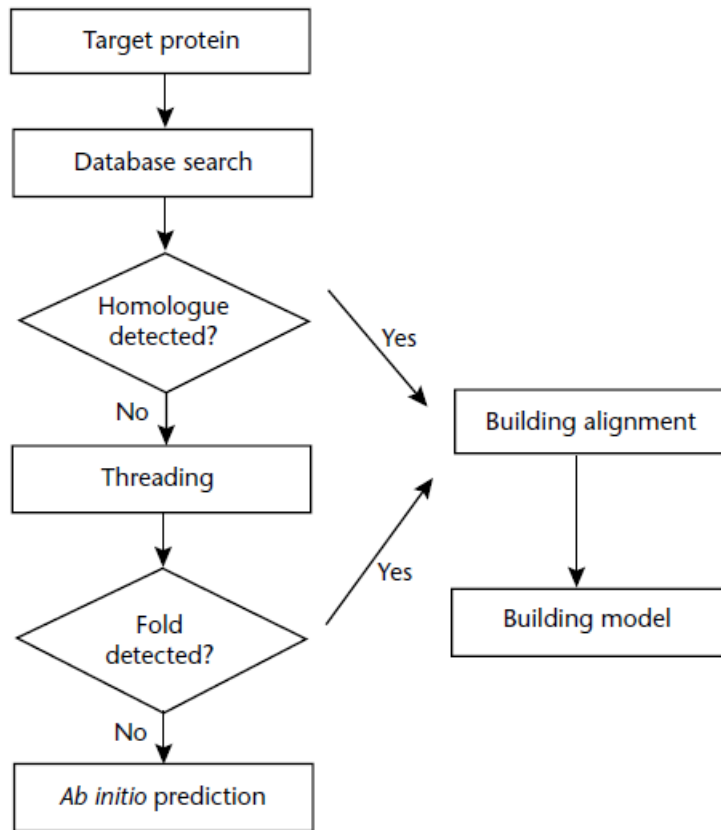
```
ADSQTSSNRAGEFSIPPNTDFRAIFFANAAEQQHKLFIGDSQEPAAYHKL TTRDGPREATLNSGNGKIRFEVSV
NGKPSATDARLAPINGKKS DGSFPFTVNF GIVVSE DGHSDYNDGIVVLQWPIG
```

PREDIKCE TERCIÁRNÍ STRUKTURY PROTEINU

The result is only a model and must be considered carefully, it isn't an experimental 3D structure! (citace z dokumentace k programu Geno3d)

This server is experimental. Some of the methods used are untested and/or unpublished. Use the server and its results at your own risk. For more information, contact the authors. (citace z dokumentace k programu HMMSTR/Rosetta)

Evaluation of template structure and model quality is a crucial step in homology modelling. (citace z dokumentace k programu SWISS-MODEL).



SAMOSTATNÝ PROJEKT

Predikujte lokalizaci svého proteinu v buňce, dále analyzujte svůj protein pomocí databáze InterPro, určete jeho možnou funkci a predikujte jeho sekundární strukturu.