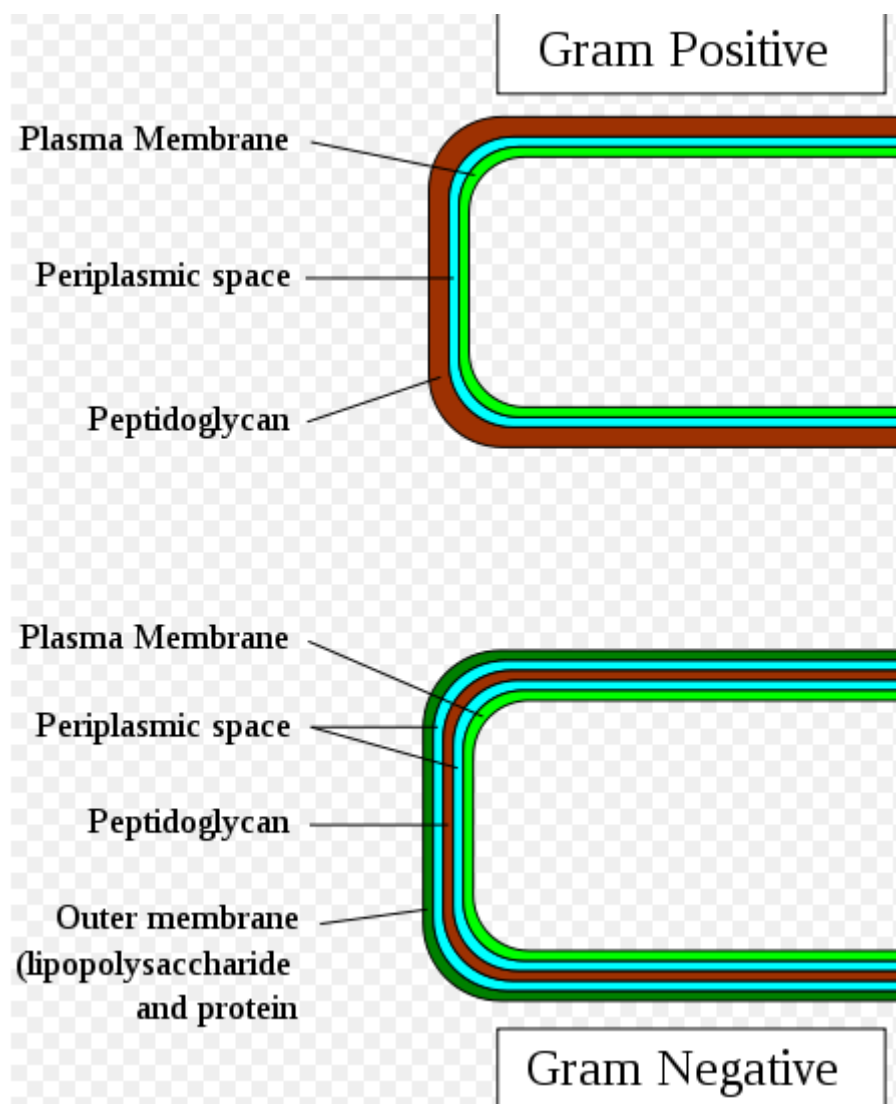


BIOINFORMATIKA V PRAXI – CVIČENÍ 7

PREDIKCE VLASTNOSTÍ PROTEINŮ DOKONČENÍ

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 grampozitivní bakterie *Micrococcus luteus*

MEINGLTWGLTIALIVGLLAFDYFAHVRKAHTPTIKEAAVWSGVYVGLALVFGLVFFAFGDTQHAVEYYTGYLE
KALSVDNLFVFLVIMASFRVPREYQKVLVLFGITFALISRTLIFILLGAAVIAAWSDFYLFGLFLLI IAGGQLKG
EMSGDAEGAQDEADNVMVRLVKRFLPASDQFDGERLFTTVDGKRLMTPMLLVMIAGATDILFAFDSIPAI FGVT
QEAYIVFTATAFSLMGLRQLYFLIDGLLDRLVYLAYGLSALLAFIGVKLILHALHENNLFPINGGENVPAE IPT
NLSLVVVVVILAITVLVSLYSPKGQALRALQNAEKYSYRYSKLAEDADPAERERAAAALMDRWTARAEALDQRWRD
QLLEHKDAWSAII RTAHETRFADPRDDDARGVSEQIVRQDGP TV

ÚKOL 2

Pomocí programu **PSORT** určete lokalizaci následujících proteinů v buňce:

Protein 1 z neznámé gramnegativní bakterie

MTKYALVGDVGGTNARLALCDIASGEISQAKTYSGLDYP SLEAVVRVYLDEHGVSVEDGCIAIACPITGDWVAMT
NHTWAFSIAEMKKNLGFSHLEIINDFTAVSMAIPMLKKEHLIQFGGGE PVDGKPIAVYGAGTGLGVAHLVHVKR
WISLPGEGGHVDFAPNSEEEAMILEILRAEIGHVSAERVL SGPGLVNLRYRAIVKSDNRLPENLRPKDITERALAD
SCIDCRRALS LFCVIMGRFGGDLALTMGTFGGVYIAGGIVPRFLEFFKASGFRGGFEDKGRFKDYVHGIPVYLIV
HDNPGLLGSGAHLRQTLGHIL

Protein 2 z neznámé bakterie

LVIVDAVTLLSAYPEASRDPAAPTVIDGRHLYVVSPGDAAQLGHND SRLFTGLSPGDQLHLRETALALRAEVSVL
FIRFALKDAGIVAPIELEVRDAATAVPDADDLLHPS CRPLKDHYWRSVDVLAAGATTCTADFAVCDRDGTVSGYFR
WETSIEIAGSQPDTKQPGFKPSS

Protein 3 z *Aspergillus fumigatus*

MKPQTAAFLLSLLGSTLAAPIQHADKGS DVKPTPTGRGAPGGFFTFGFP SGVPSGLPSGFPGGPVPGGFGGDPNG
PIPSGPVPTGAAPSGFPSFGTGPAPSGAPQGEES SFFGGQVQARSPQDFEDSGAAPSGAIPSGAIP TGAVPSG
APNGFGGFGQGGHGGPGGPEEGSGPSPTGAVPSGAI PSGAAPSGAVGGFDGFGQASENSGNAQFQSSGTS PFGA
SHSGSASGHQGGRRHGGDHRGQHNGSGAIPSGAAPS GAAPSGAAGGFPGFGQGGEGSGPSPTGAVPSGAAGFGGQ
GHGQGGQSFPTGVAPSDVPSAQPTA

SAMOSTATNÝ PROJEKT

Predikujte lokalizaci Vašeho proteinu v buňce.

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

- **Databáze mohou obsahovat:**

Proteinové **DOMÉNY** odvozené ze známých struktur

Proteinové sekvence seřazené do **SEKVENČNÍCH RODIN**

CHARAKTERISTICKÉ MOTIVY odvozené z těchto sekvencí rodin.



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

ÚKOL 3

Vyberte si tři z předcházejících proteinů a s využitím nástroje InterProScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>) je analyzujte pomocí databáze InterPro. Jaké domény/motivy proteiny obsahují? Popište.

PREDIKCE SEKUNDÁRNÍ STRUKTURY PROTEINU

ÚKOL 4

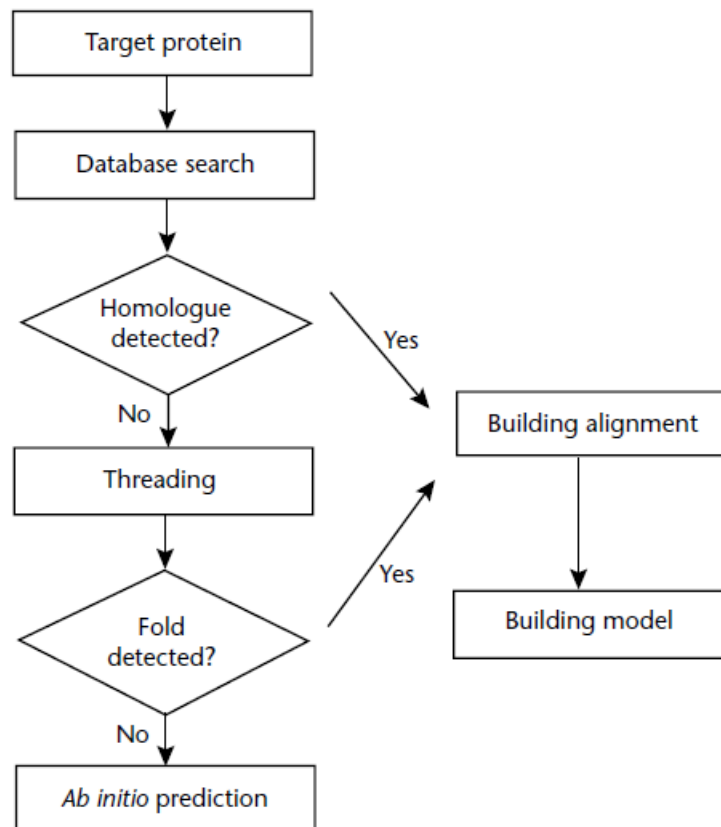
Vyberte si jeden z předcházejících proteinů a predikujte sekundární strukturu tohoto proteinu. Využijte libovolné dva programy, které najdete na <http://www.expasy.org/>. Výsledky porovnejte a komentujte.

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

Analyzujte svůj protein pomocí databáze InterPro, určete jeho možnou funkci a predikujte jeho sekundární strukturu.