

## Blok 2: Sekundární struktura proteinů

### Úloha 1:

Vyhodnoťte data získaná měřením cirkulárního dichroismu roztoku neznámého proteinu (*Input\_CD\_data.txt*). Určete, zda se ve vzorku nachází protein převážně ve stavu  $\alpha$ -helixu,  $\beta$ -skládaného listu nebo náhodného klubka (coil). K vyhodnocení využijte program **k2d3** (<http://www.ogic.ca/projects/k2d3/>).

### Úloha 2:

Na základě základních znalostí o uspořádání sekundárních struktur odhadněte **bez použití predikčních programů**, zda následující krátké peptidy jsou součástí  $\alpha$ -helixu,  $\beta$ -skládaného listu nebo náhodného klubka. Můžete využít přehlednou tabulku aminokyselin.

	NONPOLAR, HYDROPHOBIC	R GROUPS	POLAR, UNCHARGED	
Alanine Ala A MW = 89	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_3 \end{array}$		$\begin{array}{c} \text{H} - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Glycine Gly G MW = 75
Valine Val V MW = 117	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH} \begin{array}{l} \diagup \text{CH}_3 \\ \diagdown \text{CH}_3 \end{array} \end{array}$		$\begin{array}{c} \text{HO} - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Serine Ser S MW = 105
Leucine Leu L MW = 131	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_2 - \text{CH} \begin{array}{l} \diagup \text{CH}_3 \\ \diagdown \text{CH}_3 \end{array} \end{array}$		$\begin{array}{c} \text{OH} \\   \\ \text{CH}_3 - \text{CH} - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Threonine Thr T MW = 119
Isoleucine Ile I MW = 131	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH} \begin{array}{l} \diagup \text{CH}_3 \\ \diagdown \text{CH}_2 - \text{CH}_3 \end{array} \end{array}$		$\begin{array}{c} \text{HS} - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Cysteine Cys C MW = 121
Phenylalanine Phe F MW = 131	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_2 - \text{C}_6\text{H}_5 \end{array}$		$\begin{array}{c} \text{HO} - \text{C}_6\text{H}_4 - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Tyrosine Tyr Y MW = 181
Tryptophan Trp W MW = 204	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_2 - \text{C}_8\text{H}_6\text{N}_2 \end{array}$		$\begin{array}{c} \text{NH}_2 \\   \\ \text{O} = \text{C} - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Asparagine Asn N MW = 132
Methionine Met M MW = 149	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_2 - \text{CH}_2 - \text{S} - \text{CH}_3 \end{array}$		$\begin{array}{c} \text{NH}_2 \\   \\ \text{O} = \text{C} - \text{CH}_2 - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Glutamine Gln Q MW = 146
Proline Pro P MW = 115	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_2 - \text{CH}_2 \\   \quad \quad \quad   \\ \text{HN} - \text{CH}_2 \quad \quad \quad \text{CH}_2 \end{array}$		<b>POLAR BASIC</b> $\begin{array}{c} \text{NH}_3^+ - \text{CH}_2 - (\text{CH}_2)_3 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Lysine Lys K MW = 146
Aspartic acid Asp D MW = 133	<b>POLAR ACIDIC</b> $\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_2 - \text{C}(=\text{O})\text{O}^- \end{array}$		$\begin{array}{c} \text{NH}_2 \\   \\ \text{N H}_2^+ = \text{C} - \text{NH} - (\text{CH}_2)_3 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+ \end{array}$	Arginine Arg R MW = 174
Glutamine acid Glu E MW = 147	$\begin{array}{c} \text{OOC} \\   \\ \text{H}_3\text{N}^+ - \text{CH} - \text{CH}_2 - \text{CH}_2 - \text{C}(=\text{O})\text{O}^- \end{array}$		$\begin{array}{c} \text{C} = \text{CH}_2 \\   \\ \text{HN} \quad \quad \quad \text{NH} \\   \\ \text{H} \end{array} - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{N H}_3^+$	Histidine His H MW = 155

<i>Peptid</i>	<i><math>\alpha</math>-helix</i>	<i><math>\beta</math>-skládaný list</i>	<i>náhodné klubko</i>
KALSQVIRL	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
GGASPAPPK	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
LELQIRVYA	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### Úloha 3:

Seznamte se s programem **Jpred 3** (<http://www.compbio.dundee.ac.uk/www-jpred/>). Pro následující proteinové sekvence predikujte zastoupení sekundárních struktur. Patří tyto proteiny do skupiny tzv.  $\alpha/\beta$  proteinů, tj. proteinů, ve kterých jsou zastoupeny obě typické sekundární struktury ( $\alpha$ -helix,  $\beta$ -skládaný list)?

#### Protein A:

GTYRQLFHPEQLISGKEDAANNYARGHYTIGKEIIDLVLDVRVKLADQCTGLQGFLVFHSFGGGTGSGFTSLLME  
 RLSVDYGKKSLEFSIYPAPQVSTAVVEPYNSILTHTTLEHSDCAFMDNEAIYDICRRNLDIERPTYTNLNLRL  
 MSQIVSSITASLRFDGALNVDLTFQTNLVPYPRIHFPLATYAPVISA EKAYHEQLSVAEITNACFEPANQMVKC  
 DPRHGKYMACECLLYRGDVPKDVNAAIATIKTKRTIQFVDWCPTGFKVGINYQPPTVVPGGDLAKVQRAVCMLSN  
 TTAVAEAWARLDHKFDLMYAKRAFVHWYVGEEMEEGEFSEAREDMAALEKDYEEVGDADSYEDEDEGE EY

#### Protein B:

SIRLPAHLRLQPIYWSRDDVAQWLKWAENEFSLRPIDSNTEFEMNGKALLLLTKEDFRYRSPHSGDVLVYELLQHIL  
 KQAGPNIFEMLRIDEGLRLKIYKDTEGYTTIGIGHLLTKSPSLNAAKSELDKAIGRNTNGVITKDEAEKLFQDQV  
 DAAVRGILRNAKLPVYDSLDCVRRRAALINMVFMQGETGVAGFTNSLRMLQQKRWDEAAVNLA KSRWYNQTPNRA  
 KRVITTFRTGTW DAYK

#### Protein C:

ADVCMDEPEPIVRIVGRNGLCVDVDRDGRFHNGNAIQLWPCKSNTDANQLWTLKRDNTIRSNGKCLTTYGYSPGVYV  
 MIYDCNTAATDATRWQIWDNGTIINPRSSLVLAATSGNSGTTTLTVQTNIYAVSQWLPTNNTQPFVTTIVGLYGL  
 CLQANSQVWIEDCSSEKAEQQWALYADGSIRPQQNRDNCLESDSNIRETVVKILSCGPASSGQRWMPKNDGTIL  
 NLYSGLVLDVVRASDPSLKQIILYPLHGDPNQIWLPLF

### Úloha 4:

Pokuste se pomocí programu **TMpred** předpovědět, kolik transmembránových helixů obsahuje protein NMU-R1 z myši (UniProtKB/Swiss-Prot: O55040.1). Bude podle predikce jeho C-terminální doména orientována spíše dovnitř buňky nebo ven?

### Úloha 5:

Následující sekvenci analyzujte pomocí programu **CFSSP** (<http://www.biogem.org/cgi-bin/cho-fas.pl>).  
Tutéž sekvenci analyzujte pomocí programu **GOR4** ([http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_gor4.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html)) a výsledky porovnejte.

Sekvence:

MSTDKPLVIQSDKTLLLDVHSPFAQECRDSITAFSELVKSPEHVHTFLLTPLSLWNANAAGMTTEDIMGRLRTWS  
RYDIPEPVSYFITDISARFGSFVMTDIPDDADHYLLTVTI PRYAKEISSHKTVSSLLFPRGNDTFLNRYARGEV  
KLKLIKLGFPVDDRIPLKKGFPVPMNLRQQTLSGKDFSIRDYQEAARSLLGDRGPGTGY

Procentuální zastoupení	CFSSP	GOR4
$\alpha$ -helix		
$\beta$ -skládaný list		
ohyb + náhodné klubko		

### Úloha 6:

Nejspolehlivější informace o sekundární struktuře můžeme získat analýzou 3D struktury, je-li tato známa. 3D struktury jsou ukládány v tzv. **PDB databázi**. (např.). Vyhledejte na webových stránkách <http://pdb.rcsb.org> libovolný protein a podívejte se na informace o jeho sekundární struktuře.

### Úloha 7:

Stejný protein, který jste již analyzovali v PDB databázi analyzujte rovněž pomocí nástrojů na serveru **2struct** (<http://2struct.cryst.bbk.ac.uk/twostruc>). Porovnejte přiřazení sekundárních struktur v jednotlivých programech.