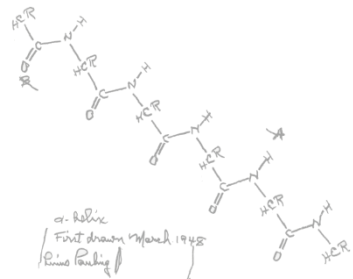


Výskyt sekundárních struktur v sekvenčním prostoru proteinů

Vjačeslav Tretjačenko



ENBIK
Národní bioinformatická konference

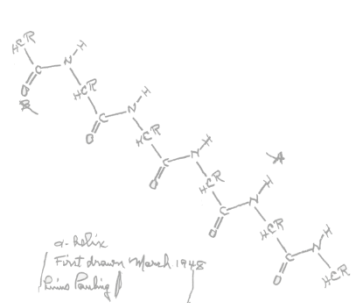


ÚOCHB AV
CR
IOCB PRAGUE

Co nevíme?

Je struktura proteinů unikátním přírodním vynálezem nebo i náhodna proteinová sekvence schovává strukturní potenciál?

Celkový odhad počtu různých proteinových sekvencí na Zemi	Počet možných kombinací (délka 100 AA, abeceda 20 AA)
$\sim 10^{13}$	$\sim 10^{130}$



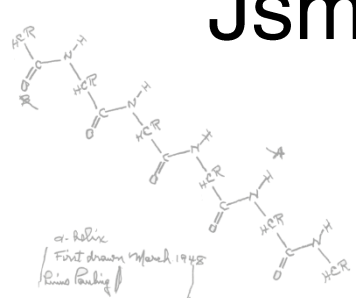
Co nevíme?

Jak častý je výskyt struktur v přírodou netknutých sekvencích?

Chiarabelli, Cristiano, et al. "Investigation of de novo totally random biosequences, Part II." *Chemistry & biodiversity* 3.8 (2006): 840-859.

- „Never Born Proteins“ – v přírodě se nevyskytující proteiny
- Phage display knihovny náhodných sekvencí - 20 % odolné k proteolýze (folding?)

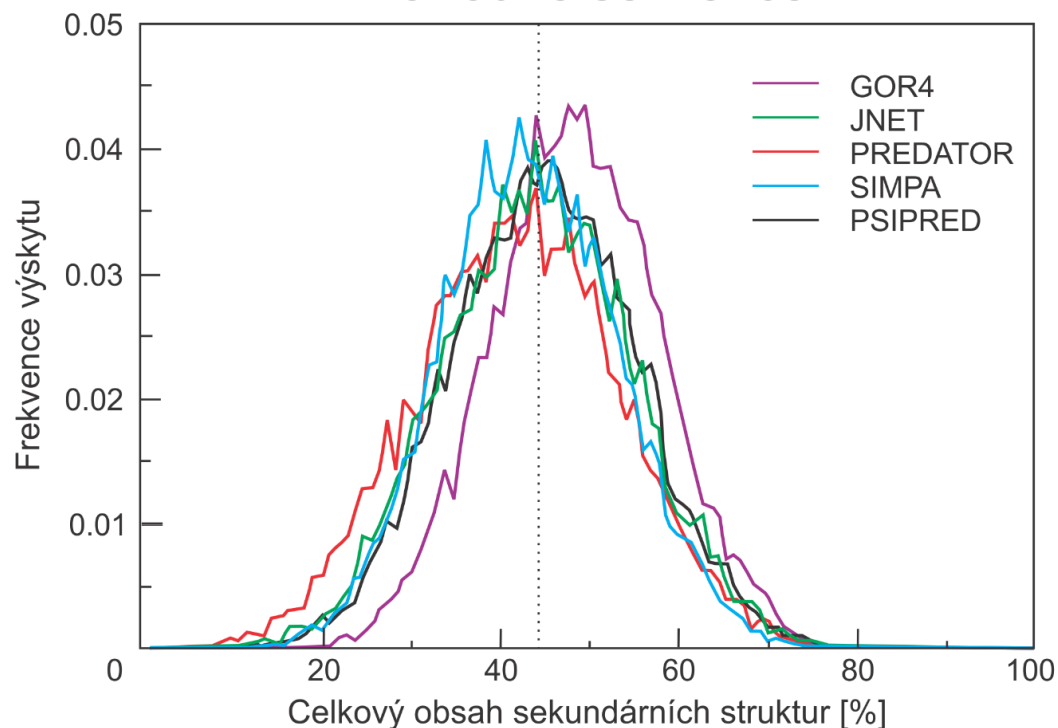
Jsme schopni předpovědět výskyt sekundární struktury v nepřírodní sekvenci?



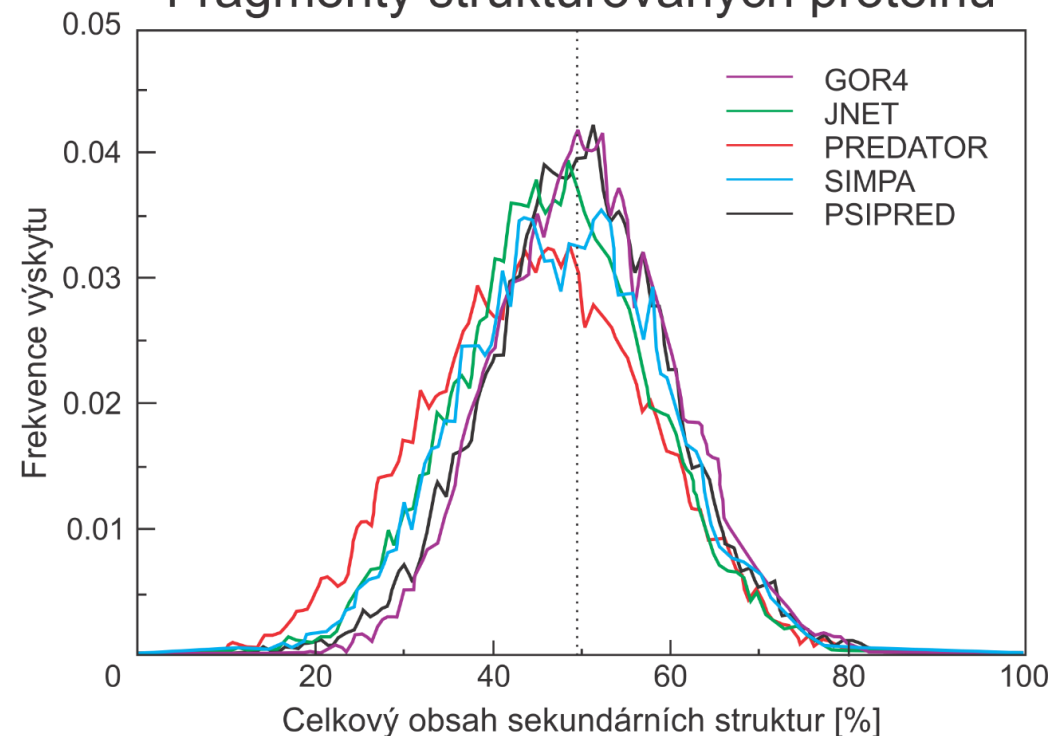
Strukturovanost prostoru náhodných sekvencí

Soubor 10 000 náhodných sekvencí, délka 100 aminokyselin, přírodní poměry obsahu aminokyselin

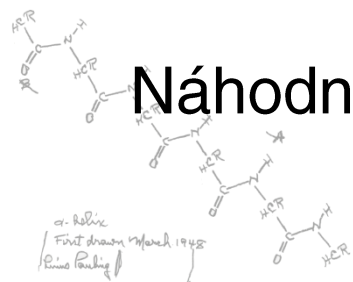
Náhodné sekvence



Fragmenty strukturovaných proteinů

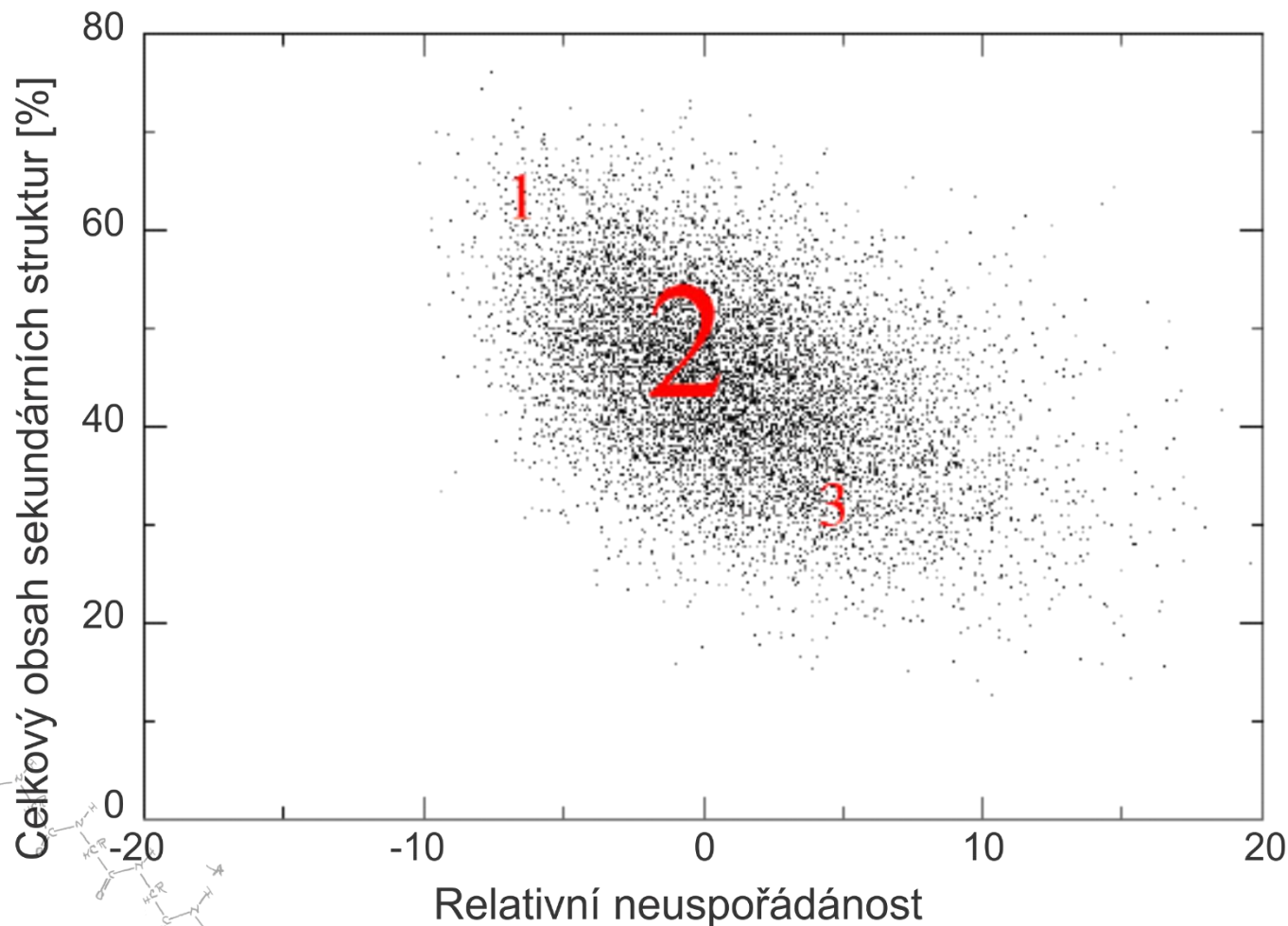


Náhodné sekvence vykazují podobnou distribuci výskytu sekundárních struktur jako existující proteiny



Rozdělení sekvenční knihovny

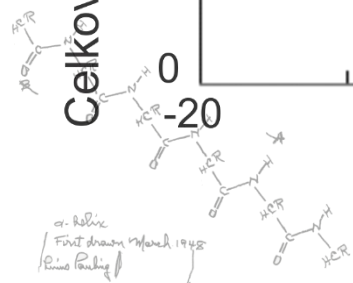
Soubor 10 000 náhodných sekvencí, délka 100 aminokyselin, přírodní poměry obsahu aminokyselin



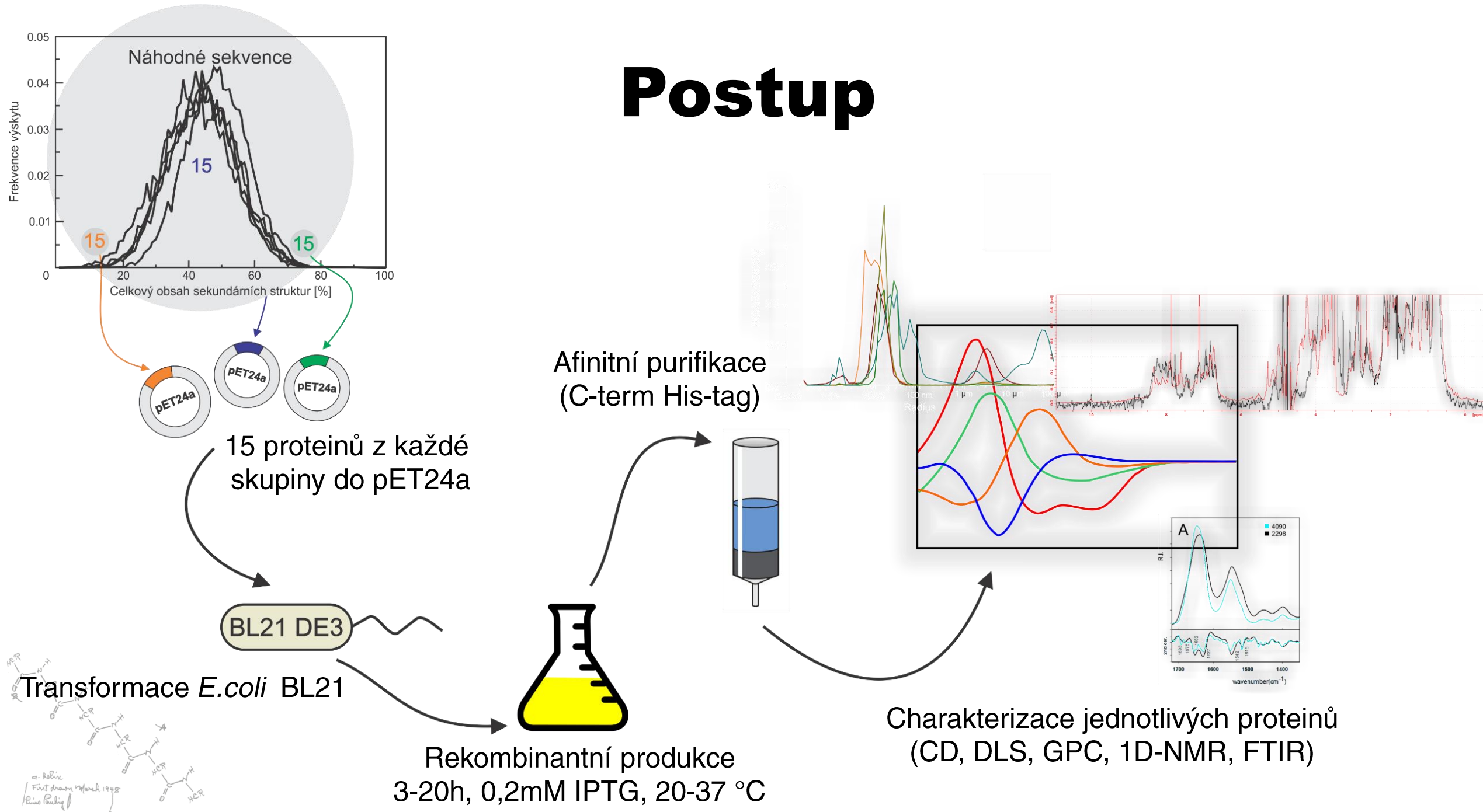
Skupina 1
nejvyšší obsah sekundárních struktur
nejnižší neuspořádanost

Skupina 2
náhodný výběr z celé knihovny

Skupina 3
nejnižší obsah sekundárních struktur
nejvyšší neuspořádanost



Postup



Rozpustnost náhodných sekvencí

Strukturované



Skupina 1

Nejvíce exprimovaných sekvencí
Nízká míra degradace

Náhodný výběr



Skupina 2

Stejné zastoupení rozpustných, degradovaných a uložených ve formě inkluzí

Nestrukturované



Skupina 3

Buď degradované nebo ~100% rozpustné v cytosolu

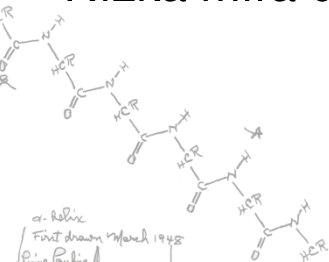
Vysoká rozpustnost

Střední rozpustnost

Nízká rozpustnost

Ner rozpustné

Degradace



A circular diagram for 'Skupina 1' (Group 1). The central circle is light grey and contains the text 'Skupina 1' in bold black font. Surrounding the center is a ring of 16 segments, each separated by a white gap. The segments are colored as follows, starting from the top and moving clockwise: 1. Dark green, 2. Medium green, 3. Light green, 4. Medium green, 5. Dark green, 6. Orange, 7. Dark grey, 8. Medium grey, 9. Dark grey, 10. Medium green, 11. Light green, 12. Medium green, 13. Dark green, 14. Medium green, 15. Light green, 16. Medium green.

$$\begin{array}{c} R \\ | \\ -C-N-H \\ | \\ HCR \end{array} - \begin{array}{c} O \\ || \\ -C-N-H \\ | \\ HCR \end{array} - \begin{array}{c} O \\ || \\ -C-N-H \\ | \\ HCR \end{array} - \begin{array}{c} O \\ || \\ -C-N-H \\ | \\ HCR \end{array} - \begin{array}{c} O \\ || \\ -C-N-H \\ | \\ HCR \end{array} - \begin{array}{c} O \\ || \\ -C-N-H \\ | \\ HCR \end{array}$$

a. 201x
 First drawn March 1948
 Linus Pauling

A circular diagram for 'Skupina 2' (Group 2). The circle is divided into 12 equal segments. The segments are colored in a sequence: 4 green segments, 4 orange segments, and 4 grey segments. The text 'Skupina 2' is centered in the middle of the circle.

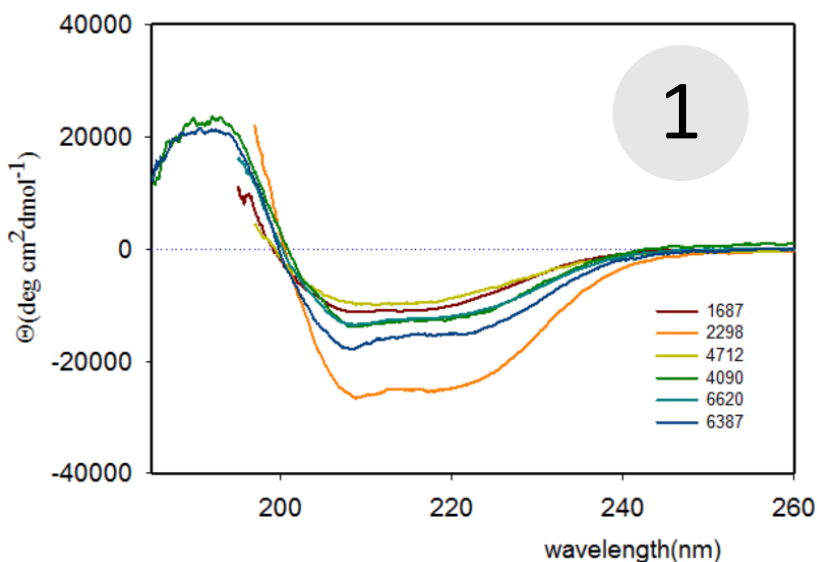
Bud' degradované nebo ~100% rozpustné v cytosolu

A circular diagram representing 'Skupina 3'. The center is a light gray circle with the text 'Skupina 3'. Surrounding this center is a ring of 12 segments. The segments are colored in a sequence: 5 green, 2 light green, and 5 orange.

-  Vysoká rozpustnost
-  Střední rozpustnost
-  Nízká rozpustnost
-  Nerozpustné
-  Degradace

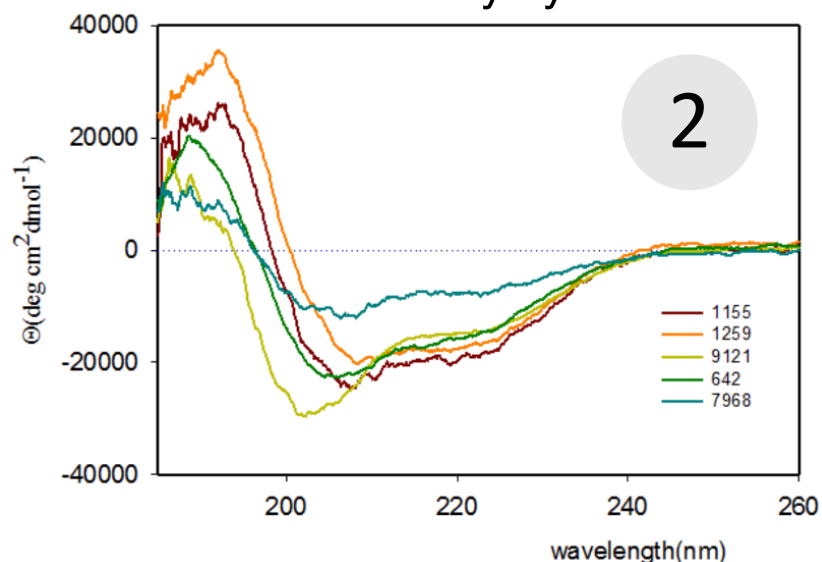
Zastoupení sekundárních struktur

Strukturované



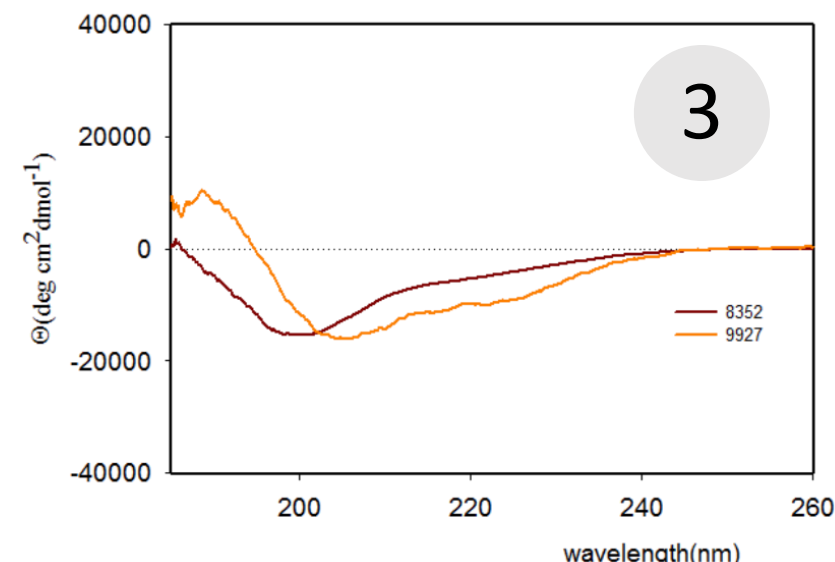
Skupina 1:

Náhodný výběr



Skupina 2:

Nestrukturované



Skupina 3:

Strukturované

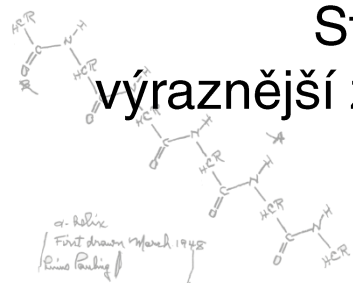
výraznější zastoupení alfa helixu

Méně strukturovaná

(posun minima k 200 nm)

Nejméně strukturovaná

(výrazný posun ke 200 nm,
nejnižší intenzity CD,
slabé maximum u 190 nm)

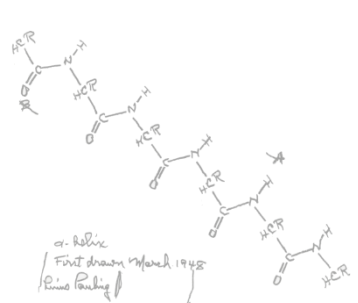


Něco navíc

Velikost částic: monodisperzní oligomery: ~10 až 600 jednotek

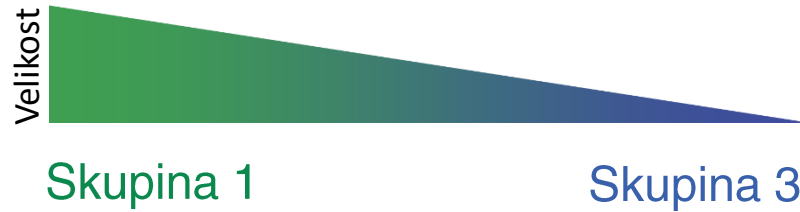


Strukturovanější proteiny – vyšší oligomerizace?



Něco navíc

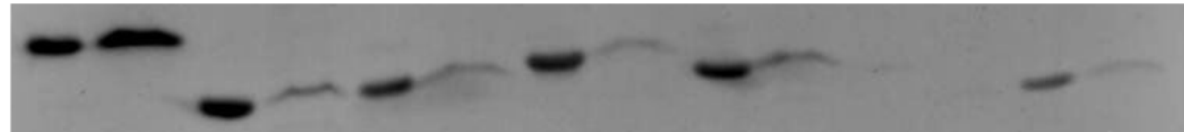
Velikost částic: monodisperzní oligomery: ~10 až 600 jednotek



Strukturovanější proteiny – vyšší oligomerizace?

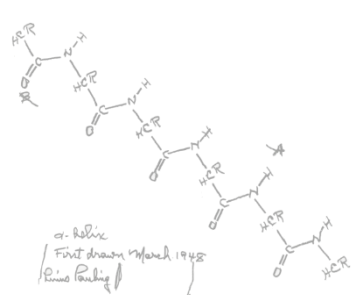
Skupina 1

<u>6387</u>	<u>7144</u>	<u>1687</u>	<u>4712</u>	<u>1856</u>	<u>8836</u>	<u>3027</u>
<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>



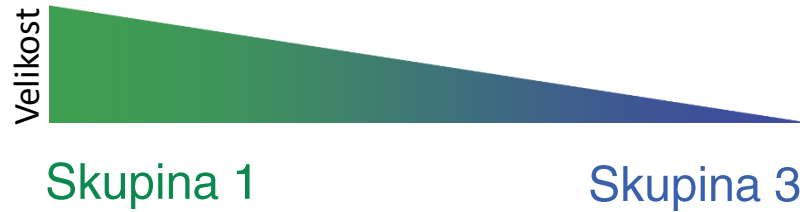
Skupina 3

<u>8352</u>	<u>8667</u>	<u>3703</u>	<u>665</u>	<u>8511</u>	<u>9927</u>	<u>6851</u>
<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>	<i>R</i>



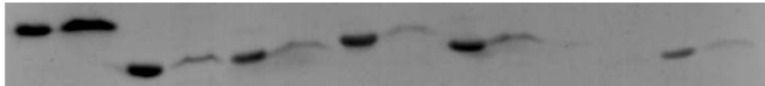
Něco navíc

Velikost částic: oligomery: ~10 až 600 jednotek



Strukturovanější proteiny – vyšší agregace?

Skupina 1 6387_R 7144_R 1687_R 4712_R 1856_R 8836_R 3027_R

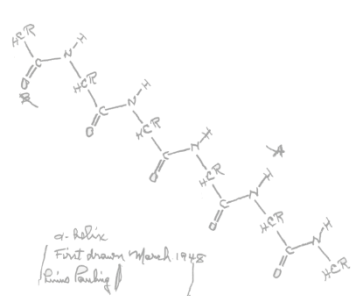
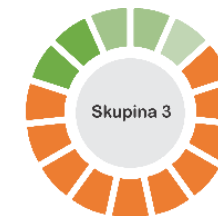
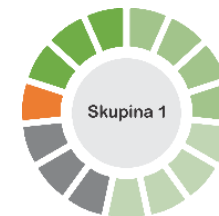


Skupina 3 8352_R 8667_R 3703_R 665_R 8511_R 9927_R 6851_R



Strukturované:
menší část produktu rozpustná, vyšší agregace

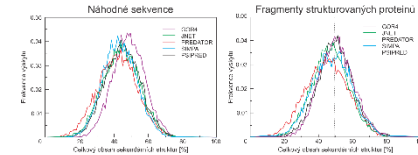
Nestrukturované:
rozpustné - lépe akceptované buňkou, více exprimované



Shrnutí

Sekundární struktura (SS) – obecná vlastnost polypeptidového řetězce

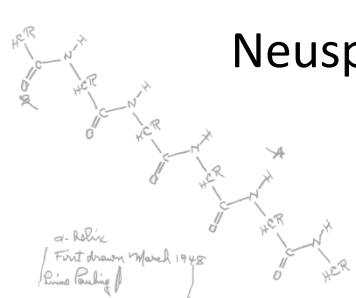
Výskyt SS v náhodných sekvencích $\sim 50\%$



Přes vysokou míru oligomerizace pozorována monodisperzní distribuce agregátů
(kvartérní struktura?)

Ověřena schopnost prediktorů pro odhalení strukturovaných oblastí v
přírodě neznámých sekvencích

Neuspořádané sekvence – buď 100% degradovány nebo buňkou tolerovány lépe než predikované
uspořádané (Never Born IDP?)



Na čem pracujeme?



Experimentální charakterizace celých knihoven náhodných sekvencí

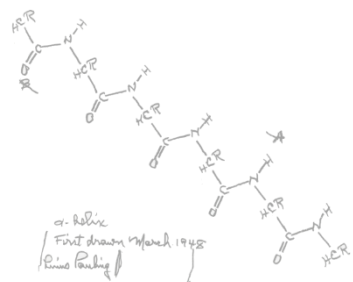
Fujishima, Kosuke, et al. "An overhang-based DNA block shuffling method for creating a customized random library." *Scientific reports* 5 (2015)



Reverzní evoluce proteinů mutagenézí moderních aminokyselin
za prebiotický set – mají současné proteiny prebiotického předka?

Charakterizace knihoven složených z různých aminokyselinových abeced
(rané, fyzchem minimum...)

Top down selekce uspořádaných sekvencí z experimentálních
knihoven proteinů



Klára Hlouchová
Jiří Vymětal
Lucie Bednarová
Vladimír Kopecký Jr.
Kateřina Hofbauerová
Jan Konvalinka
Jiří Vondrášek

Jiří Vymětal

Vladimír Kopecký Jr.

Kateřina Hofbauerov

Jan Konvalinka

Jiří Vondrášek

