

Sequencing of DNA and proteins

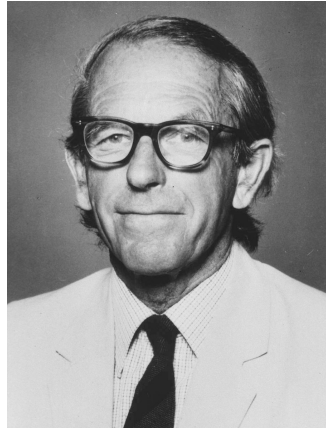
Dávid Tužinčin
460988@mail.muni.cz

C9531 Strukturní Biochemie -seminář

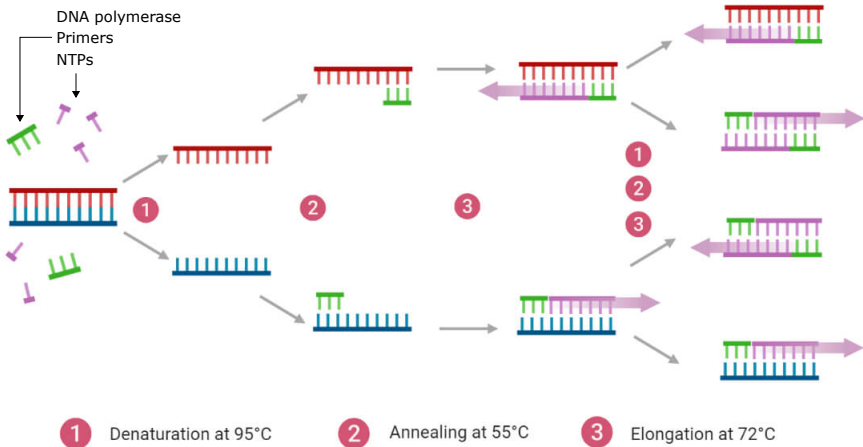
DNA sequencing

Sanger sequencing

- ▶ Frederik Sanger - britský biochemik v roce 1977 vyvinul metodu pro sekvenaci DNA - Nobelova cena za chemii
- ▶ metoda se využívá dodnes pro kratší sekvence

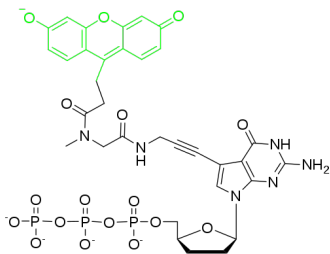


PCR - Polymerase Chain Reaction

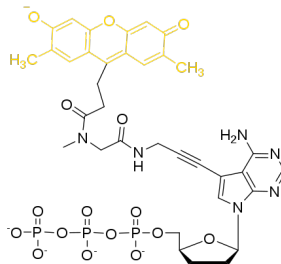


(Dahal, 2022)

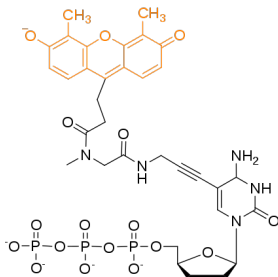
Sanger sequencing - ddNTPs



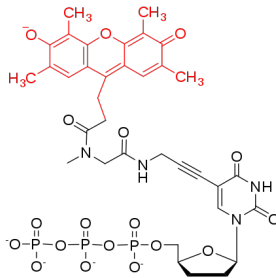
G-505



A-512



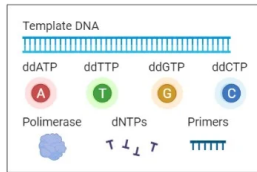
C-519



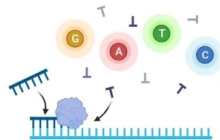
T-526

Sanger sequencing

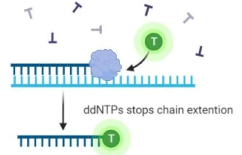
Reagents



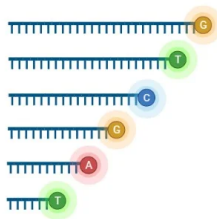
① Primer annealing and chain extension



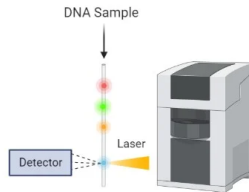
② ddNTP binding and chain termination



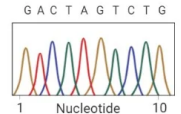
③ Fluorescently labelled DNA sample



④ Capillary gel electrophoresis and fluorescence detection



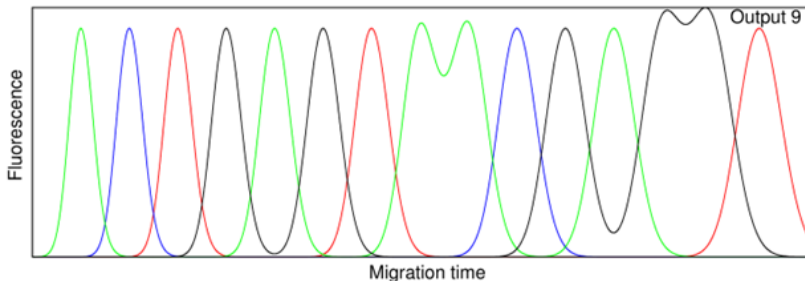
⑤ Sequence analysis and reconstruction



Genetic code

UUU } UUC } Phe UUA } UUG } Leu	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA } Stop UAG }	UGU } Cys UGC } UGA } Stop UGG } Trp
CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }
AUU } AUC } Ile AUA } AUG } Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }

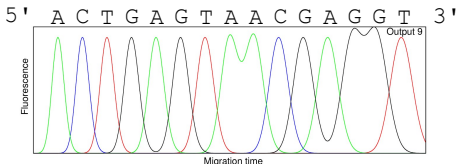
Exercise 1



1. zapište úsek sekvence vlákna DNA obsahujícího vaši cDNA, který odpovídá uvedenému fragmentu výstupu sekvenátoru:
2. zapište příslušný úsek sekvence mRNA, kódující náš protein:
3. zapište úsek sekvence proteinu, který je kódován tím úsekem vašeho genu, jehož sekvenci čteme ve výstupu sekvenátoru. Uveďte všechny aminokyseliny, které můžeme jednoznačně určit:

Postup: v úloze sekvenujeme

komplementární vlákno ke kódujícímu
vláknu



1. Přepíšeme si výstup ze
sekvenátoru →
5' ACTGAGTAACGAGGT 3'

2. Výstup ze sekvenátoru je komplementární sekvenci, která
byla sekvenována pomocí Sangerovi metody
→ 3' T G A C T C A T T G C T C C A 5'

3. Přepíšeme sekvenci z bodu 2. tak aby byla 5' → 3'
→ 5' A C C T C G T T A C T C A G T 3'

Tohle je sekvence, kterou sme posílali osekvenovat, a tudíž
řešením první úlohy (1)

4. Úloha (2). sekvence kódujícího vlákna je tedy komplementární k
nekódujícímu vláknu z úlohy (1). zároveň ho chceme zapsat od 5'
→ 3'. Je to tedy vlastně přesně sekvence, kterou sme získali ze
sekvenátoru, akorát v RNA je místo T → U
→ 5' A C U G A G U A A C G A G G U 3'

5. Úloha ③. Sekvenci mRNA z úlohy ② přeložíme pomocí genetického kódu na sekvenci aminokyselin.

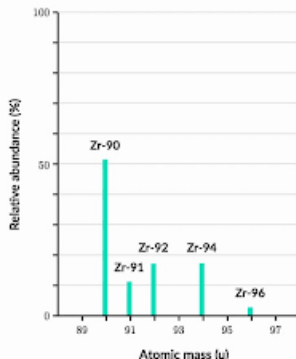
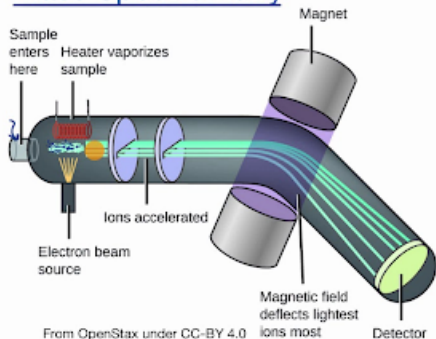
→ TE-RG

Protein sequencing

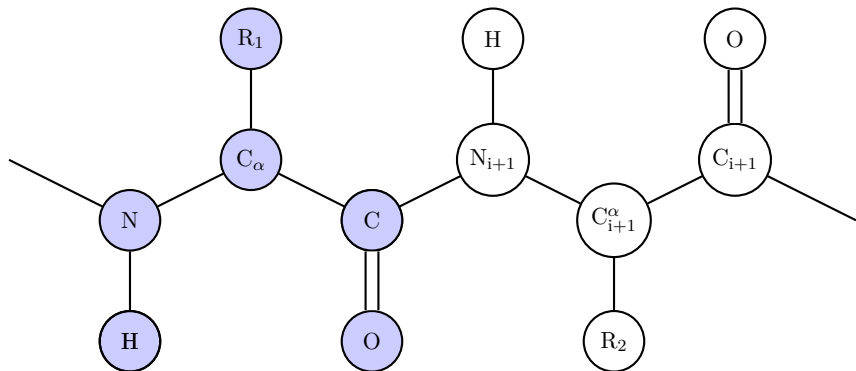
Mass Spectrometry

- ▶ 1. ionization (electrospray, MALDI)
- ▶ 2. fragmentation (collision induced fragmentation)
- ▶ 3. detection

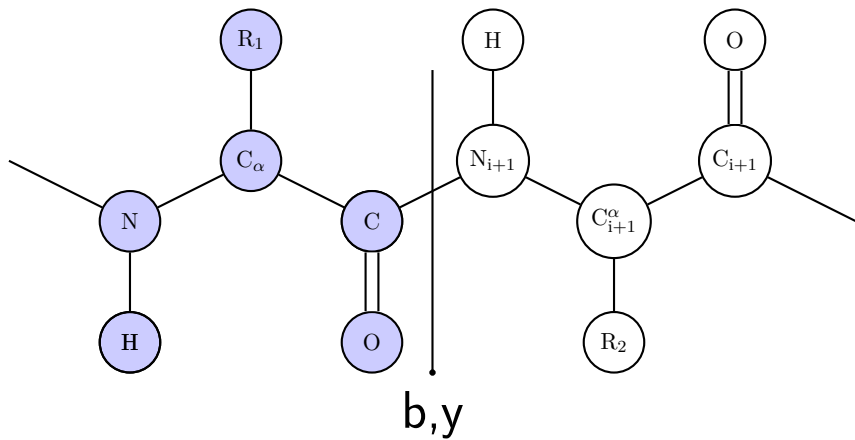
★ Mass Spectrometry



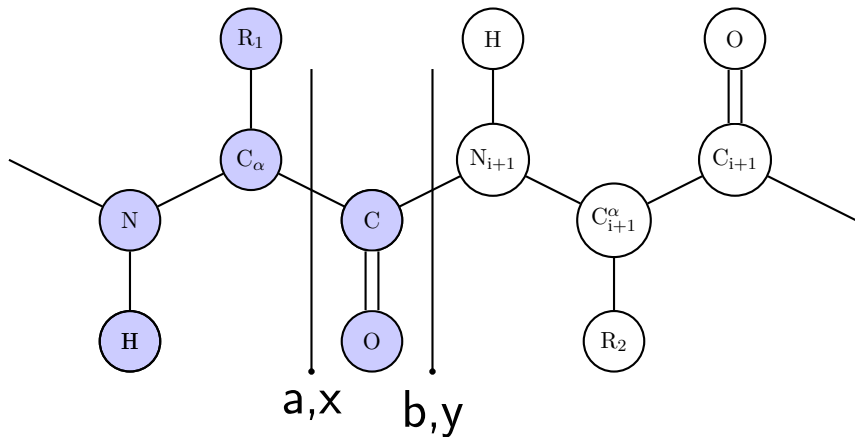
Peptide fragmentation



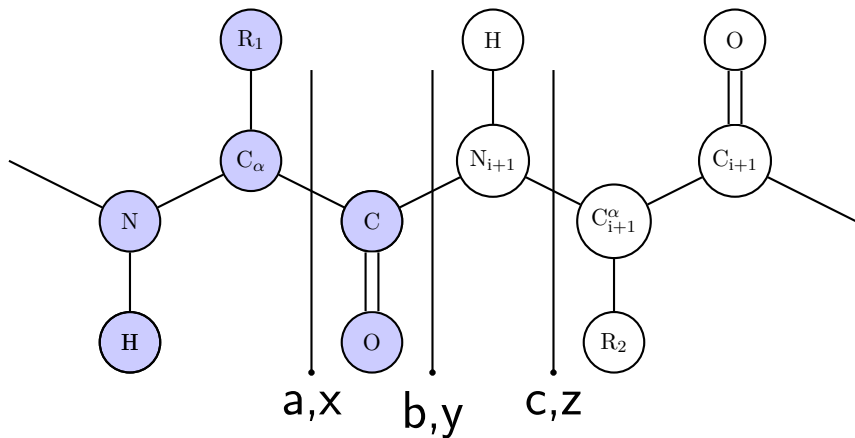
Peptide fragmentation



Peptide fragmentation



Peptide fragmentation



Peptide fragmentation

MS1 spectra: Precursor ions

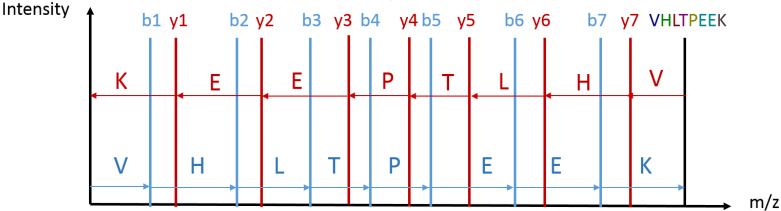
V-H-L-T-P-E-E-K



MS2 spectra: Precursor ions

b7	V-H-L-T-P-E-E	K	y1
b6	V-H-L-T-P-E	E-K	y2
b5	V-H-L-T-P	E-E-K	y3
b4	V-H-L-T	P-E-E-K	y4
b3	V-H-L	T-P-E-E-K	y5
b2	V-H	L-T-P-E-E-K	y6
b1	V	H-L-T-P-E-E-K	y7

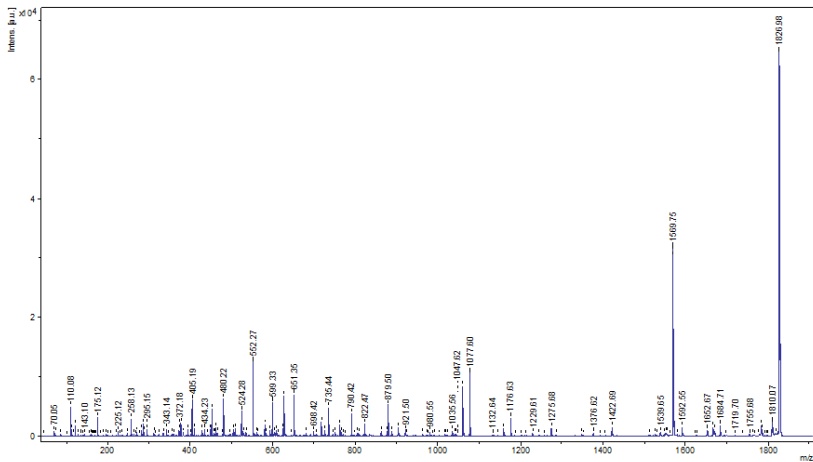
MS2 spectrum



MS example of real spectrum

Peptide $MH^+ = 1826.98$

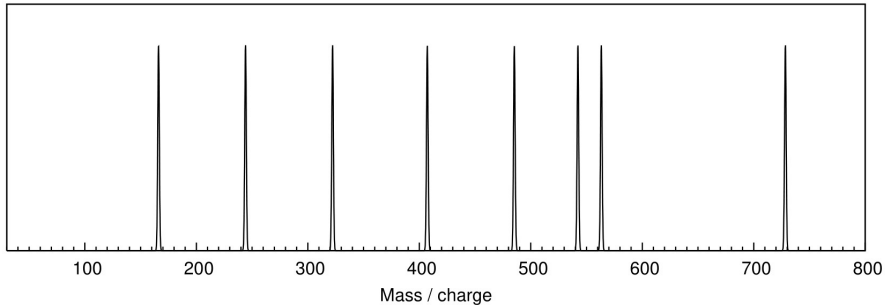
AADFFVVPTGSHFYLR



Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

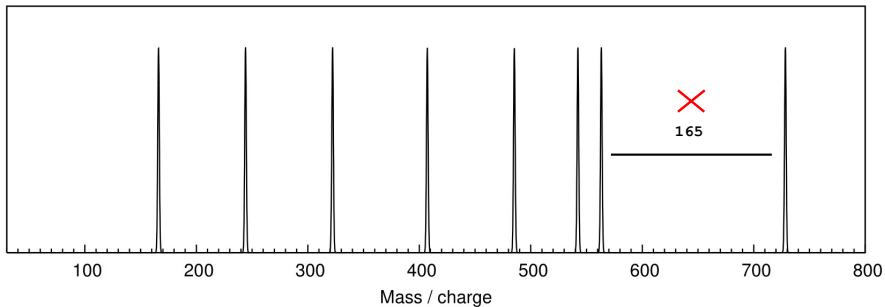
166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352



Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

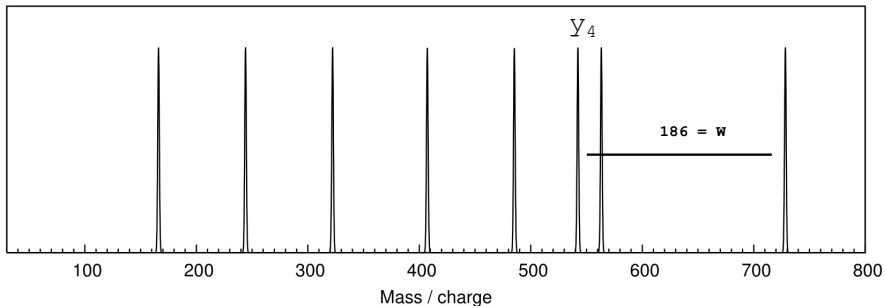


Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

W

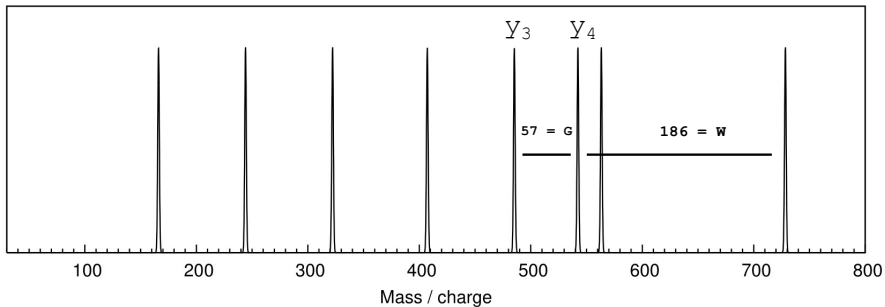


Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

W G

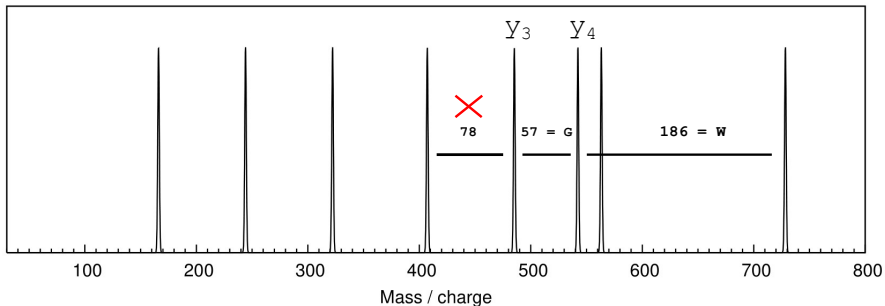


Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

W G

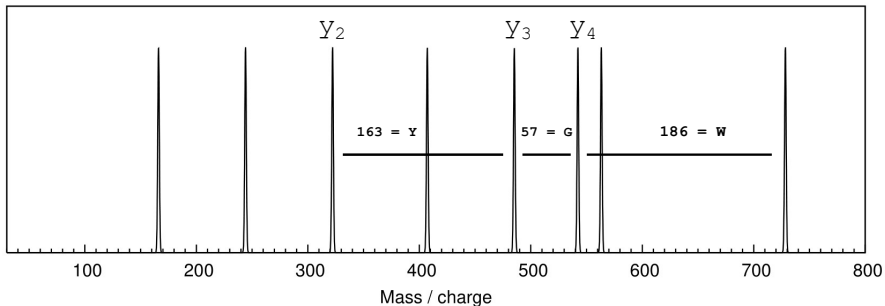


Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

W G Y

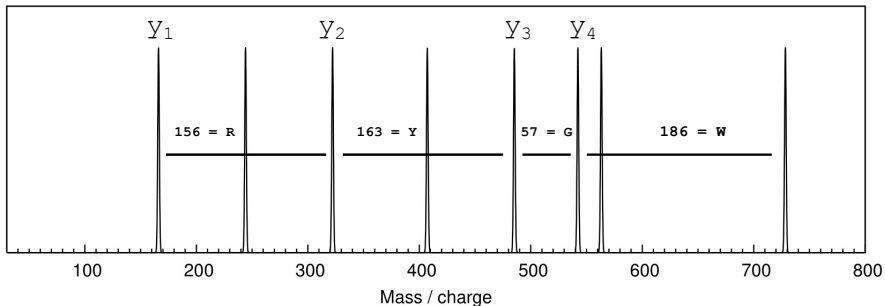


Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

W G Y R



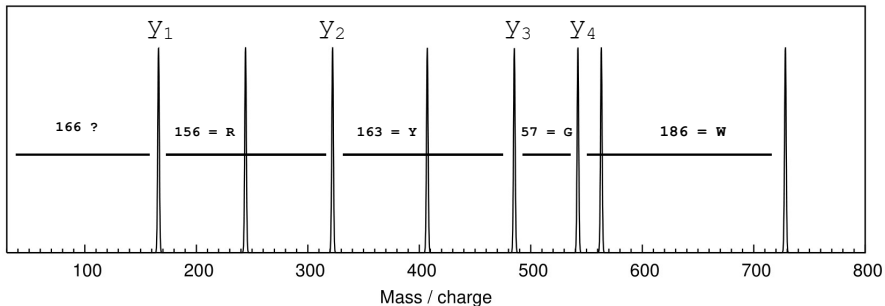
Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

W G Y R F

$$\begin{array}{r} 166 \\ -18 \\ \hline 147 = \mathbf{F} \end{array}$$



Exercise 2

Určete sekvenci pentapeptidu z hmotnostního spektra, ve kterém jsou přítomny signály celého peptidu a všech y a b fragmentu (krom fragmentu b1, který v praxi nevzniká).

166.087
244.109
322.188
407.172
485.251
542.273
563.273
728.352

W G Y R F

$$\begin{array}{r} 166 \\ -18 \\ \hline 147 = \mathbf{F} \end{array}$$

