

# Zastosowanie metod wykorzystujących połączenie technik separacyjnych i spektrometrii mas w naukach biomedyczo-farmaceutycznych

Wojciech Łuczaj, Agnieszka Gęgotek, Katarzyna Bielawska,  
Elżbieta Skrzydlewska

*Zakład Chemii Nieorganicznej i Analitycznej, Uniwersytet Medyczny w Białymstoku,*

e-materiały (wynik pracy intelektualnej) opracowane w ramach projektu  
Advanced Analytical Chemistry for Life Sciences [AACLifeSci]

Projekt został zrealizowany przy wsparciu finansowym Komisji Unii Europejskiej

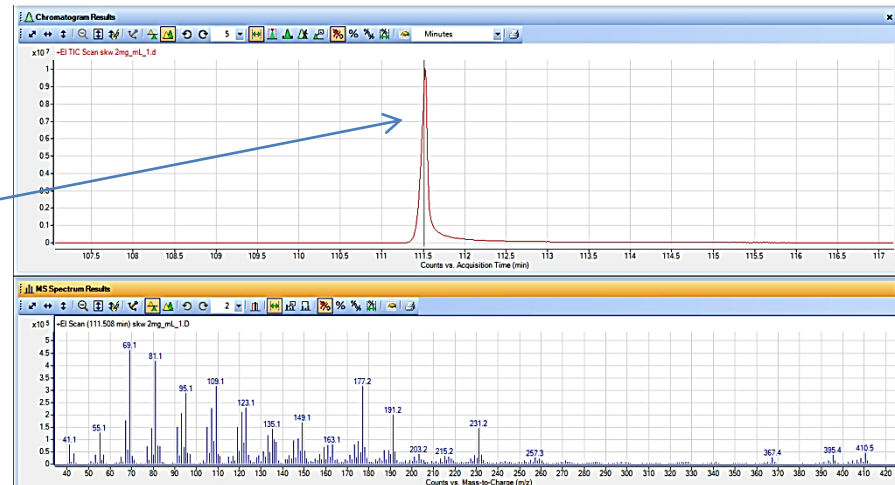
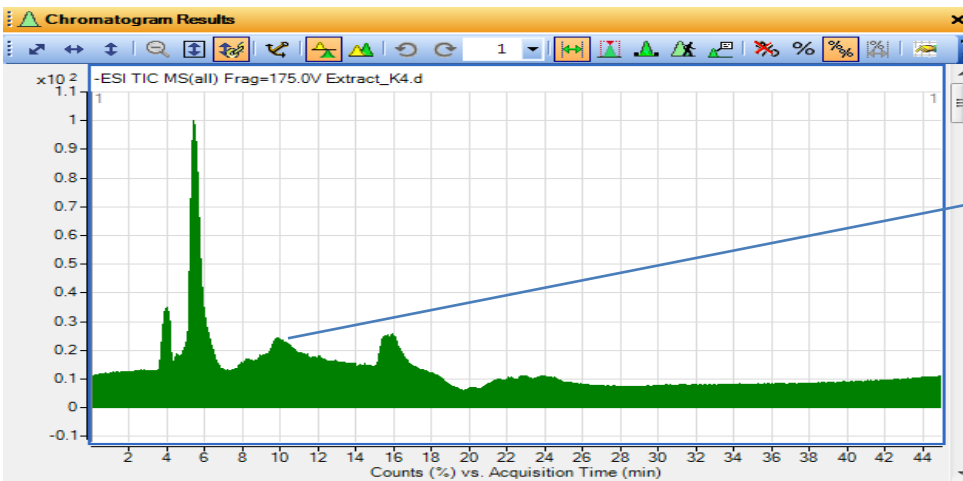


# Spektrometria mas

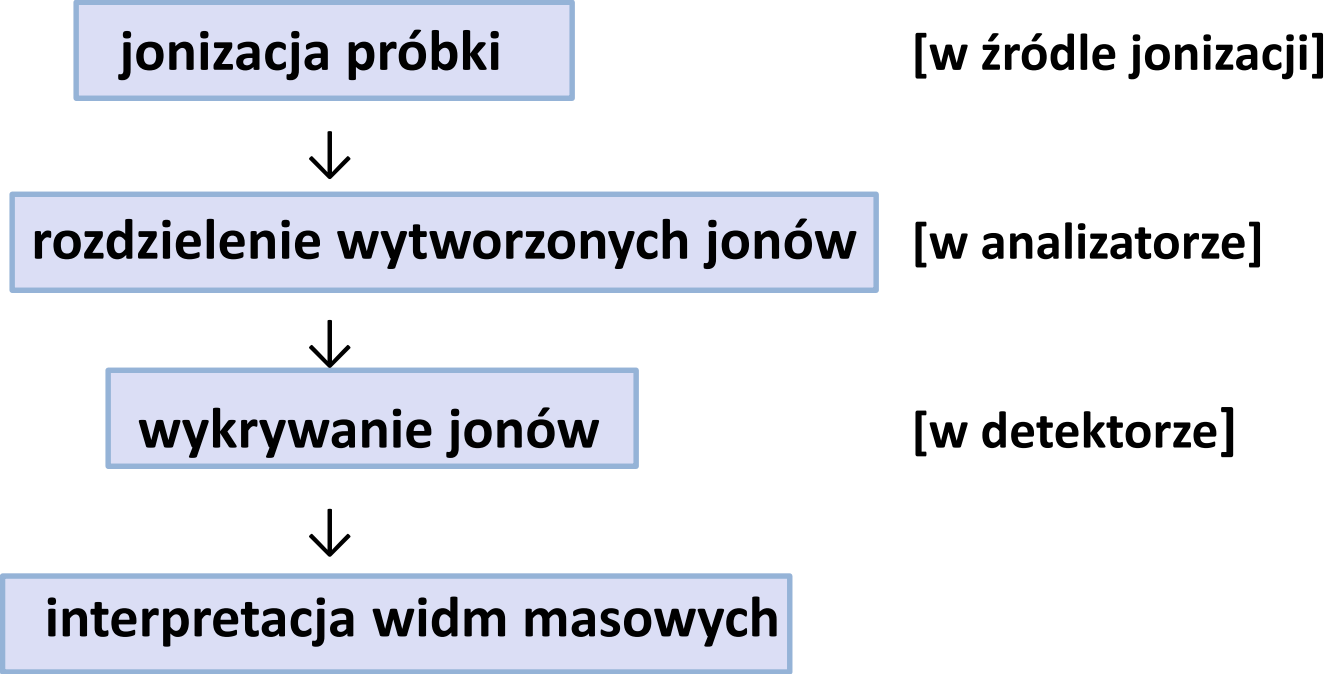
Technika analityczna, która umożliwia uzyskanie informacji na temat:

- ✓ masy cząsteczkowej
- ✓ struktury chemicznej
- ✓ ilości oznaczanych związków

Technika analityczna umożliwiająca rozdzielenie zjonizowanych związków na podstawie stosunku ich ładunku do masy ( $m/z$ )

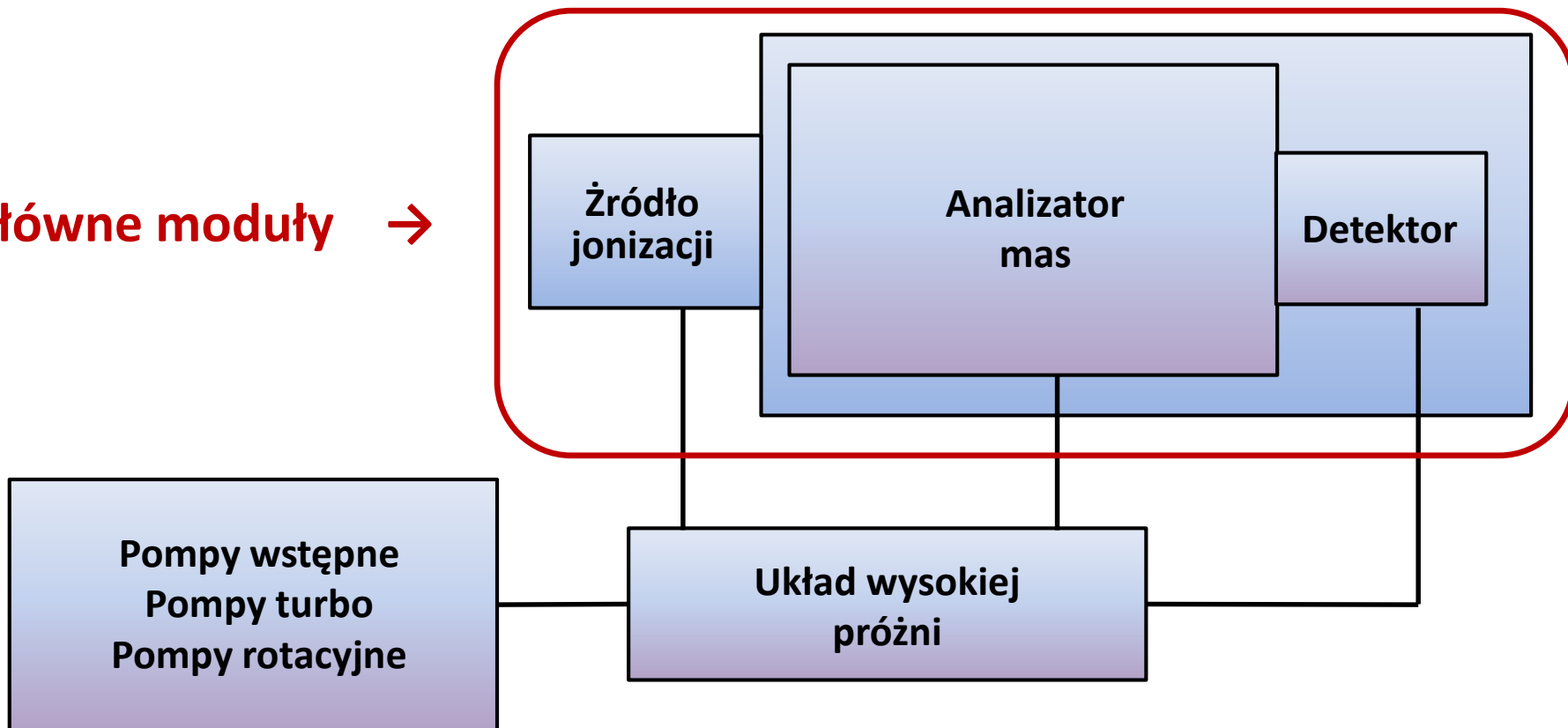


# Etapy procedury analitycznej opartej o spektrometrię mas:



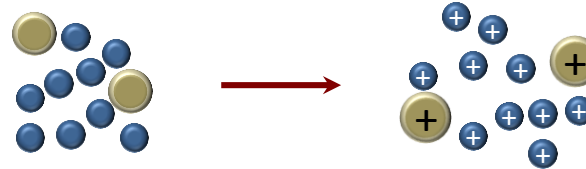
# Budowa spektrometru mas

Główne moduły →

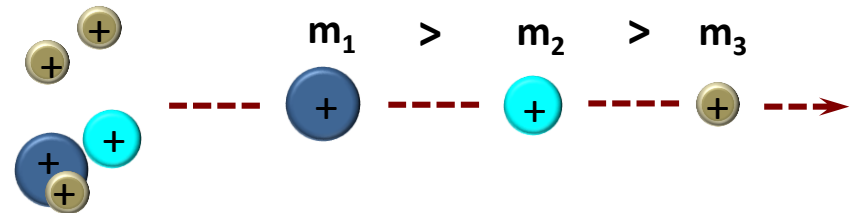


# Procesy zachodzące w trakcie analizy MS

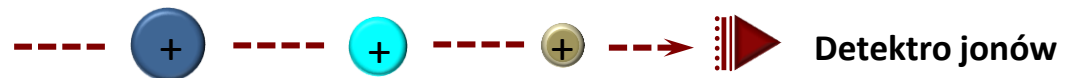
## 1. Jonizacja analitu



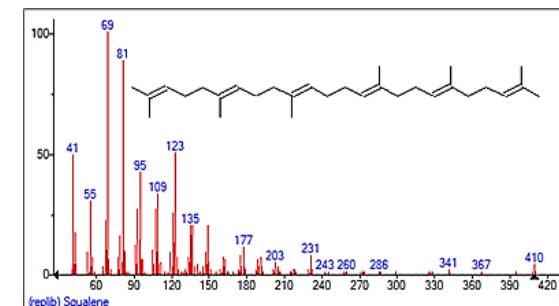
## 2. Rozdzielenie wytworzonych jonów pod względem ich stosunku $m/z$



## 3. Wykrywanie jonów



## 4. Analiza i interpretacja uzyskanych widm masowych



## Metody jonizacji

- jonizacja elektronowa [EI]
- jonizacja typu elektrosprej [ESI]
- laserowa jonizacja wspomagana matrycą [MALDI]
- Jonizacja chemiczna pod ciśnieniem atmosferycznym [APCI]

**Jonizacja twarda** – produkty: **jon molekularny** + **jony fragmentów cząsteczki**

✓ EI

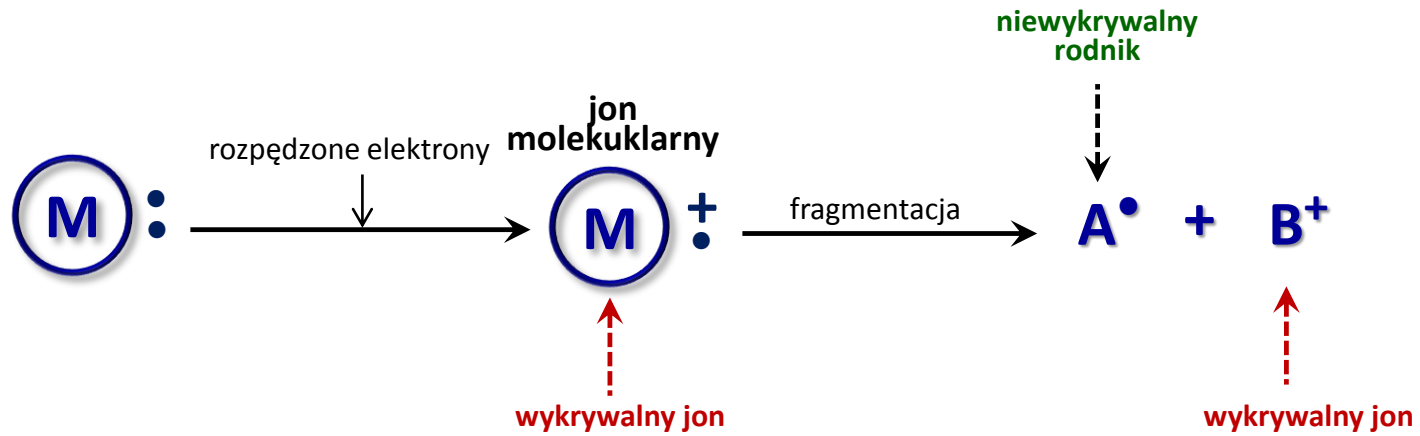
**Jonizacja miękka** – główny produkt – **jon molekularny**

✓ ESI

✓ MALDI

✓ APCI

# Jonizacja elektronowa (EI)



*Jonizacja i fragmentacja w źródle typu EI*

# Jonizacja elektronowa (EI)

## ➤ Analit:

- małowcząsteczkowe,
- niepolarne,
- lotne,
- stabilne termicznie

## ➤ Zakres mas:

- <1 kDa

## ➤ Sposób wprowadzania próbki:

- GC lub próbki ciekłe/stałe

## ➤ Zalety:

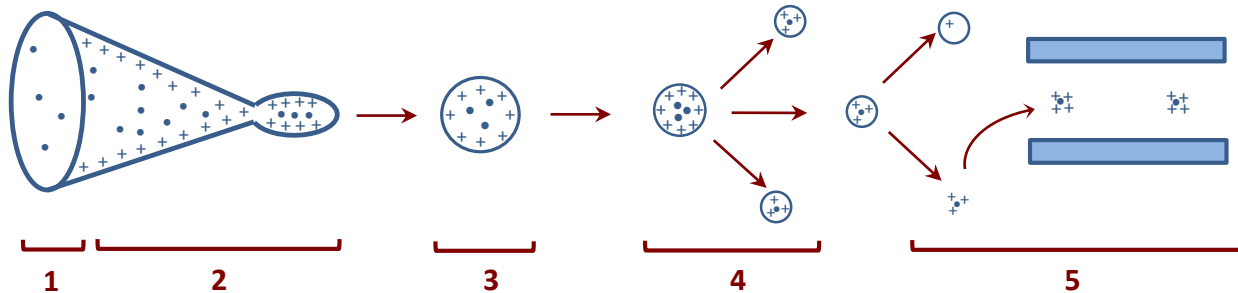
- analiza związków niepolarnych,
- brak supresji jonów,
- łatwa do aplikacji z GC,
- istnienie bibliotek widm

## ➤ Wady:

- analiza
  - ✓ tylko związków lotnych,
  - ✓ wymagana stabilność termiczna,
  - ✓ związki małowcząsteczkowe,
- jonizacja twarda,



## Jonizacja typu elektrosprej (ESI)



- 1- utworzenie jonów,
- 2- utworzenie naładowanych na powierzchni kropli,
- 3- odparowanie rozpuszczalnika,
- 4- „wyładowania Coulomba”,
- 5- utworzenie jonów rozproszonych w fazie gazowej

***Mechanizm jonizacji typu elektrosprej (ESI)***

# Jonizacja typu elektrosprej (ESI)

## ➤ Analit:

- związki polarne  
np. peptydy, białka, cukry, nukleotydy

## ➤ Zakres mas:

- <200 kDa

## ➤ Sposób wprowadzania próbki:

- LC lub roztwór próbki

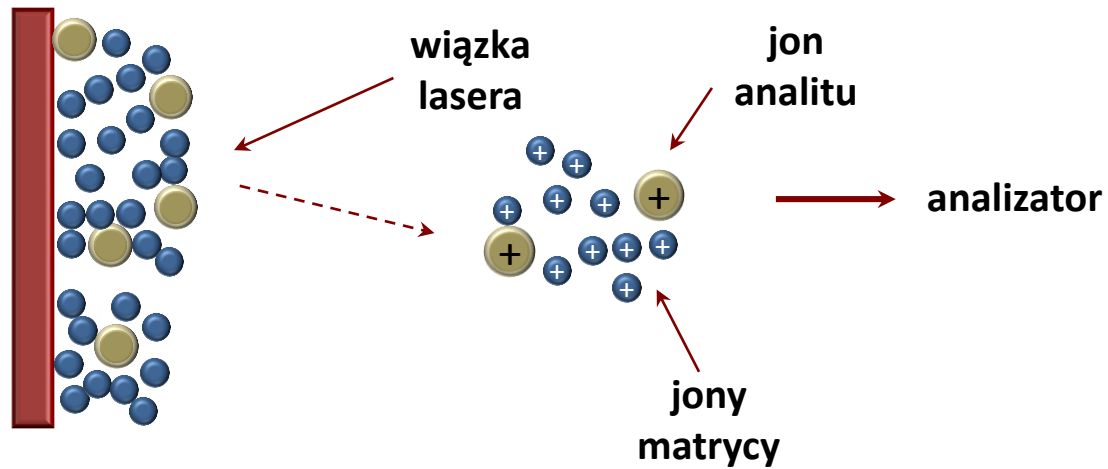
## ➤ Zalety:

- związki termicznie labilne
- związki wielkocząsteczkowe
- tworzenie jonów naładowanych wielokrotnie
- wysoka czułość
- łatwe do połączenia z LC
- jonizacja miękka

## ➤ Wady:

- wymagana rozpuszczalność w rozpuszczalnikach polarnych
- wrażliwa na obecność soli w roztworze próbki
- supresja jonów

# Laserowa jonizacja wspomagana matrycą (MALDI)



*Mechanizm laserowej jonizacji wspomaganej matrycą (MALDI)*

# Laserowa jonizacja wspomagana matrycą (MALDI)

## ➤ Analit:

- związki polarne  
np. peptydy, białka, cukry, nukleotydy

## ➤ Zakres mas:

- <500 kDa

## ➤ Sposób wprowadzania próbki:

- związki rozpuszczane w stałej matrycy

## ➤ Zalety:

- thermolabile compounds,
- związki wielkocząsteczkowe,
- wysoka czułość,
- mniejsza wrażliwość na obecność soli,
- jonizacja miękka

## ➤ Wady:

- szeroki wybór matryc,
- trudności w analizach ilościowych,
- supresja jonów

## Jonizacja chemiczna pod ciśnieniem atmosferycznym (APCI)

### ➤ Analit:

- związki polarne  
np. peptydy, białka, cukry, nukleotydy

### ✓ Zakres mas:

- <1 kDa

### ✓ Sample introduction:

- LC lub roztwór próbki

### ➤ Zalety:

- związki labilne termicznie,
- wysoka czułość,
- stosowanie dużych przepływów f.r.,
- łatwe do połączenia z LC,
- jonizacja miękka

### ➤ Wady:

- wymagana rozpuszczalność w rozpuszczalnikach polarnych,
- wrażliwa na obecność soli w roztworze próbki,
- supresja jonów

## Analizatory mas

Rozdzielają jony pod względem ich stosunku  $m/z$

- pracują pod wysoką próżnią
- najważniejsze parametry:
  - rozdzielczość
  - dokładność pomiaru masy
  - czułość
  - zakres pomiaru masy

## Rozdzielczość

możliwość rozdzielenia dwóch sygnałów o bardzo zbliżonych właściwościach

$$R = \Delta m/m$$

gdzie zdolność rozdzielczą definiuje się jako:

$$m_1/(m_2 - m_1)$$

$m_1$  -  $m/z$  lżejszego jonu;  $(m_2 - m_1)$  - różnica  $m/z$  następujących po sobie jonów

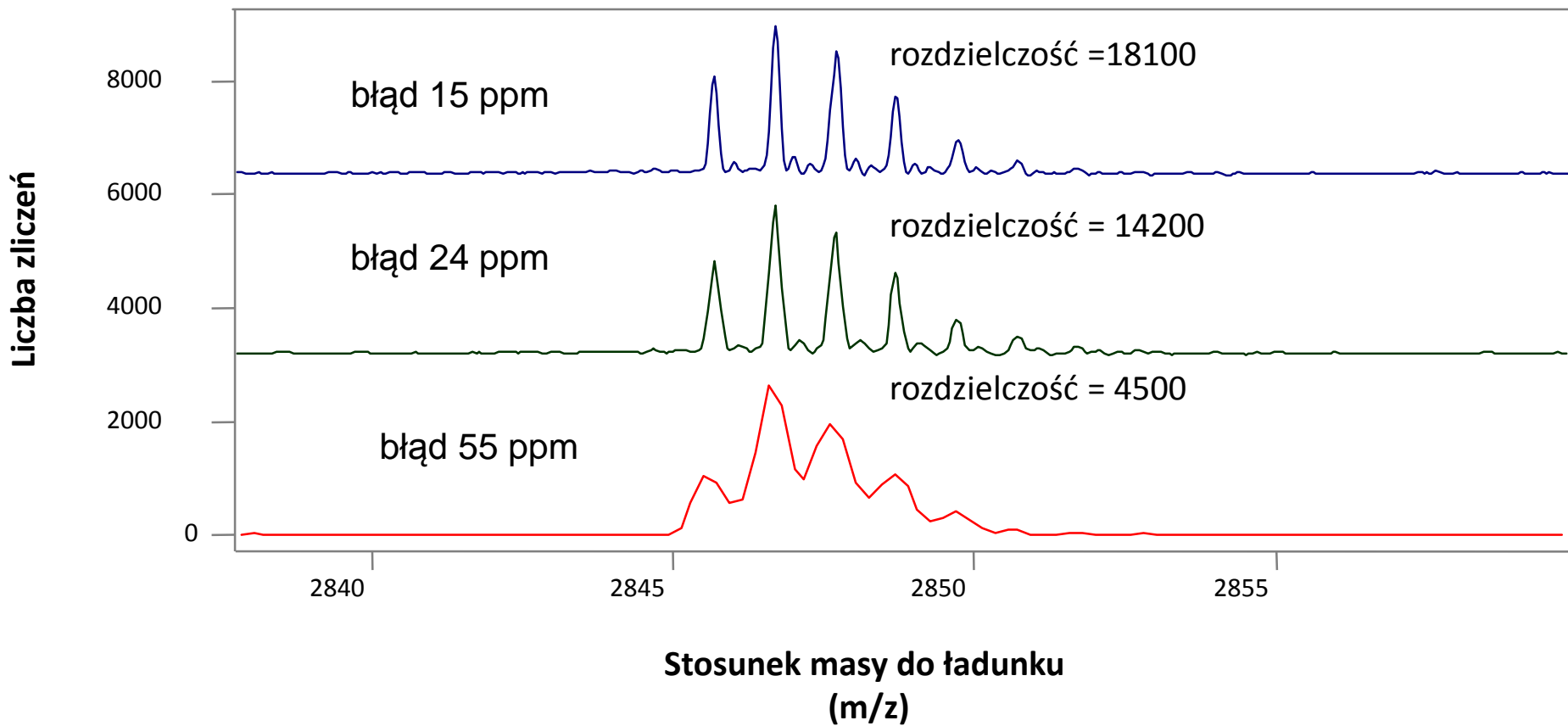
## Dokładność pomiaru masy

informuje w jakim stopniu zmierzona masa zgodna jest z wartością rzeczywistą

$$\frac{(\text{wyliczona teoretycznie masa monoizotopowa} - \text{dokładna masa zmierzona})}{\text{wyliczona teoretycznie masa monoizotopowa} \times 10^6}$$

- wyrażana w [ppm]

Im większa rozdzielczość tym mniejszy błąd pomiaru masy





## Czułość

**charakteryzuje wielkość sygnału, jako odpowiedź detektora w zależności od stężenia analitu**

- określa limit detekcji (LOD)

## Liniowy zakres dynamiczny

**zakres, w którym sygnał pochodzący od jonów analitu jest wprost proporcjonalny do stężenia tego analitu**

- kluczowe znaczenie w zapewnieniu dokładności pomiarów (analiza ilościowa)

# Analizatory mas

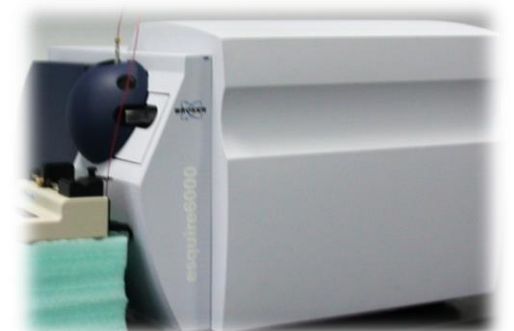
Spektrometria mas

analizatory mas

Najczęściej stosowane analizatory mas:



**Kwadrupol (Q)**



**Pułapka jonowa (IT)**



**Czasu przelotu (TOF)**

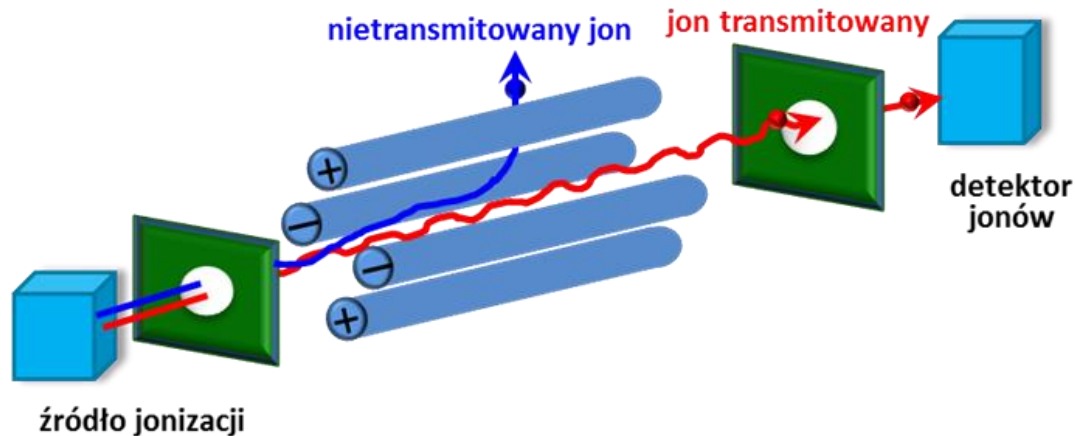


**Orbitrap**

# Analizatory mas

## Kwadrupol (Q):

- zbudowany z czterech równoległe ułożonych prętów
- wykorzystuje zarówno napięcie stałe jak i zmienne w celu filtrowania mas
- posiada kilka trybów pracy:
  - ✓ skanowanie (SCAN),
  - ✓ monitorowanie wybranych jonów (SIM)
- **niska rozdzielczość**
- **najwyższa możliwa czułość (analizy ilościowe)**



*Budowa kwadrupolowego analizatora mas*

# Analizatory mas

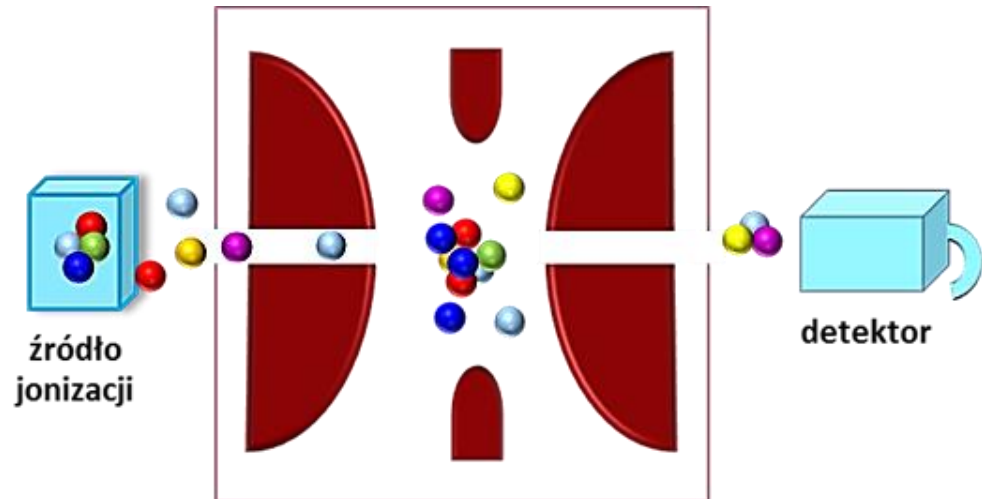
## Kwadrupol (Q):

Parametr	Kwadrupol
Szybkość zbierania danych (Hz)	<b>2-10</b>
Dokładność pomiaru masy (ppm)	<b>niska</b>
Zakres pomiaru masy (m/z)	<b>&lt;3000</b>
Rozdzielczość	<b>jednostka</b>

# Analizatory mas

## Pałapka jonowa (IT):

- wiąże jony wykorzystując kwadrupolowe pole magnetyczne
- dwa typy:
  - ✓ pałapka jonowa typu 2D (liniowa pałapka jonowa)
  - ✓ pałapka jonowa typu 3D (kwadrupolowa pałapka jonowa)
- **niska rozdzielczość**
- **duża szybkość skanowania**



*Schemat analizatora mas typu pałapki jonowej 3D*

## Analizatory mas

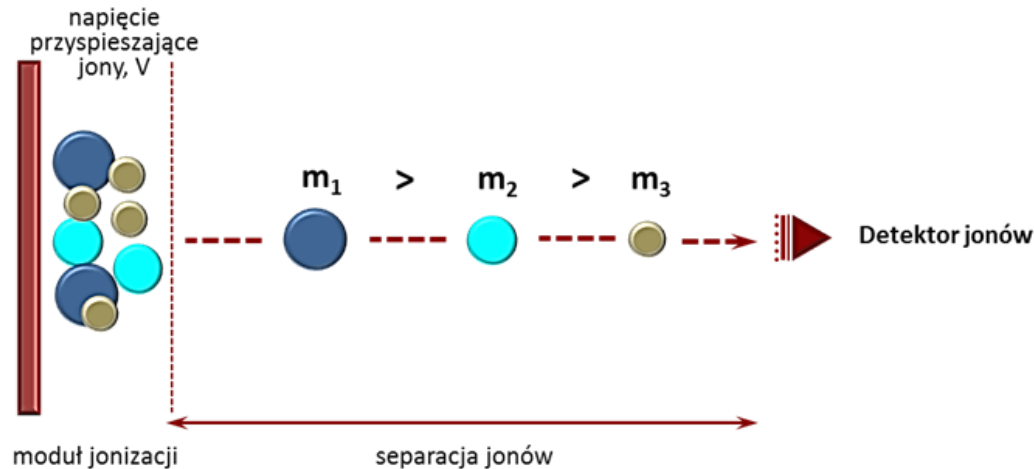
### Pułapka jonowa (IT)

Parametr	Pułapka jonowa
Szybkość zbierania danych (Hz)	<b>2-10</b>
Dokładność pomiaru masy (ppm)	<b>niska</b>
Zakres pomiaru masy (m/z)	<b>&lt;6000</b>
Rozdzielczość	<b>jednostka</b>

# Analizatory mas

## Czasu przelotu (TOF):

- jony wytwarzane są w pakietach
- mierzy czas jaki potrzebują jony aby dostać się do detektora
- lżejsze jony trafiają do detektora szybciej niż jony o większej masie
- **wysoka rozdzielczość**
- **duża dokładność pomiaru masy**
- **wysoka czułość**



*Mechanizm separacji jonów w analizatorze mas czasu przelotu (TOF)*

## Analizatory mas

### Czasu przelotu (TOF)

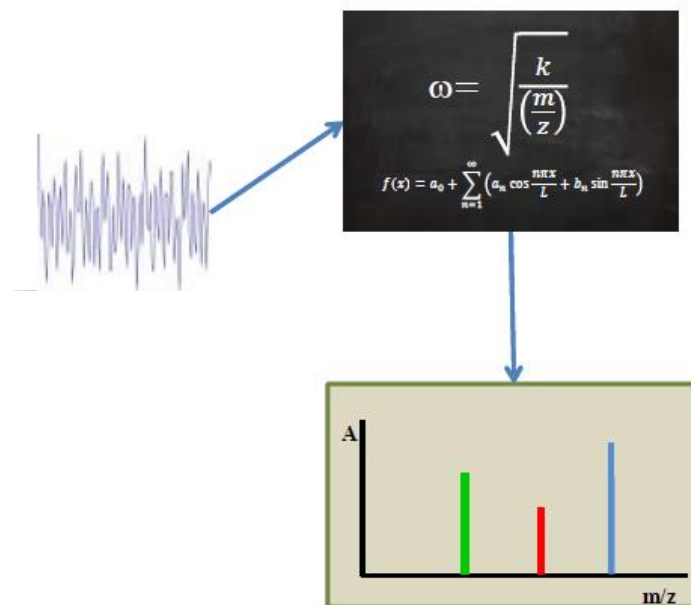
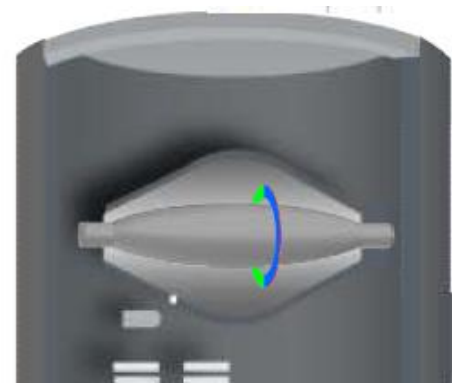
Parametr	TOF
Szybkość zbierania danych (Hz)	10-100
Dokładność pomiaru masy (ppm)	1-10 ppm
Zakres pomiaru masy (m/z)	<100,000 bez ograniczeń
Rozdzielczość	<50,000



# Analizatory mas

## Orbitrap:

- zbudowany z elektrod kształtem przypominających beczki
- wartość stosunku  $m/z$  obliczana jest przy użyciu transformacji Fouriera w oparciu o częstotliwość zatrzymywanych jonów
- **wysoka czułość**
- **duża dokładność pomiaru masy**
- **wysoka czułość**



## Analizatory mas

### Orbitrap

Parametr	Orbitrap
Szybkość zbierania danych (Hz)	1-18
Dokładność pomiaru masy (ppm)	1-5 ppm
Zakres pomiaru masy (m/z)	<6000
Rozdzielczość	<500,000

## Porównanie parametrów różnych analizatorów mas

Analizator	Q	IT	TOF	Orbitrap
<b>Zalety</b>	<p>łatwa aplikacja różnych źródeł jonizacji,</p> <p>szeroki zakres dynamiczny</p>	<p>łatwa aplikacja różnych źródeł jonizacji,</p> <p>szeroki zakres dynamiczny,</p> <p>MS<sup>n</sup></p>	<p>duża szybkość skanowania,</p> <p>szeroki zakres mas,</p> <p>duża dokładność pomiaru masy</p>	<p>duża dokładność pomiaru masy,</p> <p>możliwość analizy w dwóch polaryzacjach jednocześnie</p>
<b>Wady</b>	<p>niska rozdzielczość,</p> <p>mała dokładność pomiaru masy,</p> <p>wąski zakres mas,</p> <p>mała szybkość skanowania,</p> <p>MS/MS wymaga kilku analizatorów</p>	<p>niska rozdzielczość,</p> <p>mała dokładność pomiaru masy,</p> <p>wąski zakres mas,</p> <p>mała szybkość skanowania,</p>	<p>węższy zakres dynamiczny w porównaniu do Q,</p>	<p>mniejsza szybkość skanowania w porównaniu do QTOF,</p> <p>węższy zakres dynamiczny w porównaniu do Q,</p>

## Tandemowa spektrometria mas (MS/MS)

### Tandemowe spektrometry mas

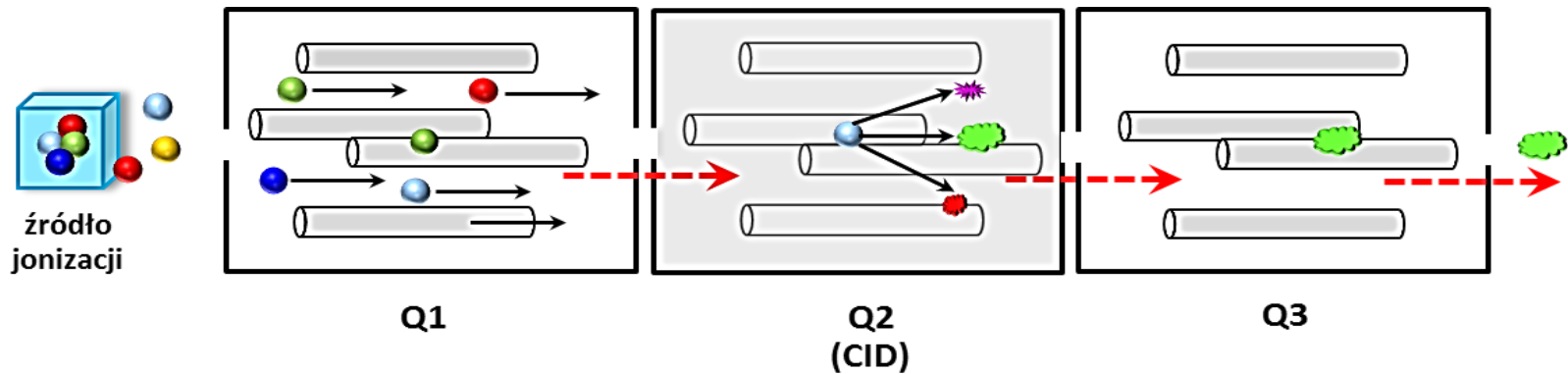
Potrójny kwadrupol [QqQ]

Kwadrupolowy analizator czasu przelotu [QqTOF]

Pułapka jonowa [IT]

Kwadrupolowy analizator typu orbitrap [QqOrbitrap]

## Tandemowa spektrometria mas



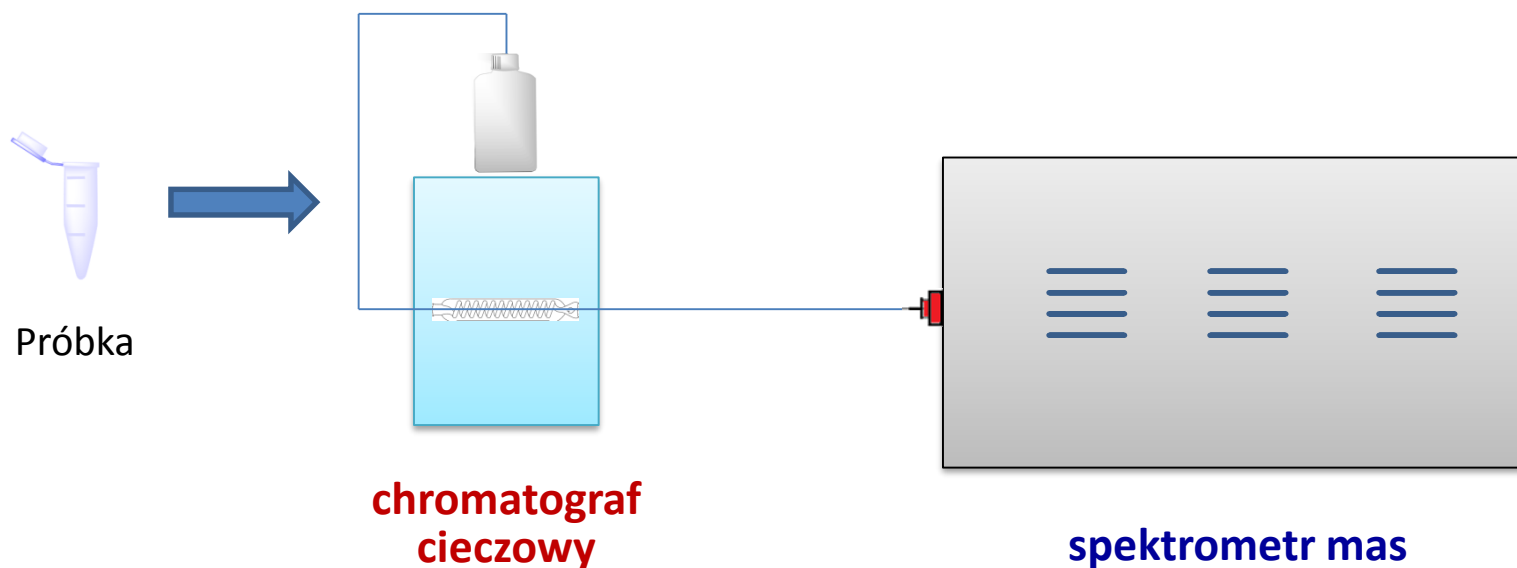
*Tandemowa spektrometria mas w oparciu o działanie potrójnego kwadrupola [QqQ]*

### Trypy analiz w trybie MS/MS:

- monitorowanie wybranych/wielokrotnych reakcji fragmentacji (**selected/multiply reaction monitoring SRM/MRM**)
- skanowanie jonów wtórnych (product ion scan)
- skanowanie jonów pierwotnych/macierzystych (precursor ion scan)
- utrata cząsteczki obojętnej (neutral loss scan)

## Wprowadzenie do techniki LC-MS

Technika LC-MS łączy możliwość rozdziału analitów z użyciem chromatografii ciekowej (LC) z możliwościami analitycznymi spektrometrii mas (MS)



LC pozwala na rozdział wielu związków na podstawie różnic w ich czasach retencji ( $t_R$ ).

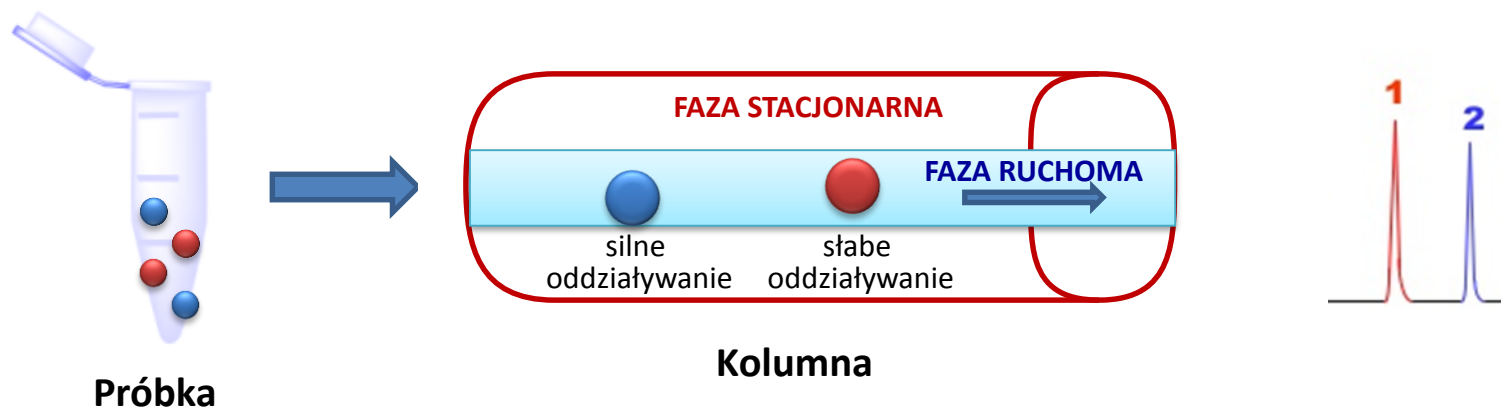
LC-MS pozwala na rozdział związków o zbliżonych czasach retencji, ale różnych  $m/z$  ulegających fragmentacji w różny sposób.

# Wprowadzenie do chromatografii cieczowej [LC]

LC – rodzaj chromatografii, w której fazą ruchomą jest ciecz ("eluent")

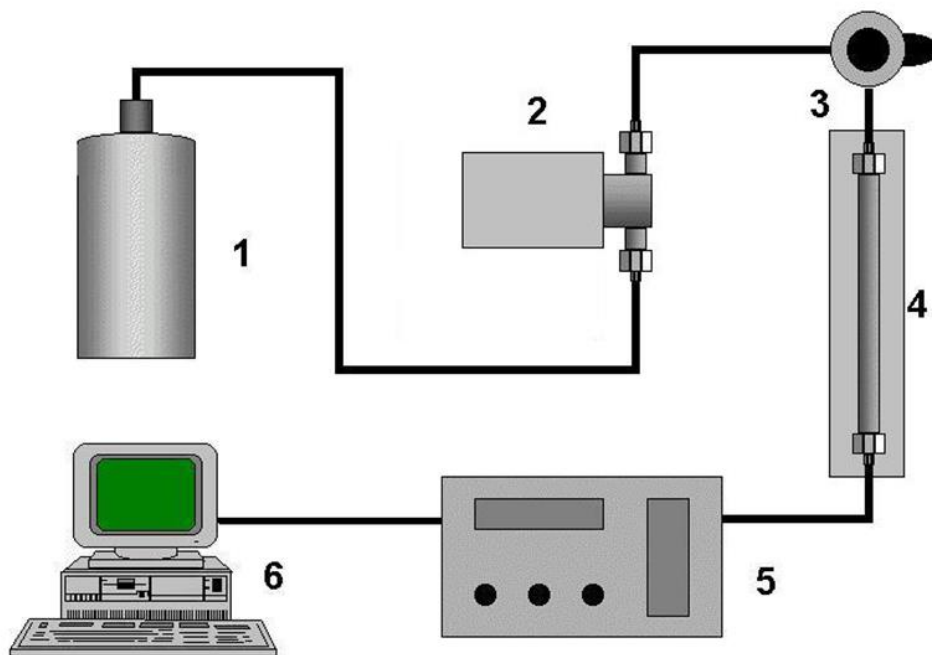
## Mechanizm rozdziału

Ze względu na różnice w oddziaływaniach między fazą stacjonarną i ruchomą oraz różnice w polarności związków zawartych w próbce, poruszają się one wewnątrz kolumny z różną szybkością i opuszczają ją w różnym czasie.



## Chromatografia cieczowa (LC)

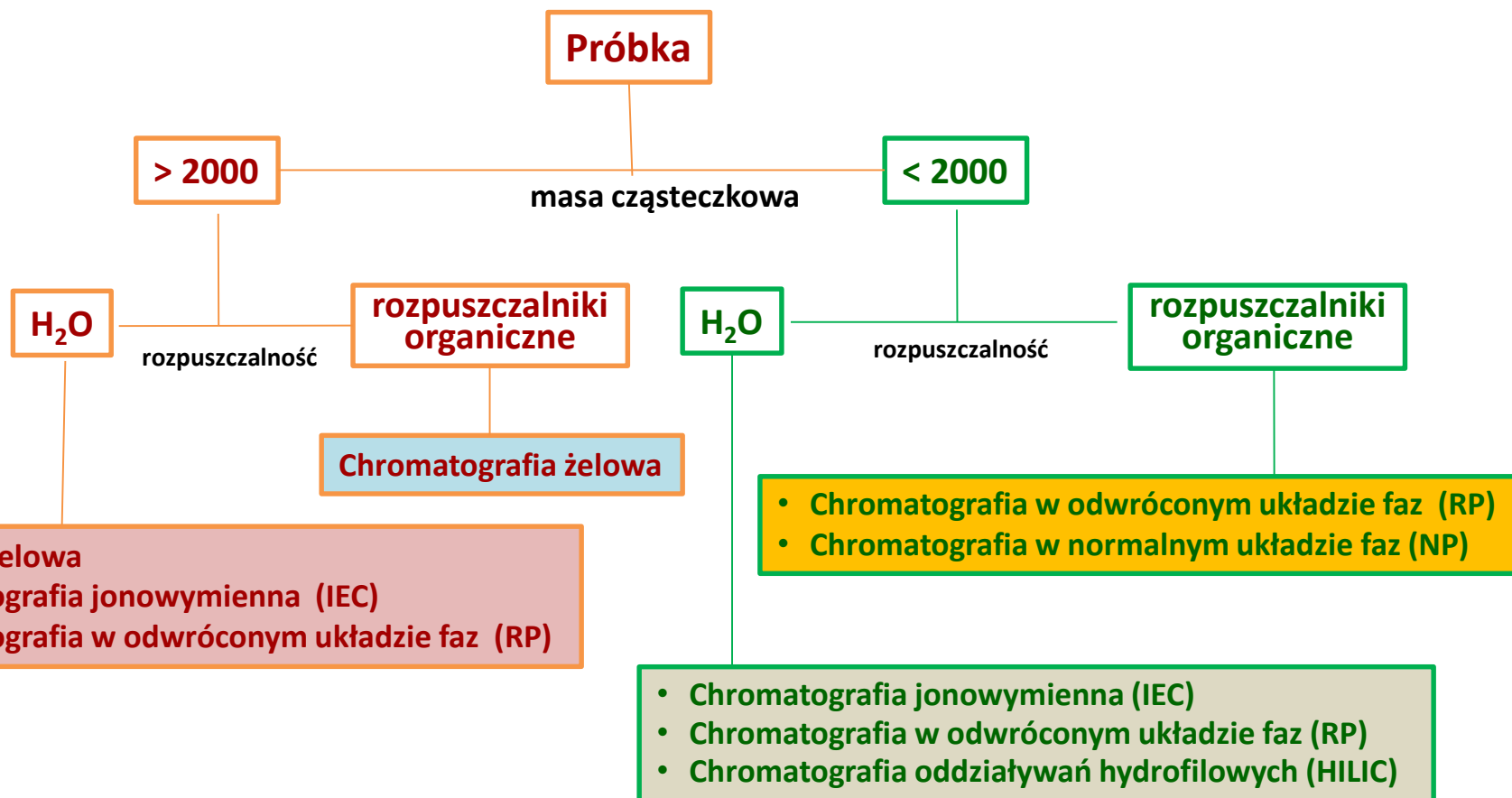
Standardowy system chromatograficzny zawiera:



- 1 – zbiornik fazy ruchomej
- 2 - pompa
- 3 – automatyczny podajnik próbek
- 4 – kolumna chromatograficzna
- 5 - detektor
- 6 – komputer



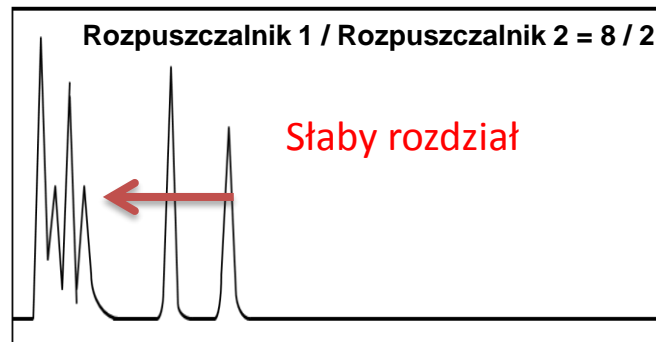
# Wybór techniki



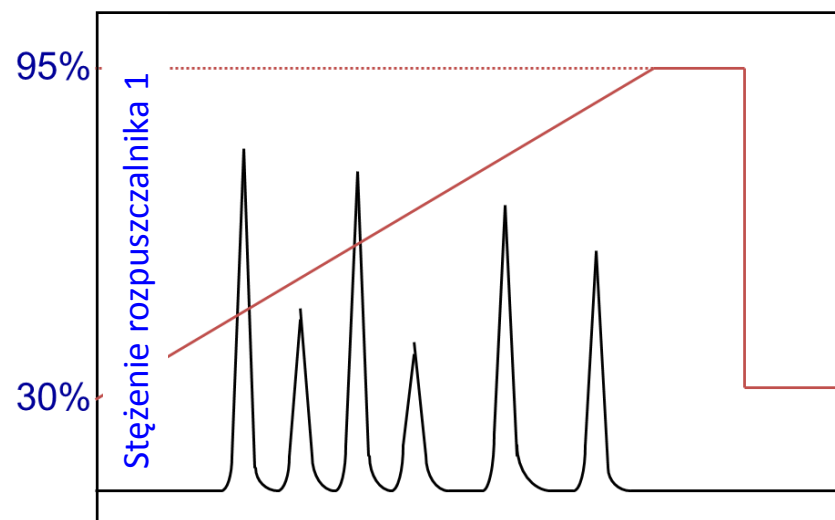
Głównym zadaniem pompy jest zapewnienie stabilnego przepływu, który zmienia się w zależności od rodzaju elucji oraz parametrów kolumny chromatograficznej.

**Pompa dostarcza fazę ruchomą do układu chromatograficznego w trybie:**

- **izokratycznym**  
(stały skład eluentu)



- **gradientowym**  
(zmienny skład eluentu)

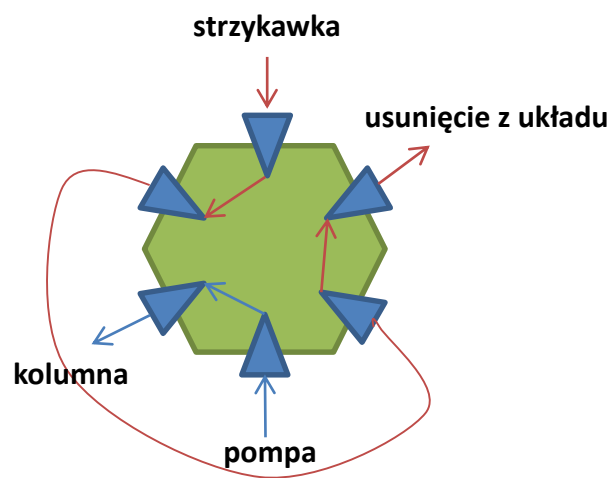


# Automatyczny podajnik próbek

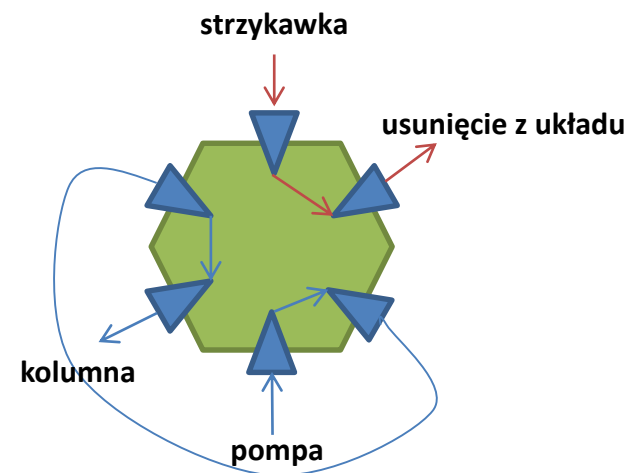
Najczęściej stosowany w LC podajnik próbek ma postać zaworu dozującego (*loop injector* lub *six-port valve injector*).

1. Próbka jest wprowadzana za pomocą mikrostrzykawki do fazy ruchomej, która wypełnia pętlę o określonej objętości.

2. Gdy próbka jest w pętli, faza ruchoma jest przepompowywana przez zawór do kolumny.



**Wprowadzanie próbki do pętli**



**Przeniesieni próbki na kolumnę**

**Dozowanie próbek powinien odbywać się:**

- z wysoką powtarzalnością
- z dużą dokładnością
- bez obecności pęcherzyków powietrza

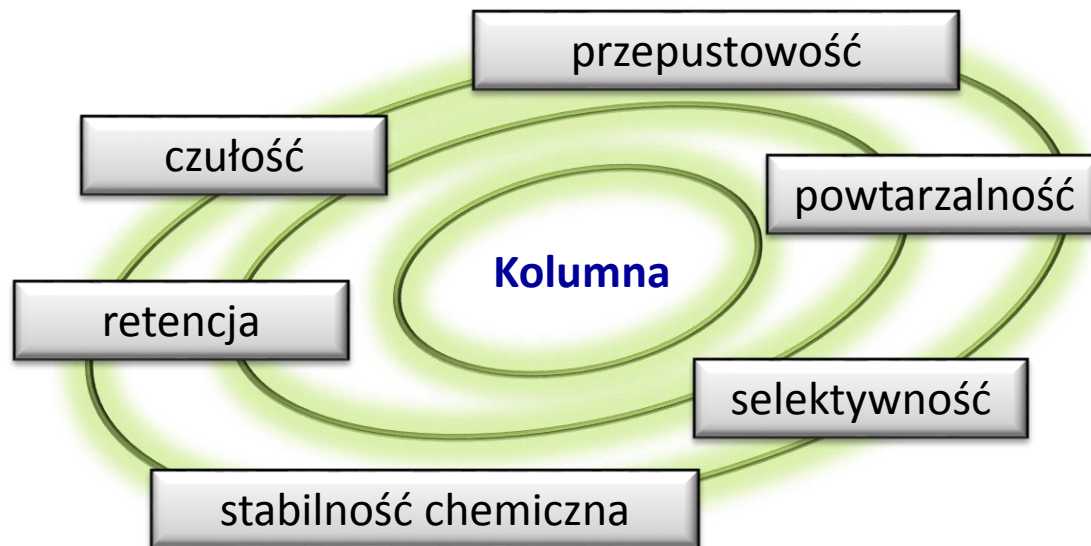
# Kolumna

Kolumny stosowane w LC charakteryzują następujące parametry:

- długości kolumny,
- średnicy kolumny,
- rozmiar ziaren fazy stacjonarnej.

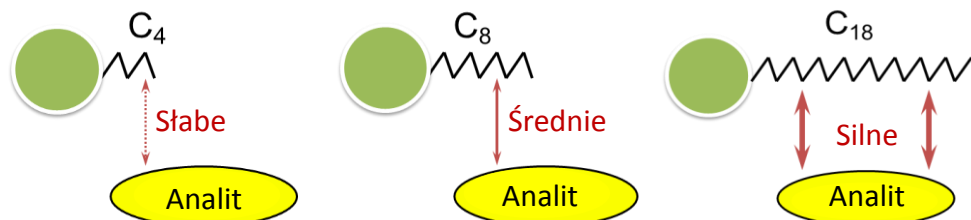
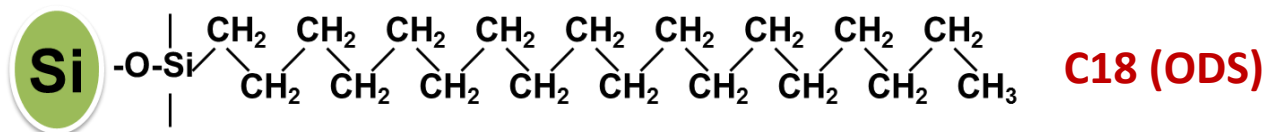
	Długość kolumny [mm]	Średnica kolumny [mm]	Średnica ziaren [ $\mu\text{m}$ ]	Optymalna prędkość przepływu [ $\mu\text{l}/\text{min}$ ]	Ciśnienie [bar]
<b>Kolumna nano-HPLC</b>	50-1000	0.05-0.1	1-3	0.3	<300
<b>Kolumna kapilarna</b>	50-1000	0.3	1-5	5	<500
<b>Mikrokolumna</b>	50-1000	0.5-1	1-5	10-50	<800
<b>Kolumna HPLC o zmniejszonym ziarnie</b>	50-250	2.1	2-5	400	<1200
<b>Kolumna HPLC</b>	30-250	4.6	2-5	1000	<400

# Kolumna



## Rodzaje wypełnień kolumn LC:

- **C18 (ODS)**
- C8 (octyl)
- C4 (butyl)
- fenyłowe
- TMS
- cyjankowe



*Oddziaływania między analitem a kolumną*

## Faza ruchoma i stacjonarna

Charakter analitu /rozdzielanych związków ma wpływ na rodzaj **fazy ruchomej** oraz **fazy stacjonarnej**.

Układ LC	Faza ruchoma	Faza stacjonarna	Rodzaj rozdzielanych związków
NP	rozpuszczalniki organiczne: dichlorometan, octan etylu	krzemionka, aminowa, cyjankowa, diol	związki organiczne nie rozpuszczalne w wodzie
RP	woda/rozpuszczalnik organiczny z/bez dodatków	C18, C8, C4, cyjankowa, aminowa	związki niepolarne, słabe kwasy i słabe zasady
HILIC	acetonitryl z wodą, dodatki substancji jonowych	polarna, czysta krzemionka	związki polarne
IEC	wodne roztwory buforowe	anionowa lub kationowa, żywice jonowymienne	związki jonowe, jony nieorganiczne

# Chromatografia w normalnym układzie faz (NP)

Rozdział chromatograficzny w normalnym układzie faz (NP) jest uzyskiwany na skutek oddziaływań związków zawartych w próbce z **polarną fazą stacjonarną** oraz **niepolarną fazą ruchomą**.

## Faza stacjonarna używana w NP:

- żel krzeminkowy:  $-\text{Si}-\text{OH}$
- cyjankowa:  $-\text{Si}-\text{CH}_2\text{CH}_2\text{CH}_2 \text{CN}$
- aminowa:  $-\text{Si}-\text{CH}_2\text{CH}_2\text{CH}_2 \text{NH}_2$
- diolowa:  $-\text{Si}-\text{CH}_2\text{CH}_2\text{CH}_2\text{OCH}(\text{OH})-\text{CH}_2 \text{OH}$

## Faza ruchoma używana w NP:

- dichlorometan
- octan etylu
- inne rozpuszczalniki niemieszające się z wodą



# Chromatografia w odwróconym układzie faz (RP)

## RP jest najczęściej stosowaną spośród metod HPLC

Do rozdziału chromatograficznego w odwróconym układzie faz (RP) dochodzi na skutek oddziaływań związków zawartych w próbce z **niepolarną fazą stacjonarną** oraz **polarną fazą ruchomą**.

### Faza stacjonarna używana w RP:

(długołańcuchowe węglowodory kowalencyjnie przyłączone do powierzchni krzemionki)

- C18
- C8
- C4
- cyjankowa
- aminowa

### Faza ruchoma używana w RP:

**woda lub roztwór buforowy i rozpuszczalniki organiczne**, wśród których najczęściej stosowane są:

- metanol
- acetonitryl

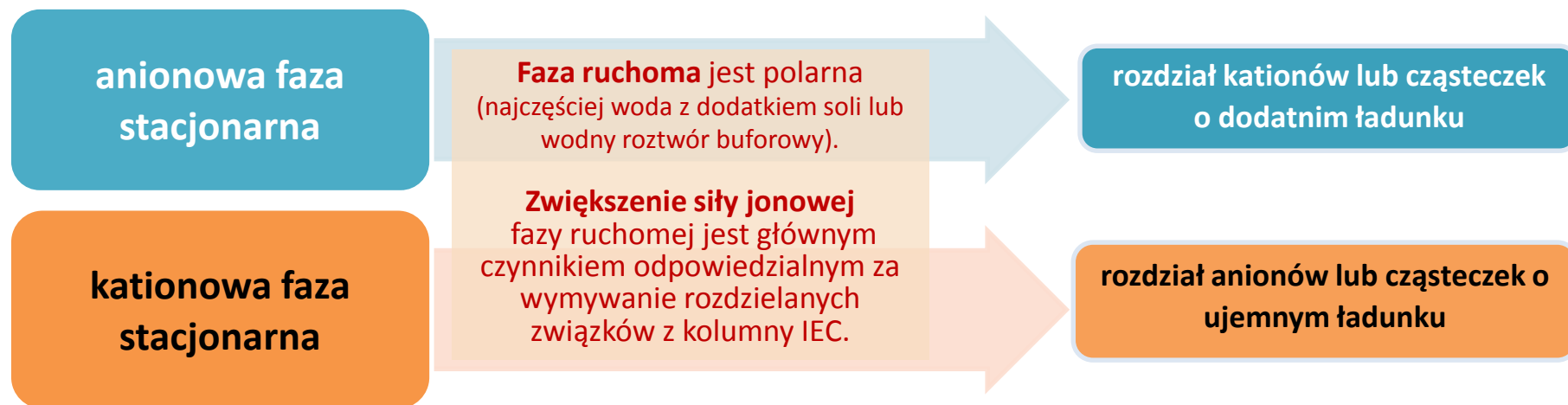
**NP-LC oraz RP-LC są wykorzystywane do różnych celów w analizie próbek biologicznych:**

- NP-LC stosuje się do rozdzielania poszczególnych klas lipidów dzięki **obecności w obrębie ich struktury „głów” polarnych**,
- RP-LC stosuje się do rozdziału kwasów tłuszczowych dzięki **różnicy w hydrofobowości ich łańcuchów**,
- RP jest również stosowana do profilowania metabolitów w badaniach metabolomicznych



# Chromatografia jonowymienna (IEC)

Chromatografia jonowymienna zapewnia rozdział analitów dzięki interakcji **jonów (lub substancji wykazujących zdolność do przechodzenia w jony)** z **fazą stacjonarną zawierającą jonowe grupy** funkcyjne o przeciwnym ładunku do analitów.



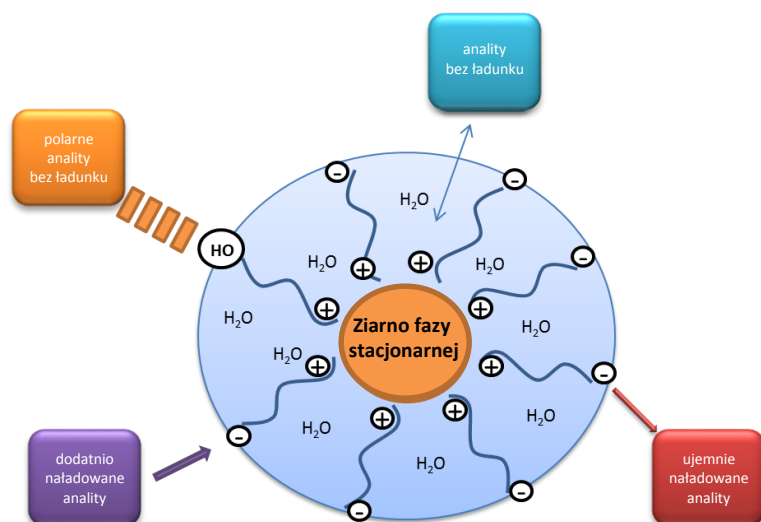
IEC jest stosowana do rozdziału zarówno **drobno-** jak i **wielkocząsteczkowych** związków o znaczeniu biologicznym (takich jak aminokwasy, kwasy karboksylowe, aminy).

Z uwagi na zjawisko supresji jonów wspomagane dużą mocą jonową fazy ruchomej IEC, bezpośrednie połączenie tego typu chromatografii cieczerwowej ze spektrometrią mas jest trudne do realizacji.

# Chromatografia oddziaływań hydrofilowych (HILIC)

**HILIC** jest stosowana do rozdzielania związków polarnych

Mechanizm rozdzielania w chromatografii **HILIC** jest oparty o zjawisko ekstrakcji w układzie rozpuszczalnik - rozpuszczalnik na granicy faz utworzonej pomiędzy warstwą wody na powierzchni polarnego wypełnienia a rozpuszczalnikiem organicznym fazy ruchomej.



Schemat interakcji między różnymi typami polarnych analitów i fazą stacjonarną w chromatografii HILIC

## Faza stacjonarna

- **hydrofilowa**
- kolumna wypełniona zazwyczaj krzemionką lub krzemionką ze związanymi na jej powierzchni grupami aminowymi bądź amidowymi

## Faza ruchoma

- **rozpuszczalniki typowe dla chromatografii RP**, zazwyczaj acetonitryl z niewielką ilością wody
- sole octanu amonu lub mrówczanu amonu dodawane są w celu zwiększenia polarności oraz siły jonowej fazy ruchomej

**HILIC łatwo jest połączyć z analizami MS**

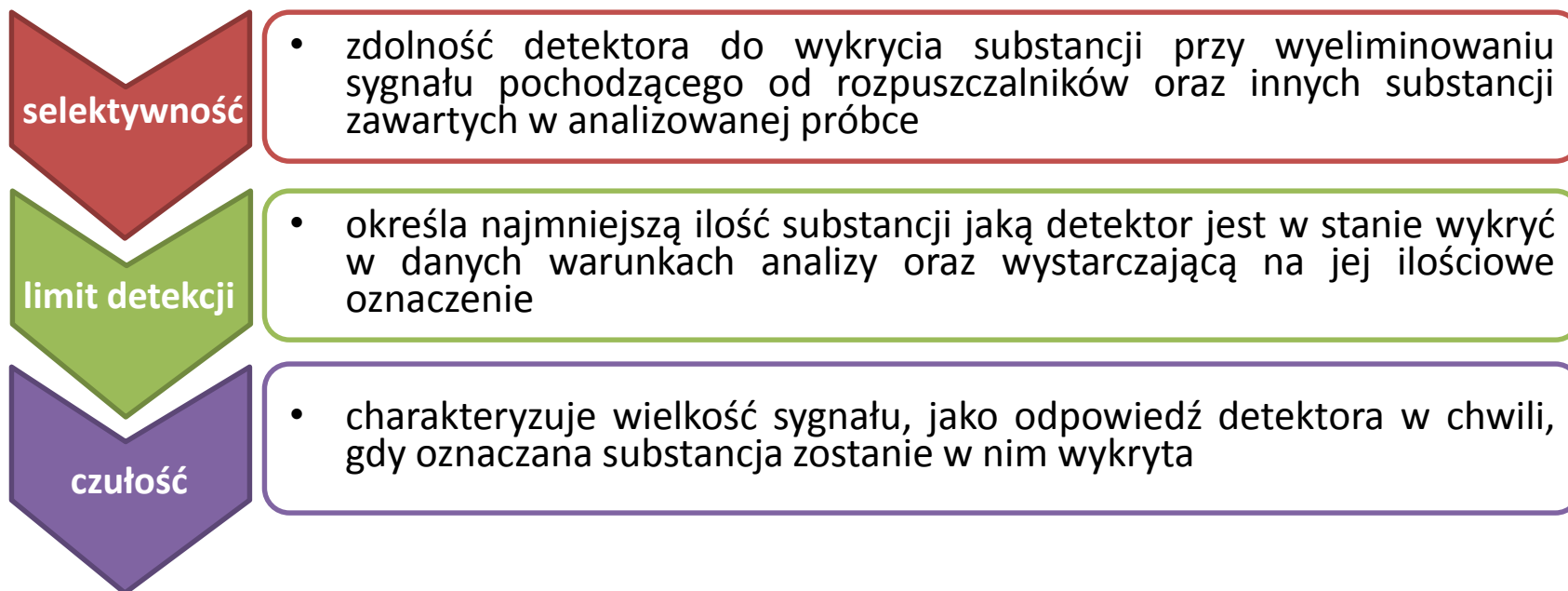
**Rozpuszczalniki organiczne obecne w fazie ruchomej dodatkowo zwiększają czułość analiz MS dzięki ograniczeniu zjawiska supresji jonów.**

# Detektory

## Detektory stosowane w chromatografii cieczowej:

- UV,
- fluorescencyjny,
- elektrochemiczny,
- konduktometryczny,
- refraktometryczny,
- **detektor mas**

PARAMETRY DETEKTORÓW



**Spektrometry mass - najlepsze detektory do analiz jakościowych i ilościowych**

# Wprowadzenie do GC-MS

**Chromatografia gazowa(GC)** - technika separacyjna służąca do rozdzielania oraz analizy składu bardzo złożonych próbek w zależności od ich **temperatury wrzenia/ciśnienia pary** oraz **polarności**.

w GC:

- **fazą ruchomą** jest gaz (Ar, He, N<sub>2</sub> lub H<sub>2</sub>)
- **fazą stacjonarną**:
  - **adsorbent** – Adsorpcyjna Chromatografia Gazowa(GSC) lub
  - **ciekła faza stacjonarna** – Podziałowa Chromatografia Gazowa (GLC)



## Rozdzielenie składników w GC zależy od

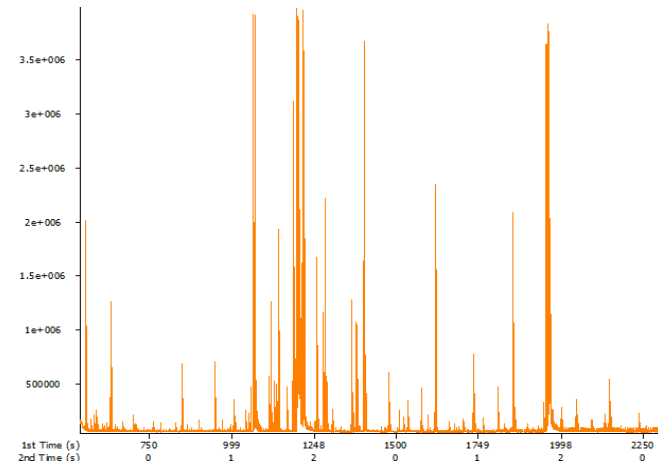
przeniesienia substancji (w postaci pary) za pomocą gazu nośnego (faza ruchoma) przez kolumnę.

Szybkość i stopień rozdzielania związków w GC zależy od :

- **powinowactwa chemicznego analitu do fazy stacjonarnej**

oraz

- **prężności par analitu** – która jest regulowana przez temperaturę kolumny



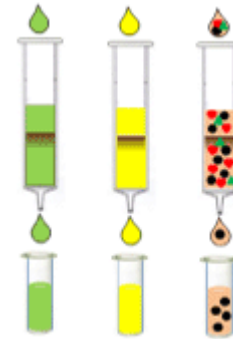
# Etapy przygotowania próbki do analizy chromatograficznej

Pobieranie próbek

gazowych  
ciekłych  
stałych

Obróbka wstępna

suszenie  
sączenia  
homogenizacja  
...

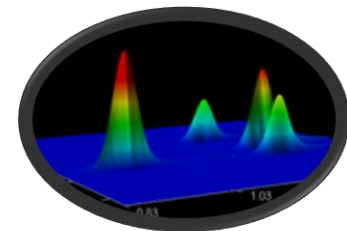
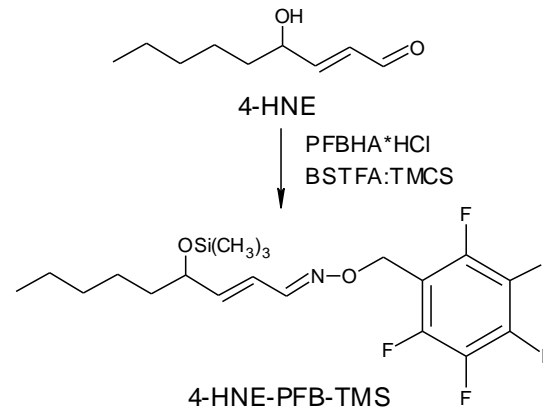


Przygotowanie próbki

SPE  
SPME  
...

Derywatywacja

sililowanie  
alkilowanie  
...



Analiza  
GC, GCxGC

FID  
MS  
...

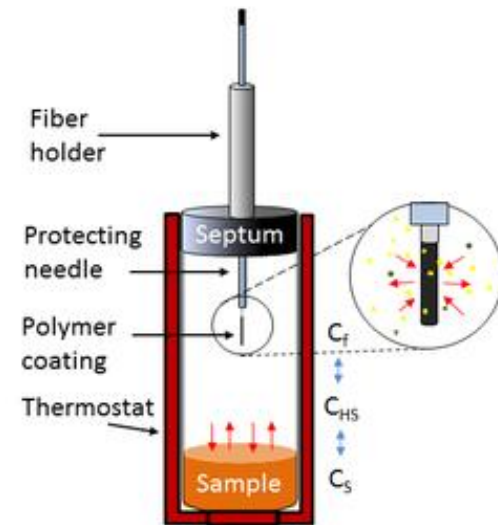
## Pobieranie próbek

Próbka analizowana w **GC** może być próbką:

- gazową
- ciekłą
- substancją zaadsorbowaną na powierzchni sorbentu (np. mikroekstrakcja do fazy stałej, SPME)

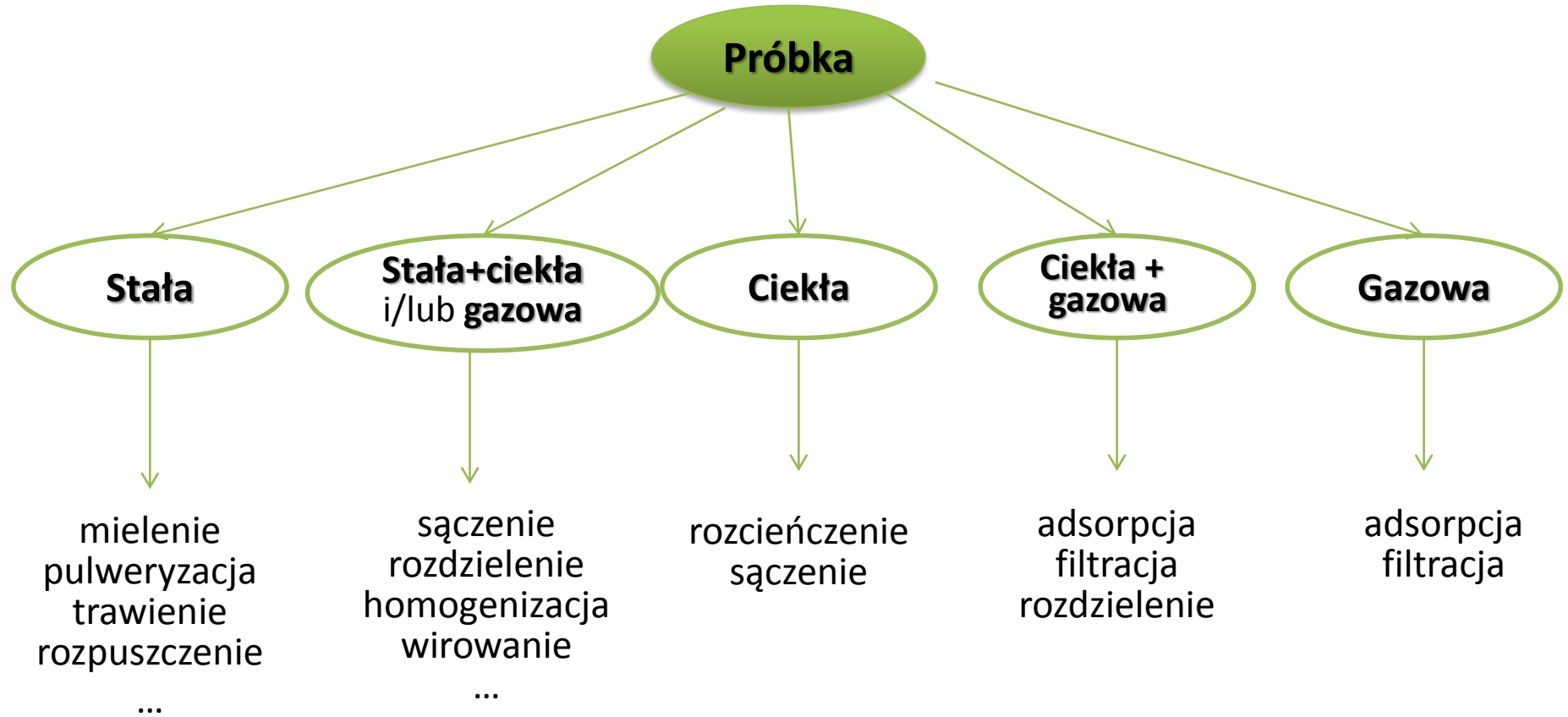
**Stabilizację składu i właściwości próbek** uzyskuje się często poprzez dodanie odpowiednich środków, które:

- hamują aktywność biologiczną organizmów występujących w próbkach
- eliminują zjawisko adsorpcji składników próbek na ścianach naczyń
- eliminują ulatnianie się składników próbek, rozkład termiczny, reakcje chemiczne, i in.

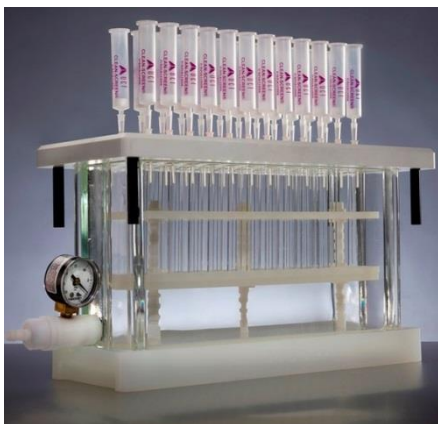


**Schemat SPME**

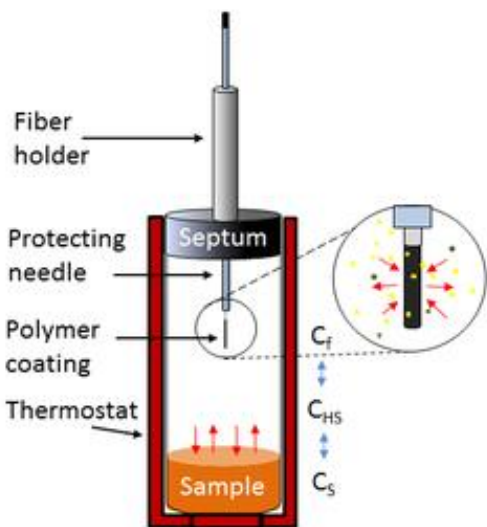
# Obróbka wstępna próbki



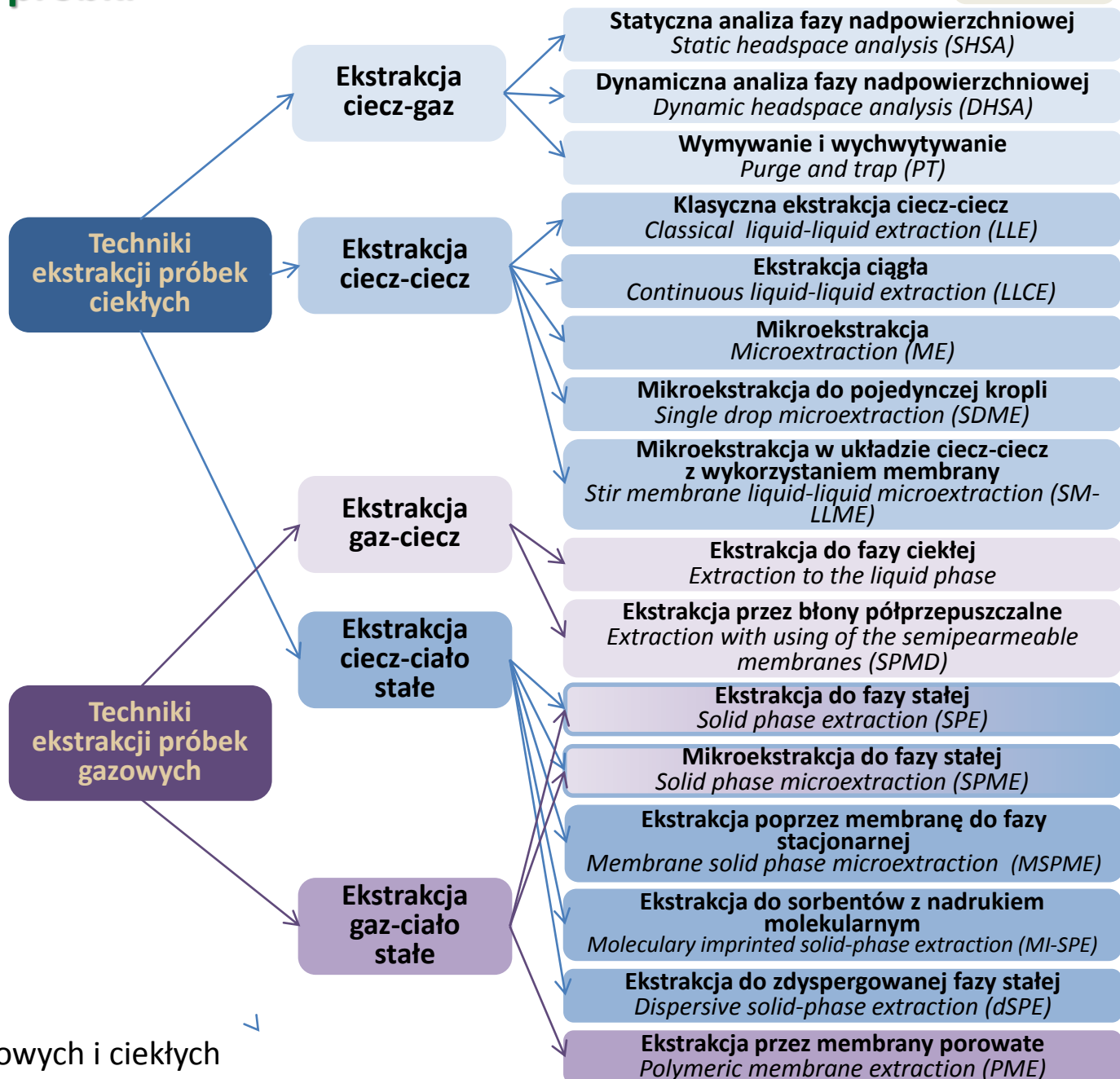
# Przygotowanie próbki



Zestaw do SPE



Schemat SPME





# Przygotowanie próbki



System do ekstrakcji mikrofalowej



Termodesorber

## Metody ekstrakcji próbek stałych

Techniki ekstrakcji próbek stałych

Ekstrakcja ciało stałe - gaz

Statyczna analiza fazy nadpowierzchniowej  
*Static headspace analysis (SHSA)*

Dynamiczna analiza fazy nadpowierzchniowej  
*Dynamic headspace analysis (DHSA)*

Termodesorpcja  
*Thermal desorption (TD)*

Ekstrakcja rozpuszczalnikiem przez wytrząsanie  
*„Shake extraction”*

Ekstrakcja strumieniem rozpuszczalnika  
*Extraction in a solvent stream*

Ekstrakcja w aparacie Soxhleta  
*Soxhlet extraction*

Homogenizacja próbki z rozpuszczalnikiem  
*Sample homogenization*

Ekstrakcja wieloetapowa (sekwencyjna)  
*Multi-stage (sequential) extraction*

Zmydlanie (saponifikacja)  
*Saponification*

Ultradźwiękami (sonikacja)  
*Ultrasound assisted extraction (USE)*

Metody klasyczne - wspomagane

Promieniowaniem mikrofalowym  
*Microwave assisted extraction (MAE)*

Przyspieszona ekstrakcja za pomocą rozpuszczalnika  
*Accelerated solvent extraction (ASE)*

Ekstrakcja rozpuszczalnikiem pod zwiększonym ciśnieniem  
*High-pressure extraction (HPE)*

Ekstrakcja rozpuszczalnikiem wspomagana promieniowaniem mikrofalowym  
*Microwave-assisted solvent extraction (MASE)*

Ekstrakcja płynem w stanie nadkrytycznym  
*Supercritical fluid extraction (SFE)*

Metody klasyczne

Nowoczesne metody ekstrakcji

# Derywatywacja

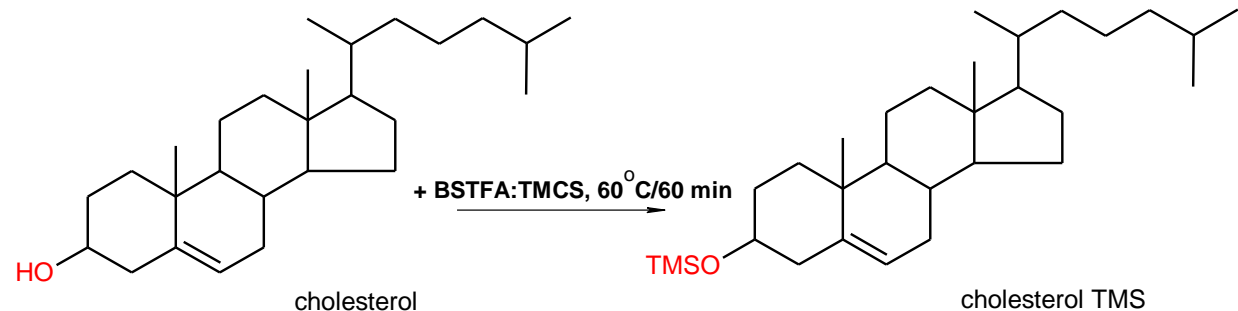
Bezpośrednia analiza mieszaniny związków w GC jest trudna ze względu na:

- zbyt niską lotność analizowanych związków
- zbyt niską stabilność termiczną analizowanych związków
- oddziaływania pomiędzy oznaczanymi związkami
- oddziaływania pomiędzy substancjami a fazą stacjonarną kolumny chromatograficznej
- zbyt niską czułość i specyficzność oznaczeń

Dlatego też ważnym etapem przygotowania próbki do analizy GC jest przekształcenie analitów w ich bardziej lotne i termicznie stabilne pochodne (derywatywacja)

Derywatywacja analitów odbywa się na drodze:

- silylowania
- alkilowania
- acylowania



Reakcja silylowania

# Derywatywacja

Rodzaj derywatywacji	Grupa funkcyjna - typ związku	Pochodna	Odczynnik
<b>Sililowanie</b>	<p>-OH -alkohole, fenole</p> <p>-CO -ketony, steroidy</p> <p>-COOH -aminokwasy, kwasy tłuszczowe, steroidy</p> <p>-(CH<sub>2</sub>OH)<sub>n</sub> -węglowodany</p> <p>-NH, -NH<sub>2</sub> -aminy, mocznik</p> <p>-CONH, -CONH<sub>2</sub> -imidy, białka</p>	<p>Etery trimetylosililowe</p> <p>Amidy trimetylosililowe</p>	<p><b>Bis (trimetylosililo)trifluoroacetamid (BSTFA)</b></p> <p>N- metylo -N-(t-butylodimetylosililo) trifluoroacetamid (MTBSTFA)</p> <p>N- metylotrimetylosililotrifluoroacetamid (MSTFA)</p> <p>Trimetylosililoimidazol (TMSI)</p>
<b>Alkilowanie</b>	<p>-OH -alkohole, fenole</p> <p>-CO -aldehydy</p> <p>-COOH -aminokwasy, kwasy tłuszczowe</p> <p>-NH, -NH<sub>2</sub> -aminy</p> <p>-CONH -amidy</p> <p>-SH -merkaptany</p>	<p>Estry metylowe (DMF)</p> <p>Trifluorooctany (TFAA)</p> <p>Estry metylowe (BF<sub>3</sub>-MeOH)</p> <p>Etery pentafluorobenzylu (PFBBr)</p> <p>Metyloamidy (TMAH)</p> <p>Estry metylowe (DMF)</p>	<p>Bromek benzylu</p> <p>Trifluorek boru (BF<sub>3</sub>) w metanolu lub butanolu</p> <p>Dimetyloformamid (DMF)</p> <p>Chlorowodorek pentafluorobenzylhydroksylaminy (PFBHA)</p> <p>Wodorotlenek tetrabutylamoniowy (TBH)</p> <p>Bezwodnik trifluorooctowy (TFAA)</p>
<b>Acylowanie</b>	<p>-OH -alkohole, fenole</p> <p>-(CH<sub>2</sub>OH)<sub>n</sub> -węglowodany</p> <p>-NH, -NH<sub>2</sub> -aminy</p> <p>-CONH -amidy</p> <p>-SH -merkaptany</p>	<p>Pentafluoropropioniany (PFPA)</p> <p>Trifluoroacetamidy (TFAI)</p> <p>Trifluoroacetamidy (MBTFA)</p> <p>Trifluoroacetamidy (TFAA)</p> <p>Etery trimetylosililowe (MBTFA)</p>	<p>Bezwodnik kwasu heptafluoromastowego (HFBA)</p> <p>N-Metylo-bis(trifluoroacetamid) (MBTFA)</p> <p>Chlorek pentafluorobenzoilu (PFBCI)</p> <p>Pentafluoropropanol (PFPOH)</p> <p>Bezwodnik kwasu trifluorooctowego (TFAA)</p>

# Chromatograf gazowy

## Budowa chromatografu gazowego:

### ▪ dozownik

próbka jest wstrzykiwana do ogrzanego portu dozownika, gdzie jej lotne składniki ulegają odparowaniu i w strumieniu gazu nośnego (obojętny, czystość  $\geq 99,9995\%$  (Ar, He, N<sub>2</sub> lub H<sub>2</sub>)) przenoszone są na kolumnę chromatograficzną

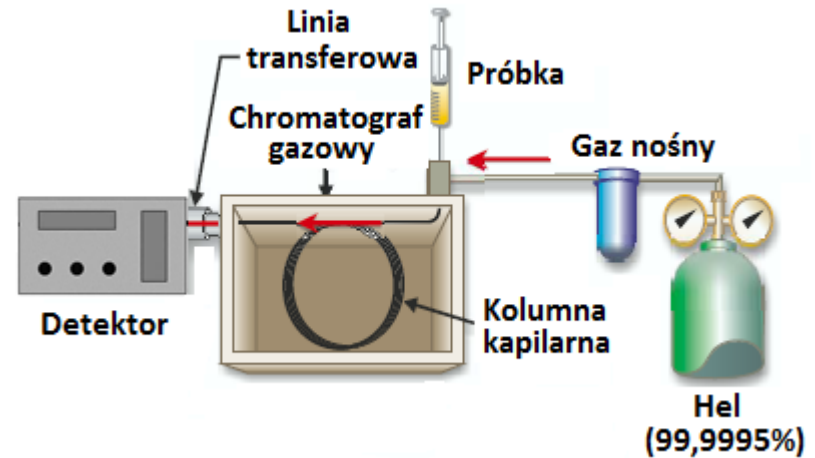
### ▪ kolumna GC

rozdzielenie składników próbki

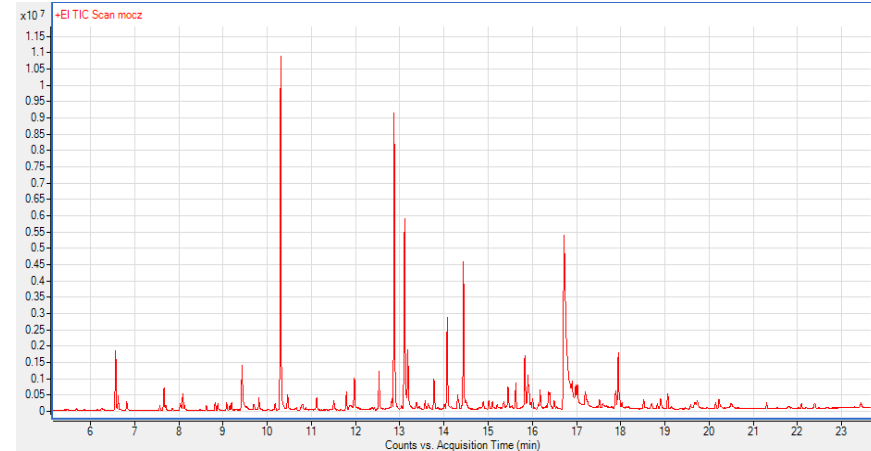
### ▪ detektor

reaguje na pewne właściwości fizykochemiczne analitu i generuje sygnał elektroniczny który jest proporcjonalny do jego ilości

**chromatogram** - wynik rozdzielenia chromatograficznego



Schemat chromatografu gazowego  
(modyfikacja z [de.leco-europe.com](http://de.leco-europe.com))



Chromatogram GC składników osocza krwi

## Dozownik GC

### Dozowanie próbek:

1. Próbką jest wstrzykiwana do ogrzanego portu dozownika
2. W porcie dozownika lotne składniki próbki ulegają odparowaniu
3. W strumieniu gazu nośnego lotne składniki próbki przenoszone są na kolumnę GC

Najpopularniejszym dozownikiem stosowanym w GC jest dozownik typu **Split/Splitless (SSI)**

### *Dozowanie z dzieleniem strumienia (typu Split)*

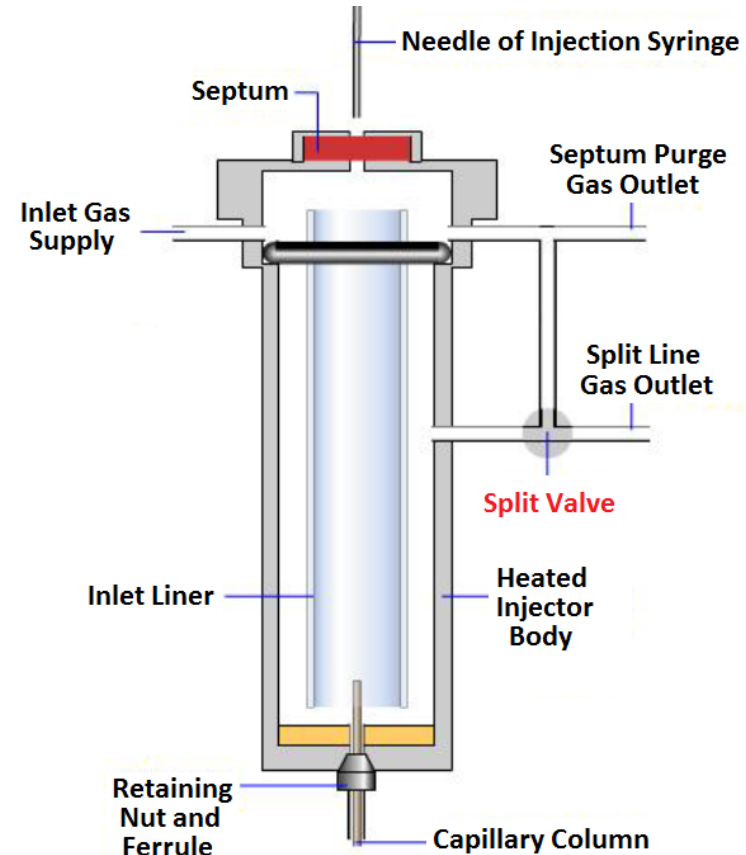
tylko niewielka część dozowanej próbki wprowadzana jest na kolumnę (zwykle 1/20 do 1/500)

- stosowane głównie w analizie próbek lotnych, w których oznaczane składniki występują w wysokich stężeniach (non-trace analysis)

### *Dozowanie bez dzielenia strumienia (typu Splitless)*

całość wprowadzonej próbki po odparowaniu jest kierowana do kolumny

- stosowane głównie do analizy składników na poziomie śladów i ultra-śladów



## Kolumny GC

**Kolumny pakowane** (analityczne, mikropakowane) wypełnione fazą stacjonarną (adsorbentem):

- adsorbenty węglowe
- krzemionka
- tlenek glinu
- sita molekularne
- porowate polimery syntetyczne



Adsorbenty są obecnie mniej popularnymi fazami stacjonarnymi ze względu na:

- niższą powtarzalność wyników
- dłuższe czasy retencji
- pojawianie się „pików ogonujących” o znacznie niższej efektywności rozdzielania

**Kolumny kapilarne** (kapilarne, mikrokapilarne)

▪ o przekroju otwartym, wypełnione ciekłą fazą stacjonarną:

- silikony
- skwalen
- glikol polietylenowy

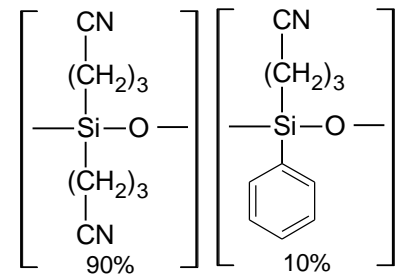
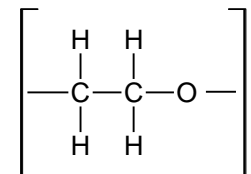
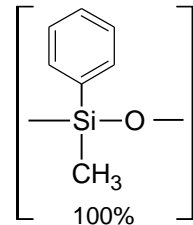
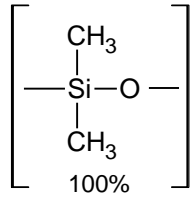
▪ ciekła faza stacjonarna powinna być:

- obojętna chemicznie
- zdolna do rozpuszczania rozdzielanych składników
- wysoce selektywna względem składników mieszaniny
- o niskiej lotności
- o wysokiej stabilności termicznej w warunkach pracy kolumny
- szczególnie przydatna do rozdzielania składników lotnych z wysoką efektywnością rozdzielczą

▪ powszechnie stosowane



# Ciekłe fazy stacjonarne



Wzrost polarności

Dimetylpolisiloksan

5% fenylo-Dimetylpolisiloksan

50% fenylo-Dimetylpolisiloksan

7% cyjanopropylo, 7% fenylo-Dimetylpolisiloksan

25% cyjanopropylo, 25% fenylo-Dimetylpolisiloksan

Glikol Polietylenowy (PEG)

70% cyjanopropylo-Metylpolisiloksan

≥90% cyjanopropylo-Polisiloksan

Wzrost selektywności



## Rozdział chromatograficzny

Szybkość i stopień rozdzielania związków na kolumnie GC jest funkcją ich podziału pomiędzy dwie fazy (**ruchomą i stacjonarną**) pozostające w równowadze

Rozdzielenie składników jest zatem wynikiem różnego tempa migracji spowodowanej różnymi wartościami **współczynnika podziału** ( $K_S$ )

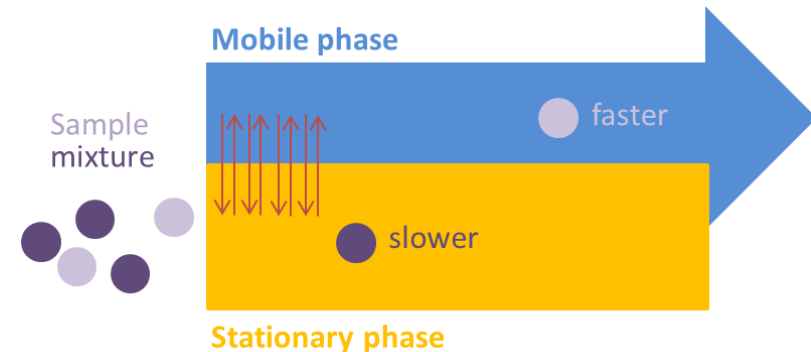
**Współczynnik podziału** można wyrazić **równaniem Nernsta**:

$$K_S = C_L / C_G$$

$C_L$  - stężenie analitu w fazie ciekłej (stacjonarnej)

$C_G$  - stężenie analitu w fazie gazowej (ruchomej)

Im wyższe powinowactwo do fazy stacjonarnej, tym wyższa wartość  $K_S$  i dłuższy czas retencji ( $t_R$ )



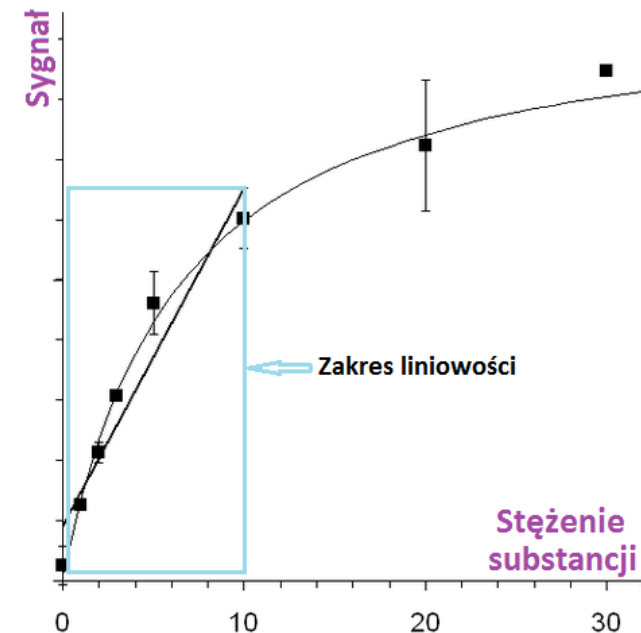
$$K_{\text{slower}} > K_{\text{faster}}$$



# Detektory GC

## Charakterystyka idealnego detektora GC

- wysoka czułość
- wysoka stabilność i powtarzalność wskazań
- szeroki liniowy zakres odpowiedzi (kilka rzędów wielkości)
- szeroki zakres temperaturowy
- krótki czas reakcji, niezależny od natężenia przepływu
- wysoka niezawodność i łatwość użytkowania
- podobieństwo odpowiedzi wobec wszystkich oznaczanych substancji
- niedestrukcyjny

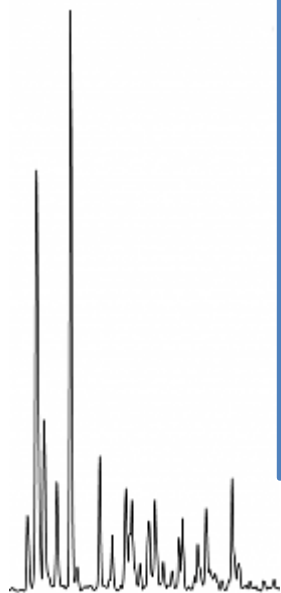
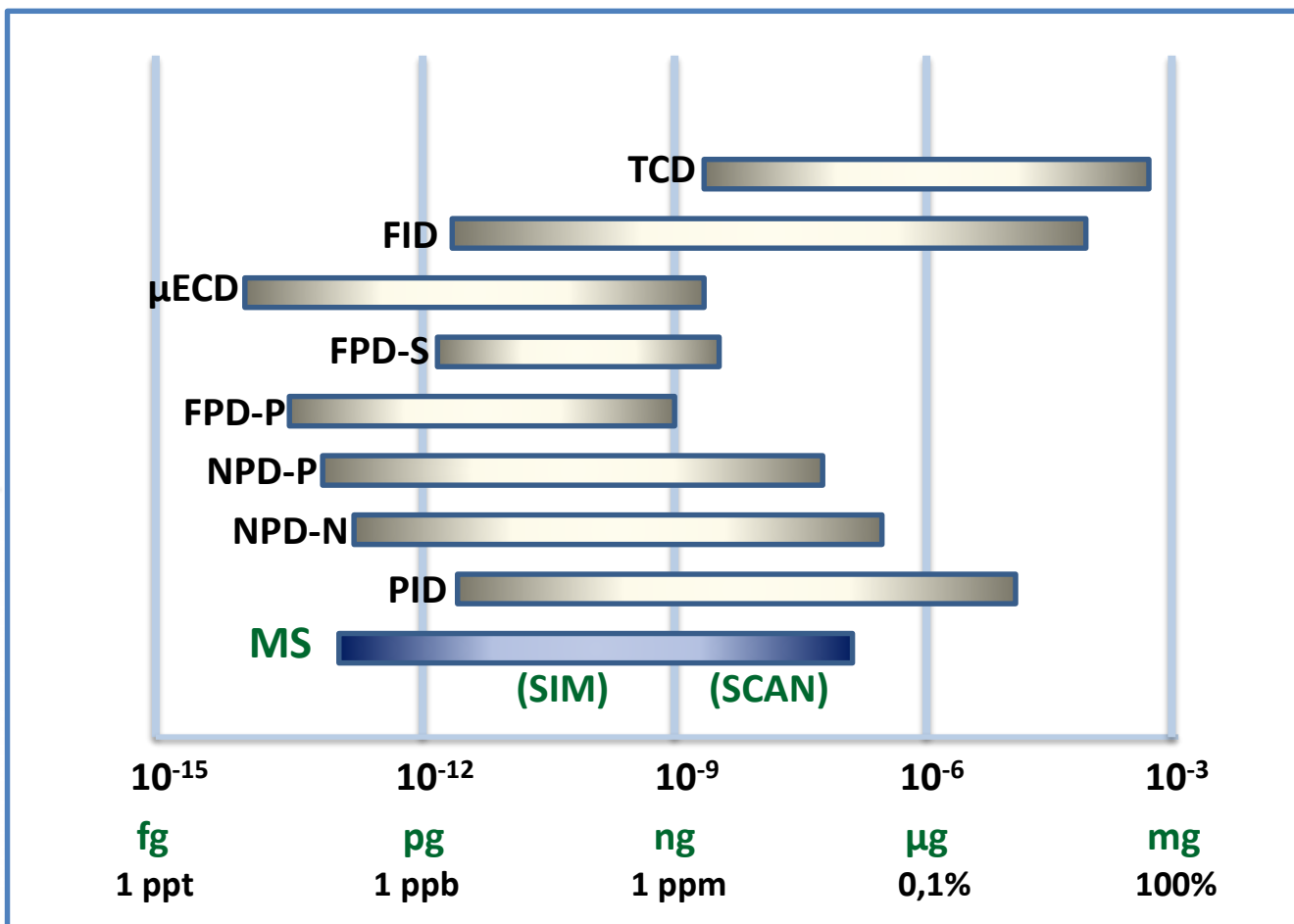


# Detektory GC

Rodzaj detektora	Rodzaj odpowiedzi	LOD [g/sec]	Zakres liniowości	Właściwości
Cieplno-przewodnościowy (TCD)	uniwersalny	$10^{-5}$ - $10^{-6}$	$10^3$ - $10^4$	pomiar zmian w przewodnictwie cieplnym
Płomieniowo-jonizacyjny (FID)	uniwersalny	$10^{-12}$	$10^6$ - $10^7$	pomiar prądu jonowego pochodzącego z pirolizy
Wychwytu elektronów (ECD)	selektywny	$10^{-14}$	$10^2$ - $10^3$	do wykrywania związków zawierających atomy o wysokim powinowactwie elektronowym
Płomieniowo-fotometryczny (FPD)	selektywny	$10^{-13}$	$10^2$	do wykrywania związków zawierających S, P
Azotowo-fosforowy (NPD)	selektywny	$10^{-8}$ - $10^{-14}$	$10^5$ - $10^7$	do wykrywania związków zawierających N, P
Fotojonizacyjny (PID)	selektywny	$10^{-8}$ - $10^{-12}$	$10^5$	selektywny dzięki zastosowaniu lamp emitujących promieniowanie UV o różnej długości fali
Spektrometria w podczerwieni z transformacją Fouriera (FTIR)	selektywny	$10^{-10}$	zależy od grup funkcyjnych	do wykrywania związków polarnych
Spektrometr mas (MS)	uniwersalny	$10^{-12}$	zależy od rodzaju analizatora MS, trybu pracy i oznaczanych związków	<b>niedestrukcyjny</b> , jedno z najdokładniejszych i najskuteczniejszych narzędzi w analizie związków organicznych, <b>najważniejszy detektor w GC</b>

# Detektory GC

Porównanie czułości i zakresu dynamiki detektorów w GC

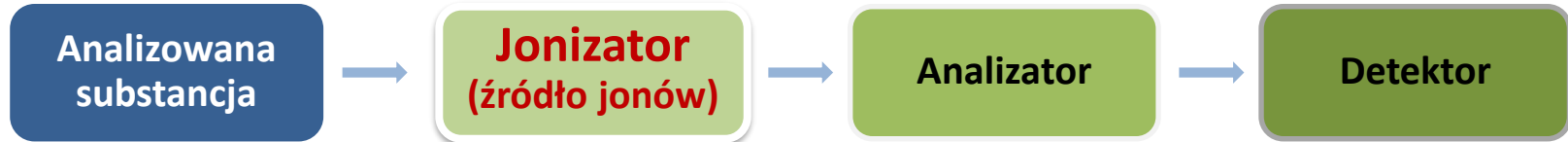


# Chromatografia gazowa sprzężona ze spektrometrią mas (GC-MS)

## GC-MS

- jest wszechstronnym narzędziem do:
  - rozdzielania
  - identyfikacji
  - ilościowego oznaczania nieznanymi substancji
- jest najskuteczniejszą techniką w analizie lotnych związków organicznych w złożonych matrycach, w szerokim zakresie stężeń (od ppb do ppm)
- charakteryzuje się wysoką selektywnością i czułością, zapewniając **szeroki zakres zastosowań**:
  - medycyna, farmacja
  - analiza materiału biologicznego
  - aplikacje sądowe i kryminalistyczne
  - monitoring środowiska
  - analiza środków bojowych
  - analiza żywności/smaków/zapachów
  - aplikacje chemiczne/przemysłowe
  - badania geochemiczne
  - analizy petrochemiczne

# Aparatura do GC-MS



## Metody jonizacji

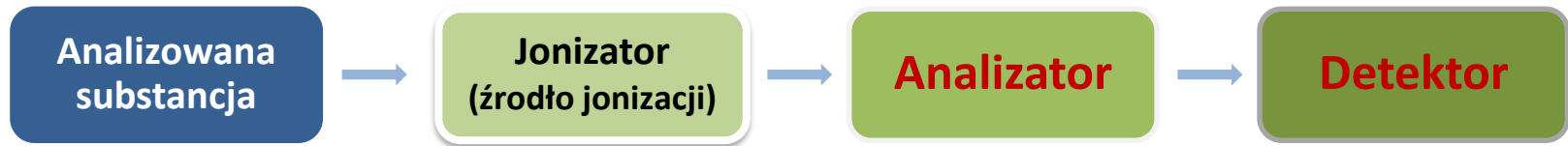
**EI – jonizacja elektronowa-** najczęściej stosowana metoda jonizacji w GC-MS, „twarda jonizacja”

**CI – jonizacja chemiczna-** stosunkowo łagodna (tzw. „łagodna, miękka jonizacja”)

**Pozostałe metody jonizacji stosowane w GC-MS:**

- fotojonizacja (PI)
- jonizacja polem (FI)
- desorpcja polem (FD)
- desorpcja laserowa (LD)
- bombardowanie szybkimi atomami (FAB)
- desorpcja plazmą (PD)
- spektrometria jonów wtórnych (SIMS)
- desorpcja laserowa wspomagana matrycą (MALDI)

# Aparatura do GC-MS



**Analizator** - rozdział jonów ze względu na stosunek ich masy do ładunku ( $m/z$ )

- **Q** - kwadrupol (pojedynczy, potrójny)
- **IT** - pułapka jonowa (liniowa, sferyczna)
- **TOF** - detektor czasu przelotu
- **hybrydowe analizatory MS:**
  - QqQ
  - Q/IT
  - Q/TOF
  - Orbitrap
- **pozostałe analizatory stosowane w GC-MS:**
  - sektor magnetyczny (B)
  - sektor elektryczny (E)
  - sektor magnetyczno-elektryczny
  - jonowy rezonans cyklotronowy (ICR)
  - analizator cyklotronowego rezonansu jonów z transformacją Fouriera (FT-ICR MS)

**Detector** – rejestracja danych

- puszka Faradaya
- powielacz elektronowy
- detektor mikrokanalikowy
- fotopowielacz

## GC

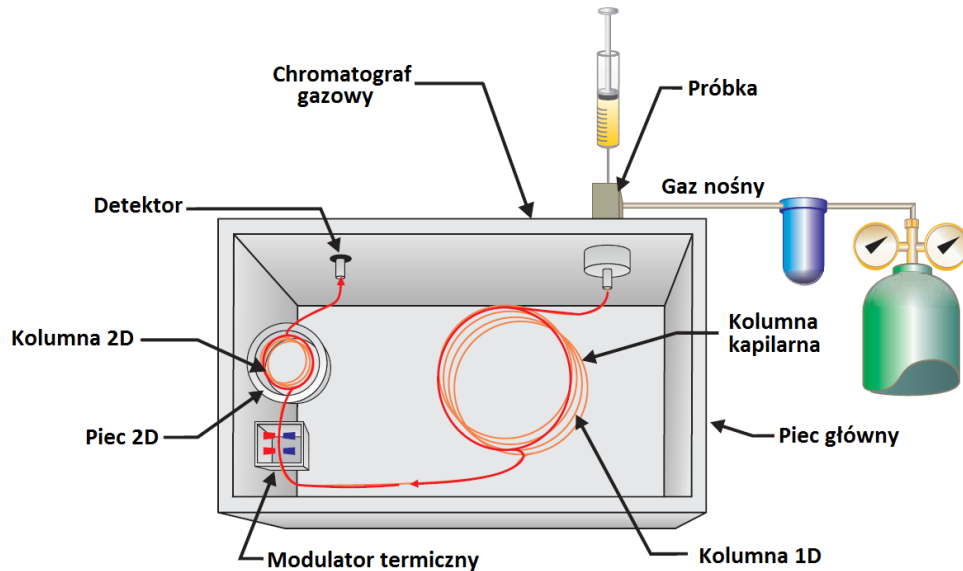
## Zalety

- Szybka analiza
- Wysoka wydajność - prowadząca do wysokiej rozdzielczości
- Czułe detektory (ppb)
- Wysoka dokładność (<1% RSD)
- Technika niedestrukcyjna (gdy detektor MS)
- Mała próbka (<1 mL)
- Wytrzymała i niezawodna technika
- Dobrze ugruntowana dzięki obszernej literaturze i aplikacjom

## Wady

- Ograniczona do próbek lotnych lub po derywatywacji
- Nie ma zastosowania w analizie próbek ulegających degradacji w podwyższonych temperaturach (termicznie nietrwałych) – wymagane jest przeprowadzenie derywatywacji
- Nie ma zastosowania w chromatografii preparatywnej
- Do analizy strukturalnej analitów wymagany jest detektor MS
- Detektory destrukcyjne (wyj. MS)
- Ograniczona pojemność pików w analizie bardzo złożonych próbek – **wymagany rozdział GCxGC**

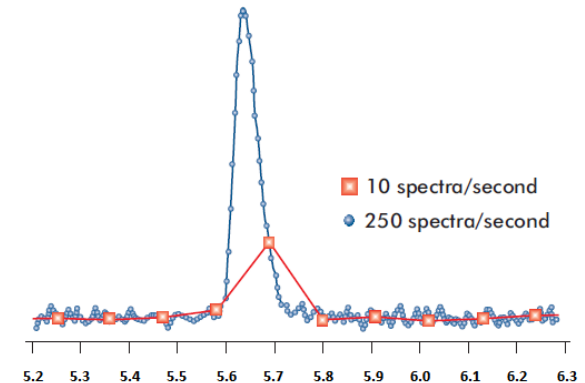
# Dwuwymiarowa chromatografia gazowa (GC×GC; 2D GC)



Schemat system do GC×GC ([de.leco-europe.com](http://de.leco-europe.com))

## Detektory GCxGC

- **FID** (maks. 300 Hz)
- **TOFMS** (maks. 500 Hz)
- **μ-ECD** (50-100 Hz)



Szybkość akwizycji danych

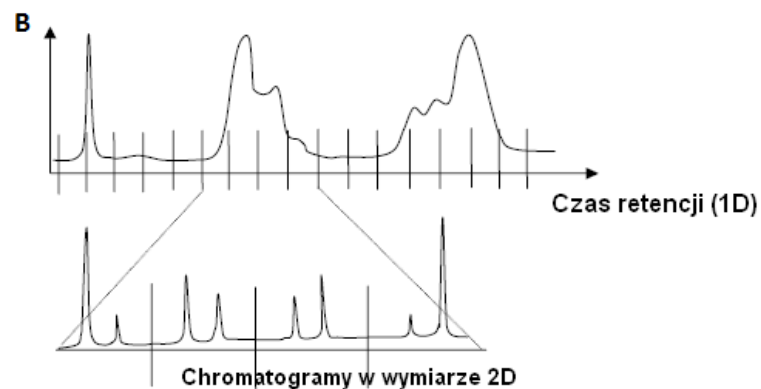
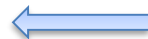
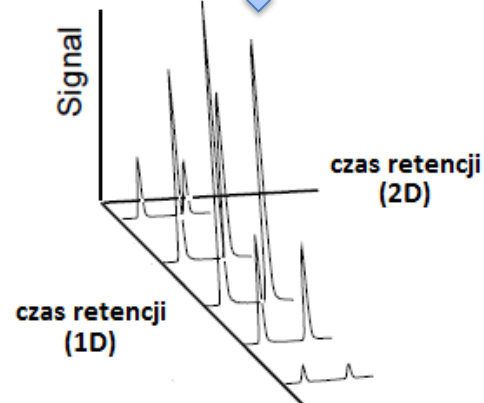
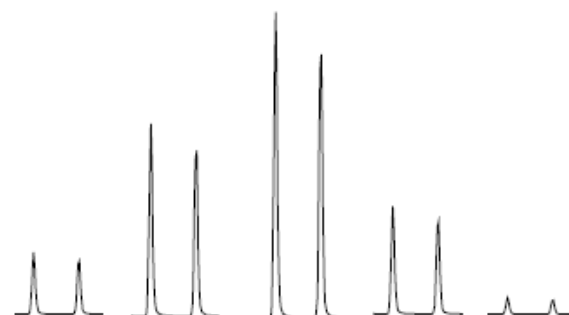
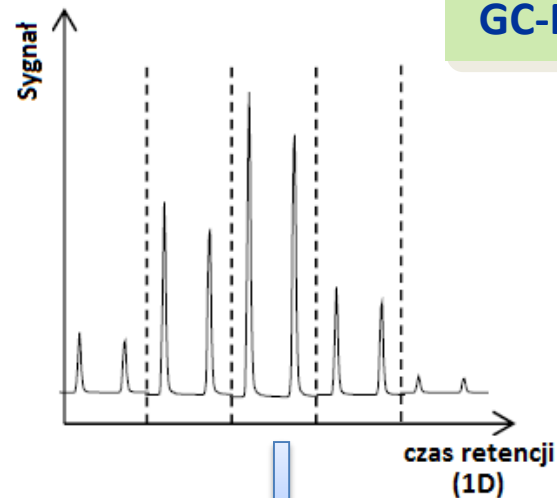
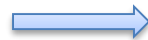
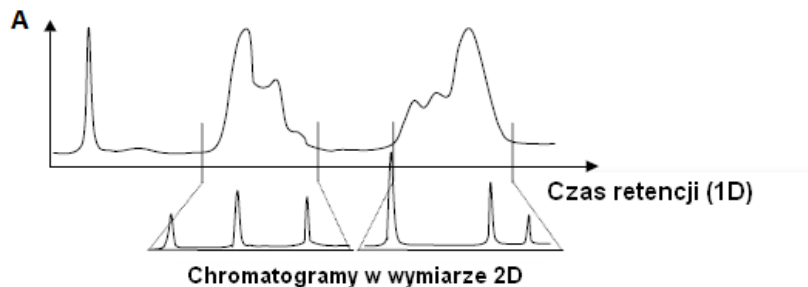
## Kolumny GC×GC

Parametry	kolumna 1D	kolumna 2D
długość	15-30 m	0,5-2m
średnica wewnętrzna	0,25 mm	0,1 mm
grubość filmu fazy stacjonarnej	0,25-1 μm	0,1-0,25 μm
rodzaj fazy stacjonarnej	niepolarna 100% dimetylopolisiloksan lub 5% fenylo/95% dimetylosiloksan	polarna 50% fenylo/50% dimetylosiloksan lub poliglikolowa(Carbowax)



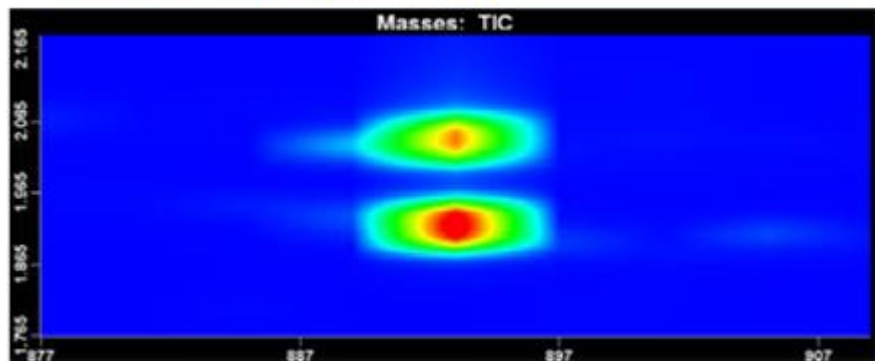
# GC×GC

GC-MS

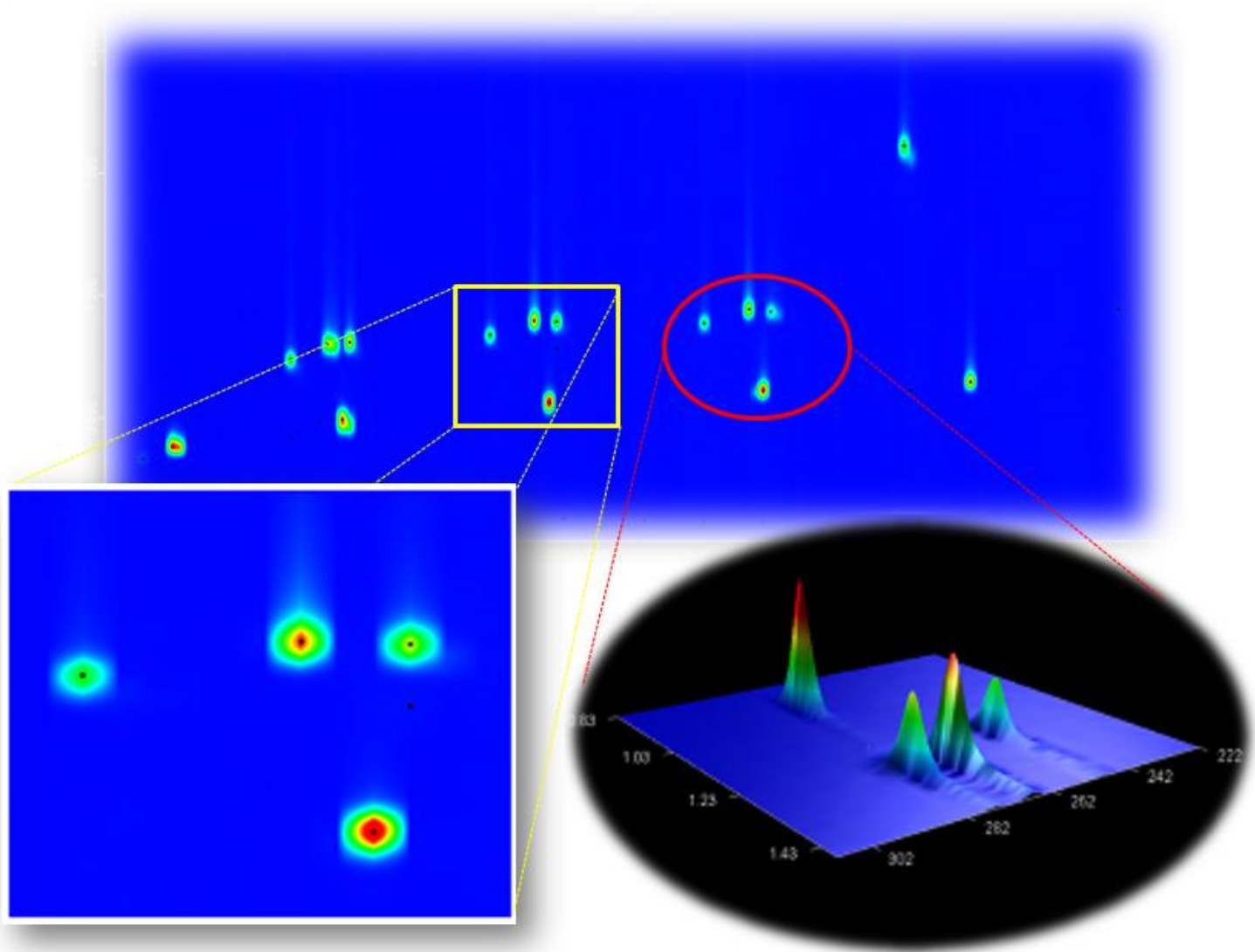


(A) Dwie frakcje opuszczające I kolumnę kierowane są po kolei do II kolumny  
(B) kompletna analiza dwuwymiarowa

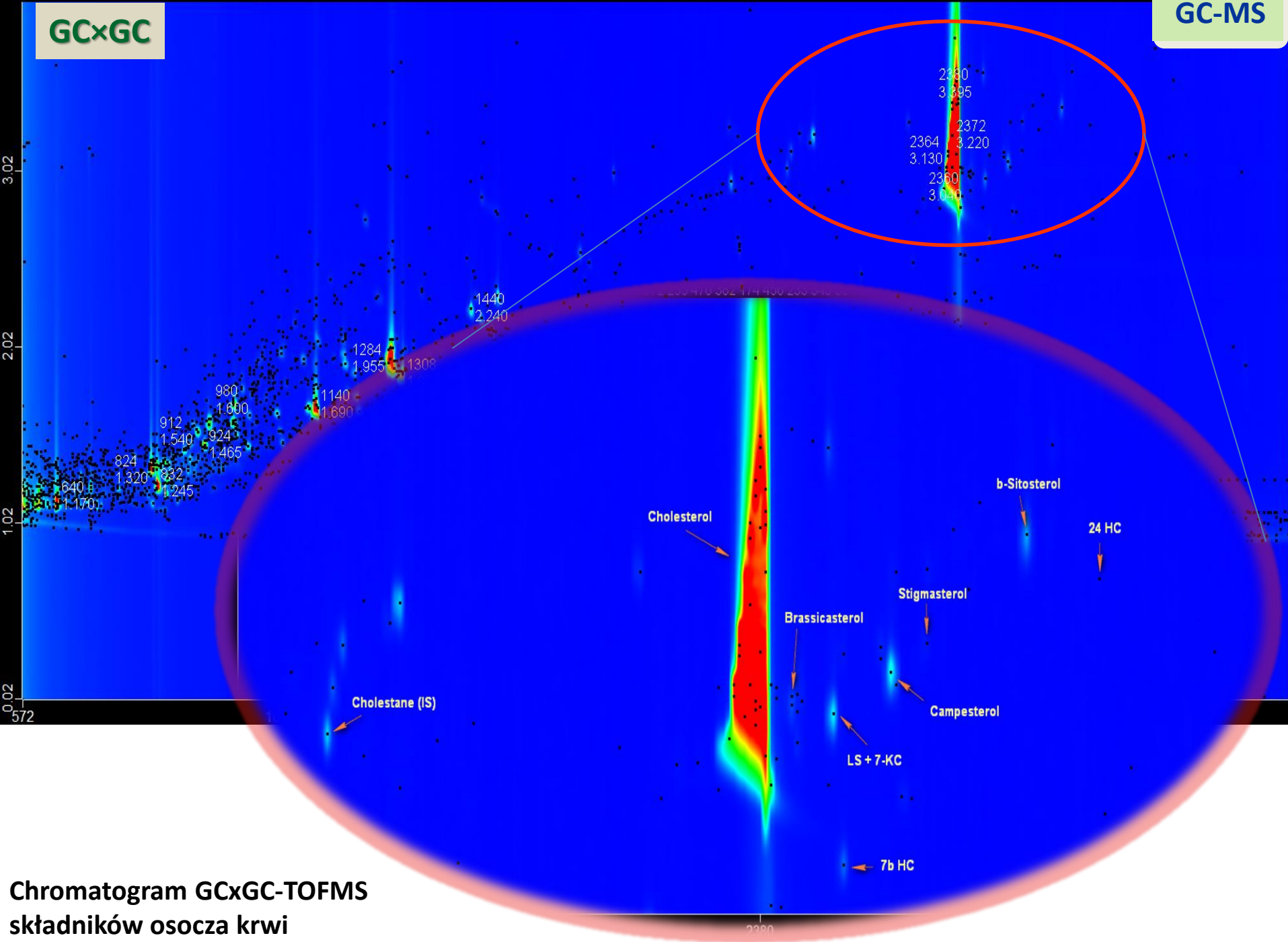
## Chromatogram 3D (widok z góry)



GC×GC



Chromatogram 2D GCMS mieszaniny związków lotnych (alkany, alkohole, aldehydy, ketony, aminy)  
[Pegasus 4D; Leco]



Chromatogram GCxGC-TOFMS  
składników osocza krwi

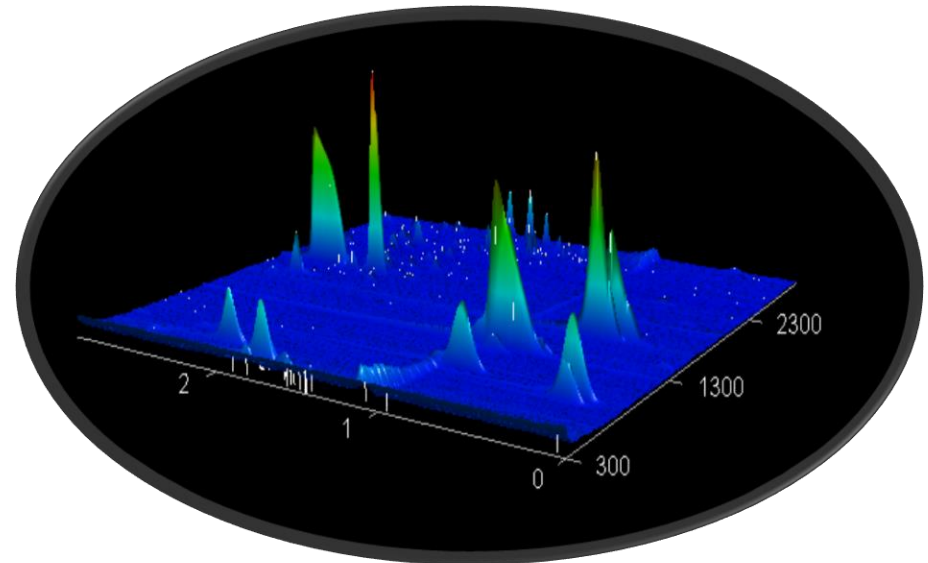
## GC×GC

## Zalety

- Zwiększona pojemność pików - możliwa analiza bardzo złożonych próbek
- Całkowite rozdzielenie wszystkich składników próbki
- Większa czułość oznaczeń w porównaniu z 1D GC (wzrost stosunku sygnału do szumu S/N)
- Uzyskuje się chromatogramy strukturalne
- Chromatogram 2D GC zawiera więcej informacji niż 1D GC - łatwiejsza i bardziej wiarygodna identyfikacja nieznanymi substancjami
- Uporządkowana natura chromatogramów

## Wady

- Wymagane detektory o dużej szybkości akwizycji danych (TOF,  $\mu$ ECD, FID)
- Duża pojemność zarejestrowanych danych – czasochłonne przetwarzanie danych
- Bardzo wysoki koszt aparatury





e-materiały opracowane w ramach projektu  
**Advanced Analytical Chemistry for Life Sciences [AACLifeSci]**

Projekt został zrealizowany przy wsparciu finansowym Komisji Unii Europejskiej

Niniejsze opracowanie odzwierciedla jedynie poglądy autorów,  
a Komisja nie ponosi odpowiedzialności za jakiekolwiek wykorzystanie informacji zawartych w tej publikacji

# Elektroforeza kapilarna- spektrometria mas (CE-MS)



# Elektroforeza kapilarna

- Separacja analitów oparta na ruchliwości elektroforetycznej
- Prosta aparatura
- Podstawowe zastosowanie w bioanalizie
  - Sekwencjonowanie DNA
  - Analiza fragmentów DNA
- Wiele udoskonalień w celu większej selektywności jonów obojętnych
  - MEKC
  - CEC



# Zalety i Wady CE

## Zalety

- Oferuje nową selektywność, alternatywę dla HPLC
- Łatwa i przewidywalna selektywność
- Wysoka wydajność separacji (10<sup>5</sup> do 10<sup>6</sup> pól teoretycznych)
- Małe rozmiary próbek (1-10  $\mu$ l)
- Szybka analiza (od 1 do 45 minut)
- Może być zautomatyzowana
- Oznaczenia ilościowe
- Różne „tryby pracy”

## Wady

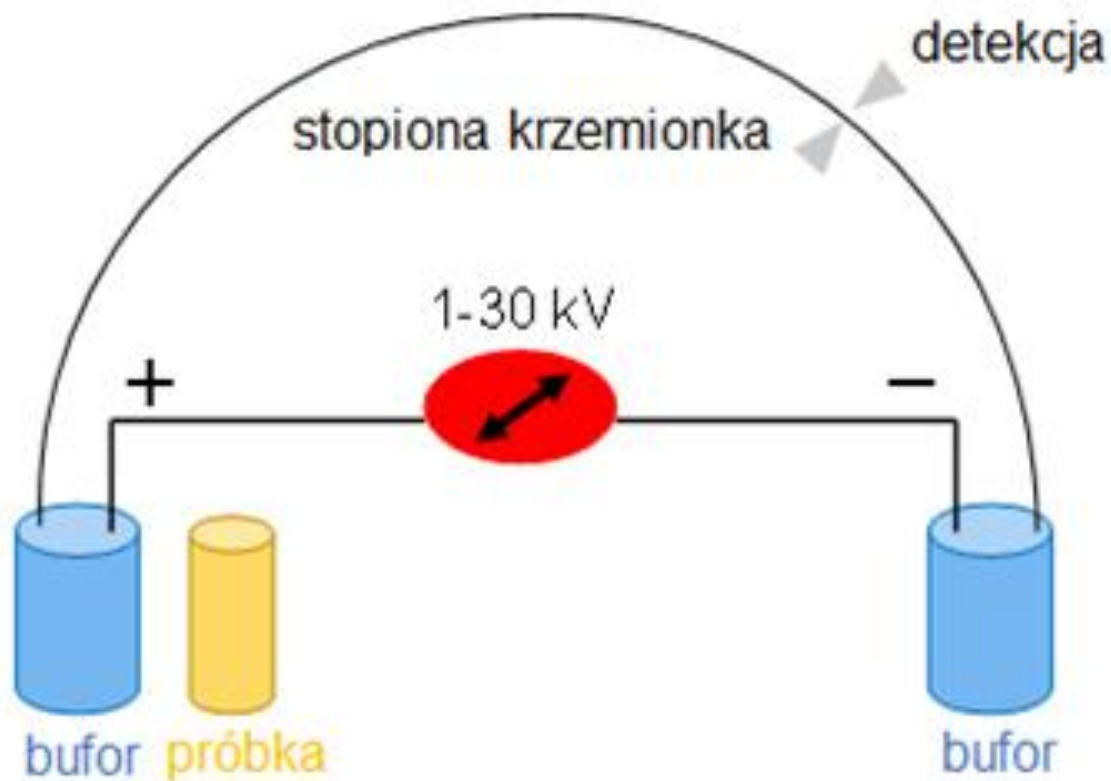
- Nie można wykonać skali preparatywnej separacji
- Niskie stężenia i mała objętość nastrzyku
- „Lepkie” związki
- Związki trudnorozpuszczalne
- Problemy z odtwarzalnością



# Aplikacyjność CZE

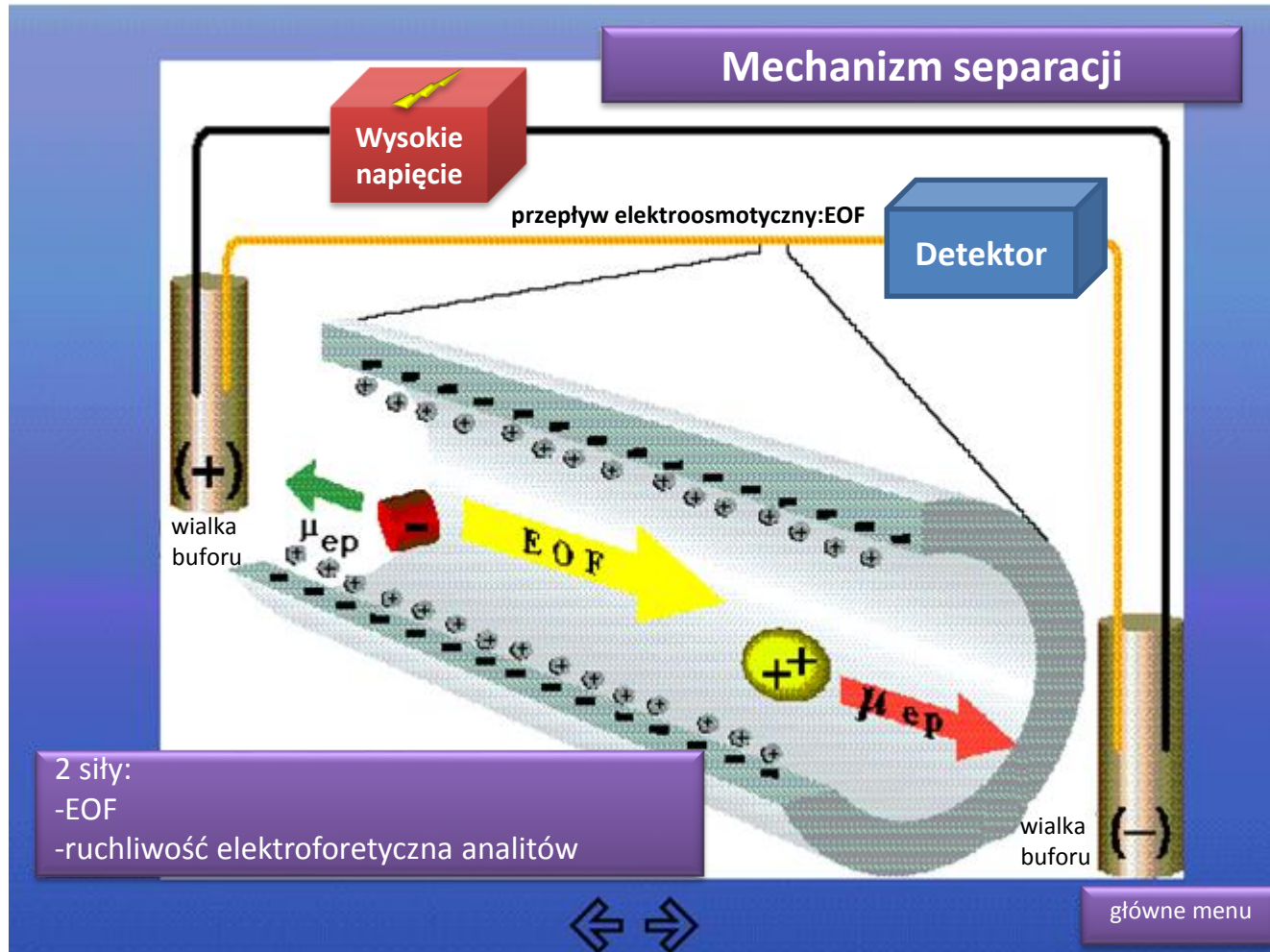
- Szeroki zakres zastosowań
  - Mikrocząsteczki
  - Makrocząsteczki (białka, peptydy)
- Ograniczenia
  - Muszą znajdować się jony o różnych ładunkach
  - Niska siła jonowa próbki
- Zalety
  - Prosta metoda
  - Bezpośrednia analiza złożonych systemów biologicznych

# Elektroforeza kapilarna (CE)



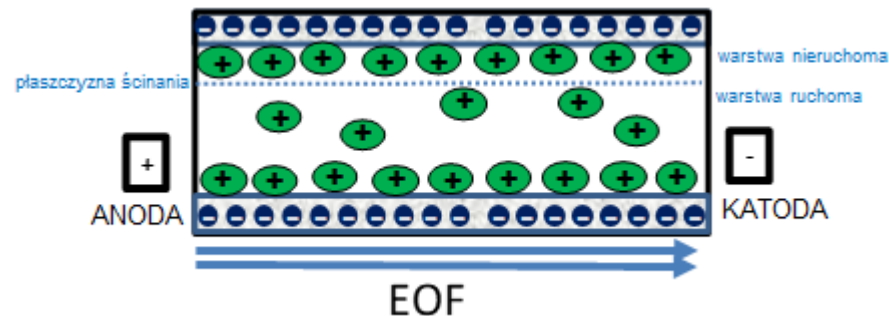
- Kapilarna elektroforeza strefowa (CZE)
  - Tryb podstawowy wykorzystujący otwarte kanały
- Micelarna chromatografia elektrokinetyczna (MEKC)
  - Rozdział związków za pomocą miceli
- Kapilarna elektroforeza żelowa
  - Wykluczanie z wykorzystaniem rozmiarów analitów - żele przesiewowe
- Elektrochromatografia kapilarna
  - Połączenie CE i HPLC
- Ogniskowanie izoelektryczne
- Enancjomery

# Kapilarna elektroforeza strefowa (CZE)



# Przepływ elektroosmotyczny

- Mechanizm przepływu zależy od zastosowanego potencjału oraz pH
- Zapewnia przepływ objętościowy w kapilarze, który modulują bufony o niskim stężeniu
- Zależy od siły jonowej oraz od chemicznych właściwości powierzchni



# Ruchliwość elektroforetyczna

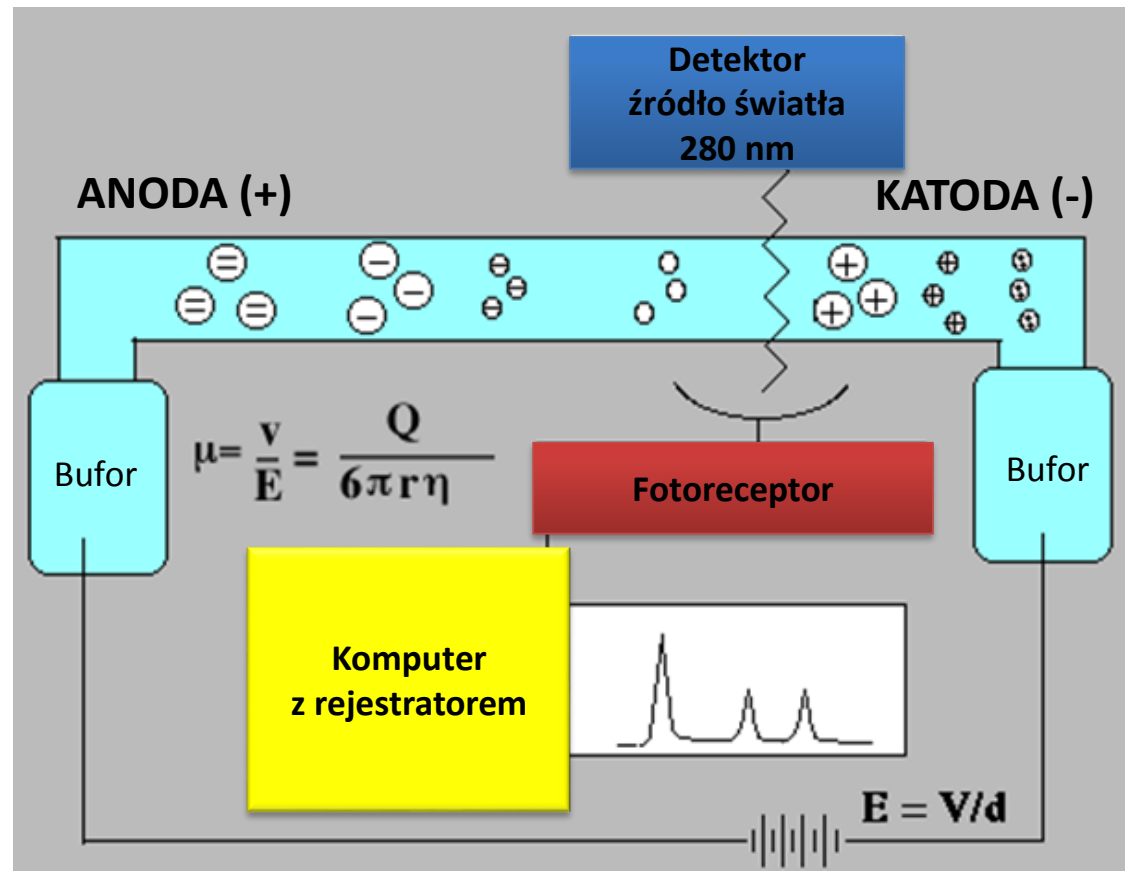
$$\mu = \frac{q}{6\pi\eta r}$$

$\mu$  = ruchliwość elektroforetyczna

$Q$  = ładunek cząsteczki

$\eta$  = lepkość ośrodka

$r$  = promień jonowy cząsteczki

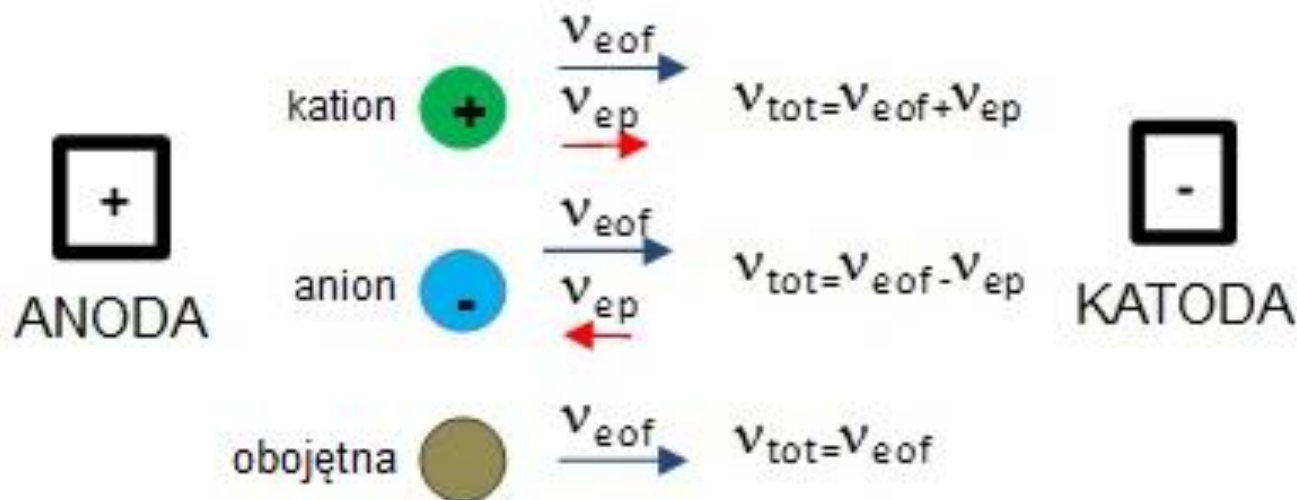


# Elektroforeza i Elektroosmoza

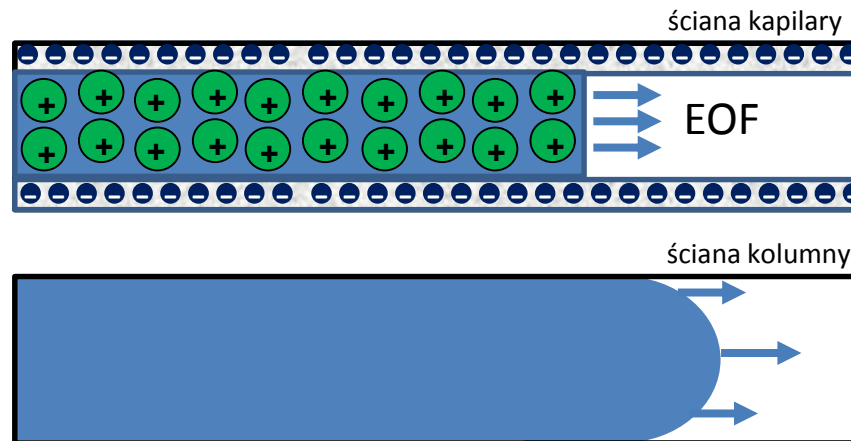
Obrazowe przedstawienie efektu połączenia obu zjawisk w kapilarze, gdy EOF jest szybsze niż EP

$$v = (\mu_{ep} + \mu_{eo})E = (\mu_{ep} + \mu_{eo})\frac{V}{L}$$

Figure from R. N. Zare, Stanford



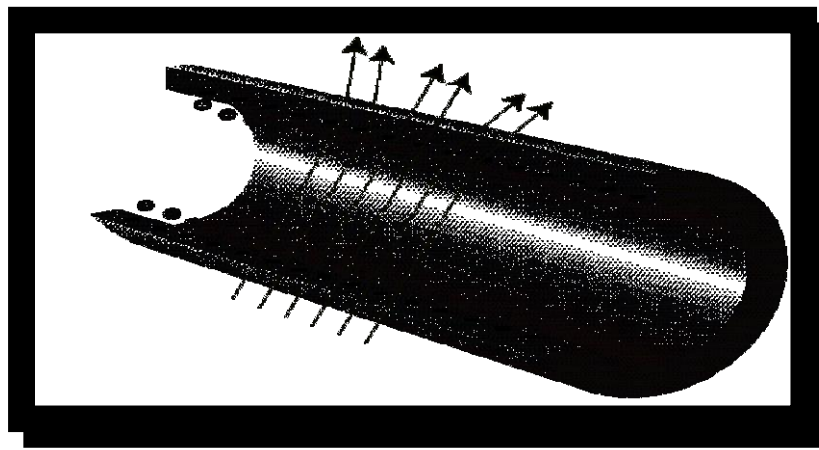
# Schemat przepływu w CE i HPLC





# Rodzaje detekcji

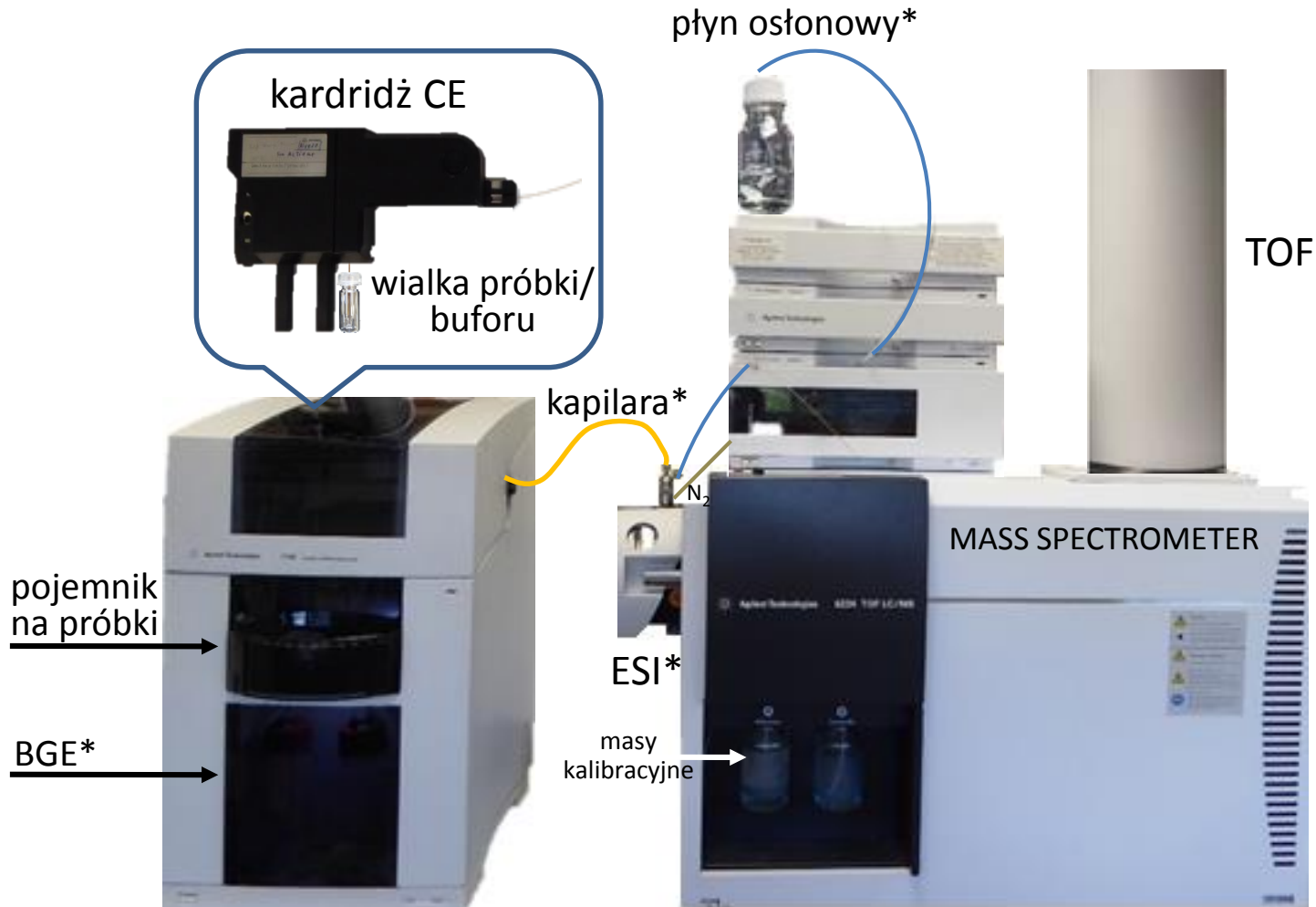
Detektor	Zalety	Charakterystyka	Limit detekcji (M)
UV/VIS	<ul style="list-style-type: none"><li>- Możliwość bezpośredniej i pośredniej detekcji</li><li>- często stosowany detektor</li></ul>	<ul style="list-style-type: none"><li>-Uniwersalny</li></ul>	<ul style="list-style-type: none"><li>-<math>10^{-3}</math></li><li>-<math>10^{-6}</math> związki aromatyczne</li></ul>
LIF	<ul style="list-style-type: none"><li>-czuły i selektywny</li><li>- Używany do związków fluorescencyjnych lub pochodnych</li></ul>	<ul style="list-style-type: none"><li>-Selektywny</li></ul>	<ul style="list-style-type: none"><li>-<math>10^{-6}</math> - <math>10^{-9}</math></li></ul>
MS	<ul style="list-style-type: none"><li>-Informacje jakościowe i ilościowe</li><li>- wysoka czułość i selektywność</li></ul>	<ul style="list-style-type: none"><li>-Uniwersalny</li><li>-Selektywny</li></ul>	<ul style="list-style-type: none"><li>-<math>\approx 10^{-5}</math> (zależy to od rodzaju detektora MS i rodzaju metabolitów)</li></ul>



# Optymalizacja parametrów separacji w CE

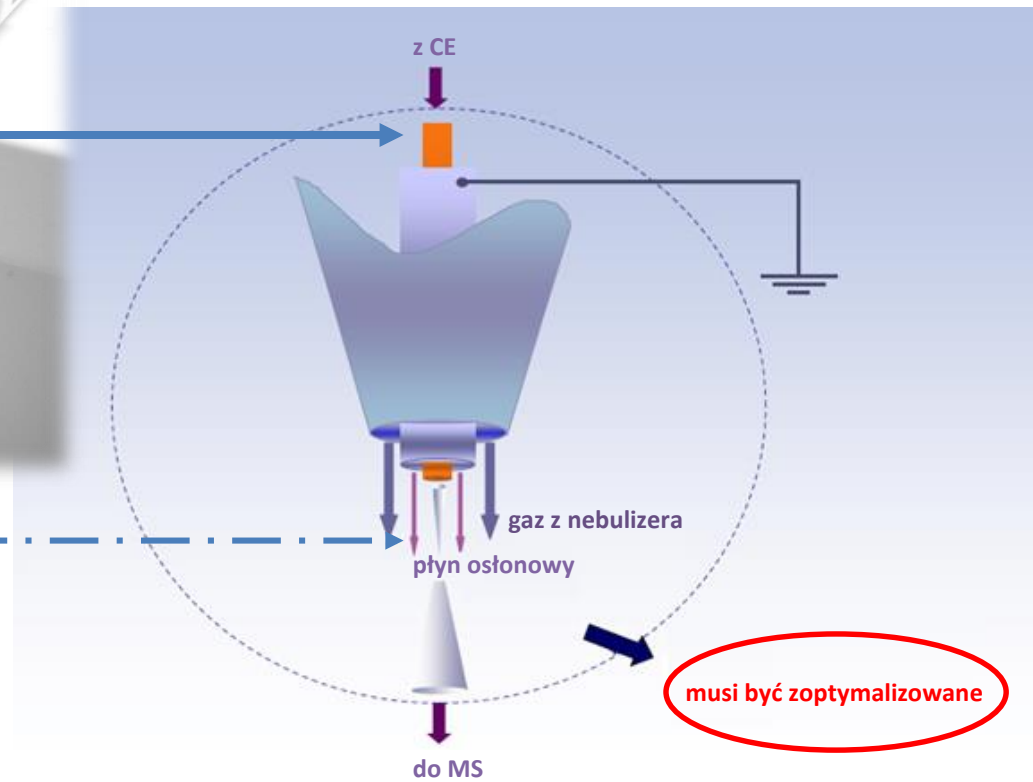
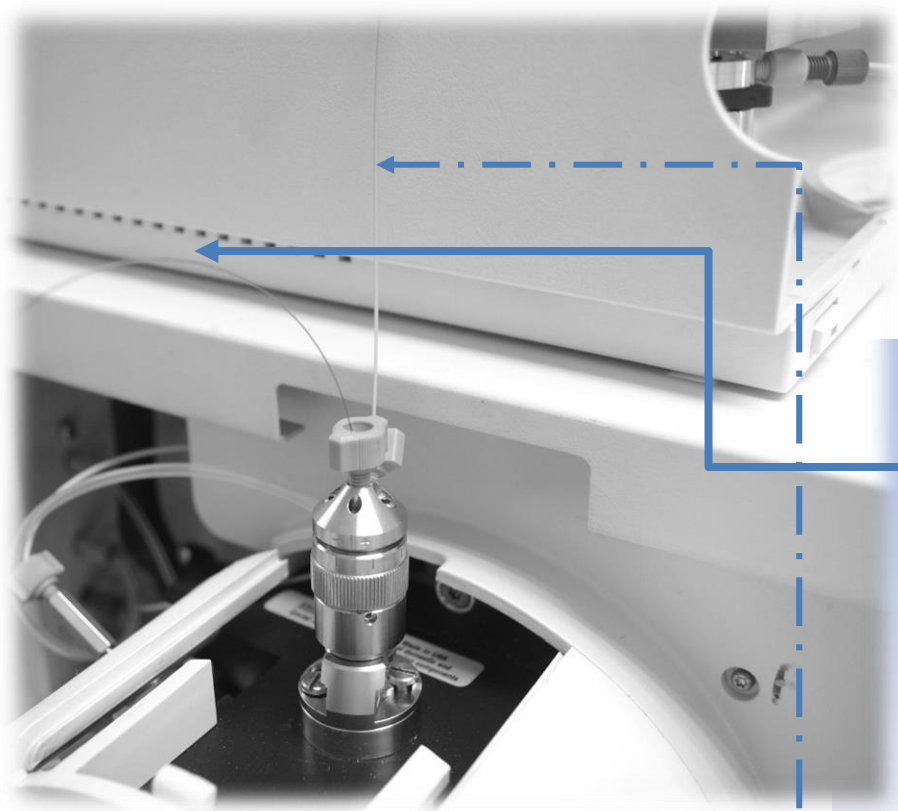
- pH
  - Pierwszy parametr do kontroli
  - Efekt EOF i ruchliwości (ładunek)
- Rozpuszczalnik organiczny
  - Solwatacja analitu
- Czynniki interakcji
  - Parowanie jonów, solwatacja, itp.
- Warunki bezwodne
  - Solwatacja i ładunek
- Temperatura
  - Solwatacja, równowaga chemiczna

# Aparatura CE-MS

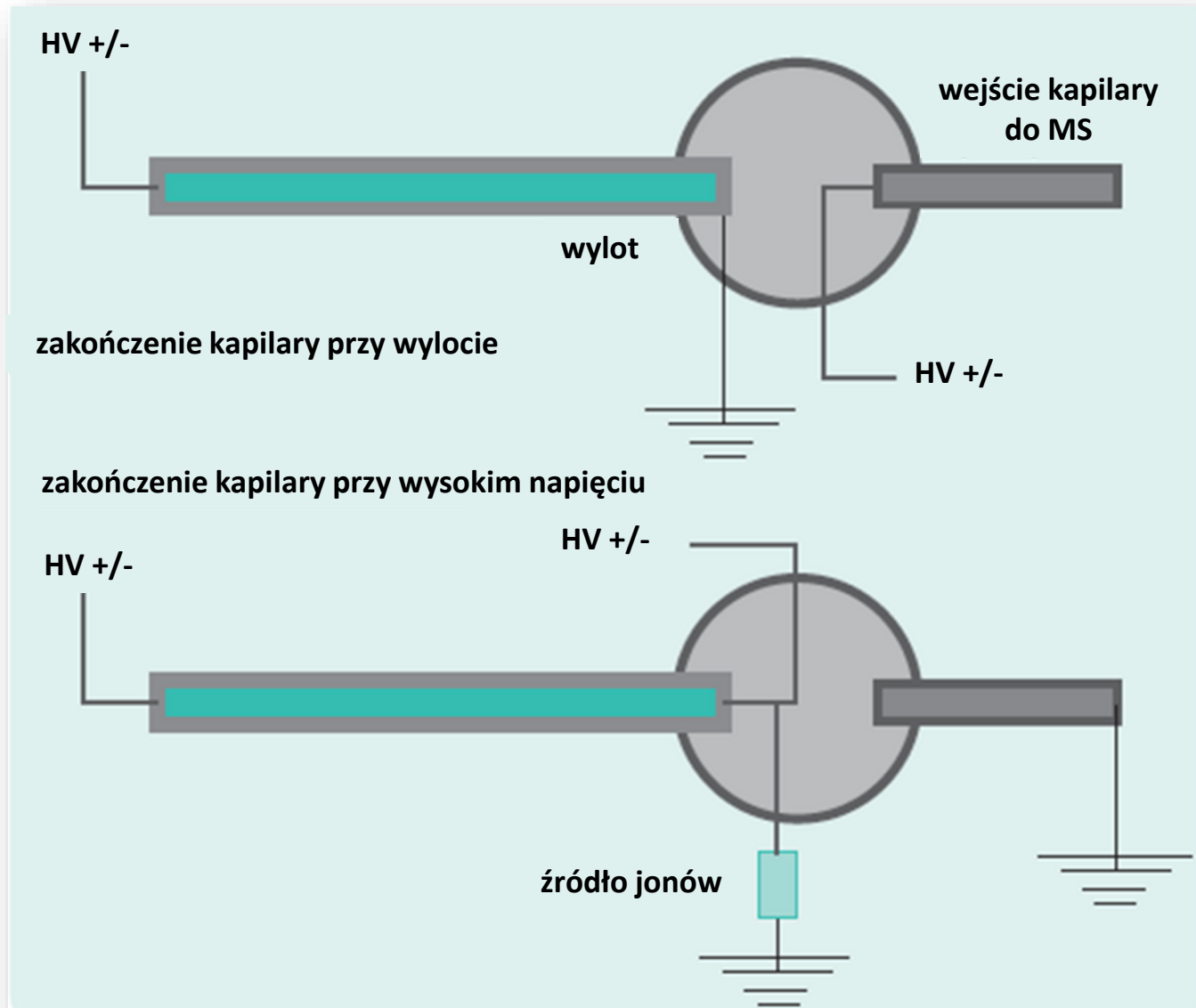


\* WYMAGANA OPTYMALIZACJA

# Źródło ESI CE-MS

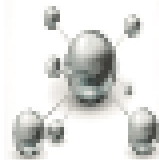


# CE-MS: Interfejs elektryczny



# METABOLOMIKA

AAALifeSci



Erasmus+

# STRESZCZENIE

## Wstęp do metabolomiki

### 1. Podejścia analityczne w metabolomice

- Przepływ pracy w badaniu metabolomicznym
- Kontrola jakości i procedura zapewnienia jakości w metabolomice

### 2. Przetwarzanie danych i identyfikacja metabolitów

- Przetwarzanie danych metodą pipeline
- Nietargetowe przetwarzanie danych metabolomicznych
- Identyfikacja metabolomiczna
- Analiza statystyczna

### 3. Analiza danych

- Od identyfikacji danych do szlaków biochemicznych
- Walidacja biomarkerów

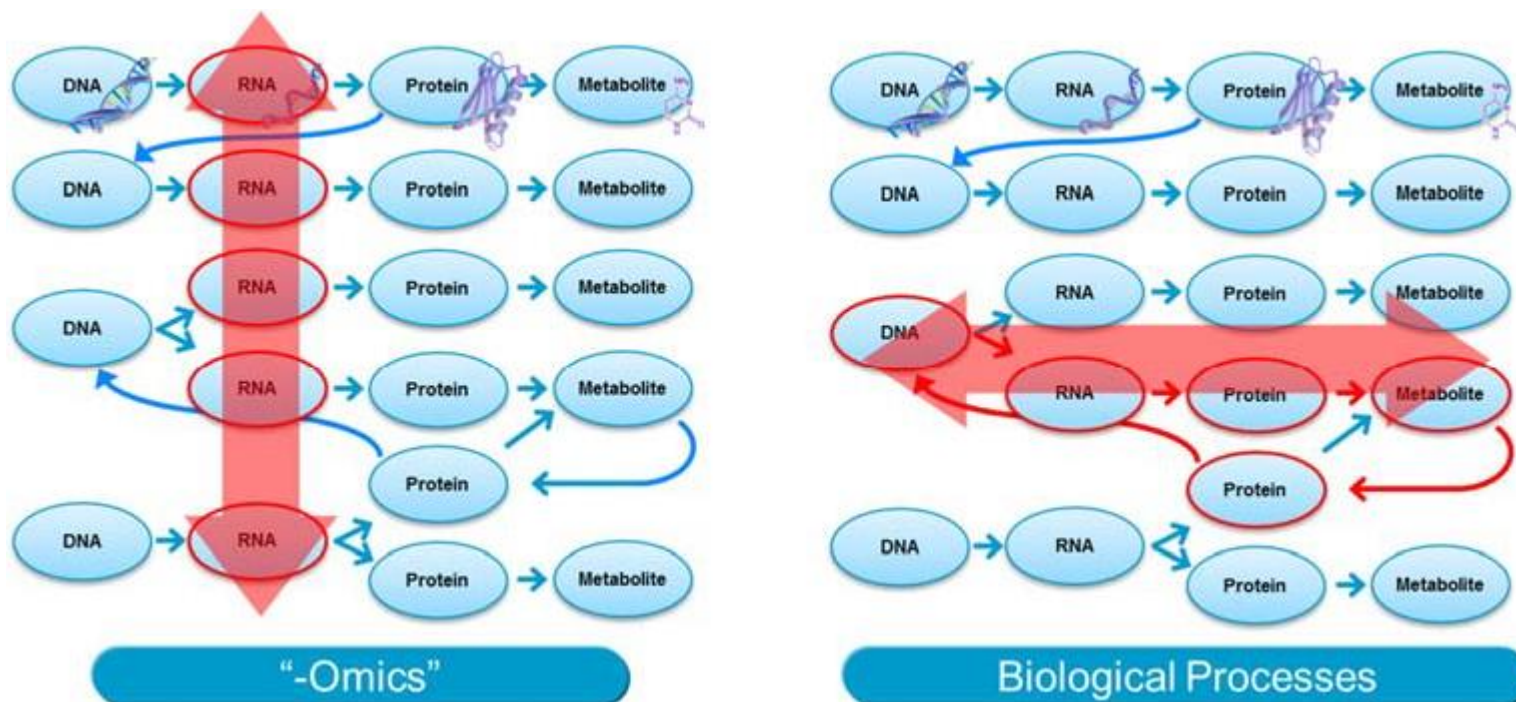
### 4. Część praktyczna

- Metabolomika targetowa i nietargetowa
- Metabolomika wraz z wolnodostępnymi narzędziami online



## Metabolomika

Nowa wschodząca dziedzina badań "omicznych" (obejmująca genomikę, proteomikę i metabolomikę), zajmuje się kompleksową charakterystyką cząsteczek o niskich masach cząsteczkowych, czyli metabolitów obecnych w systemach biologicznych.





## **Dodatek: definicja metabolomiki**

- Pomiar dynamicznej, wieloparametrycznej odpowiedzi metabolicznej systemów żywych organizmów w stosunku do bodźców patofizjologicznych lub genetycznych modyfikacji (Nicholson, 1999)
  - ilościowy pomiar związanego z czasem "całkowity" odpowiedzi metabolicznej na bodźce patofizjologiczne (żywieniowe, ksenobiotyczne, chirurgiczne lub toksyczne)
- MetaboLomics - obraz, MetaboNomics – film
- W dzisiejszych czasach wszystko jest Metabolomiką

## Definicja metabolomu

- kompletny zestaw wszystkich metabolitów o niskiej masie cząsteczkowej, które, zmieniają się zgodnie z fizjologią, rozwojem stanu patologicznego komórki, tkanki, narządu lub organizmu ... "(Oliver 2002)
- Pochodzenie: Endometabolom, Microbiom, Xenobiom, Nutriom...
- Natura: Glikom, Lipidom, Sfingolipidom, Peptydom ...
- Metabolom      Fenotyp



Gospodarz



Microbiome



Nutriome

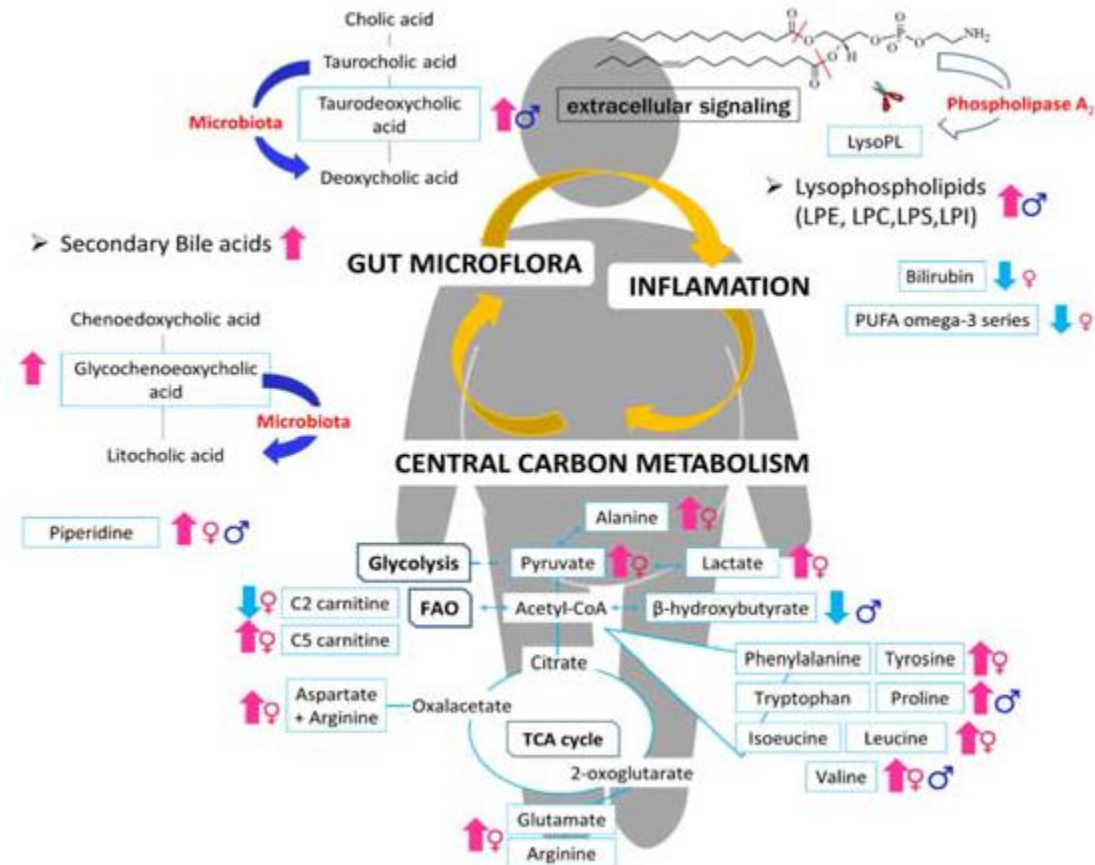


Xenobome

# Co może zapewnić metabolomika (I)

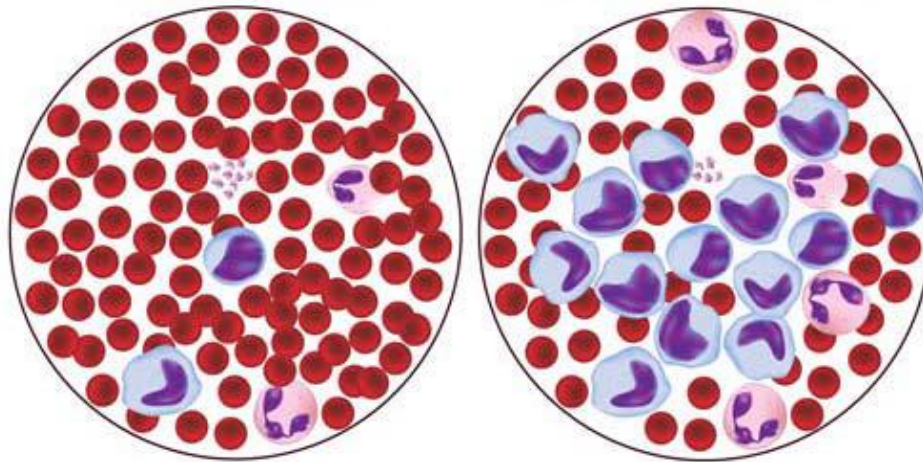
- Przegląd stanu metabolicznego i globalnych zdarzeń biochemicznych związanych z systemem komórkowy lub biologicznym.

-stany patologiczne bez znanego mechanizmu, tj. związek pomiędzy otyłością i insulinoopornością



## Co może zapewnić metabolomika (II)

- Identyfikacja (propozycja) nowych biomarkerów, ważnych w procesie odkrywania nowych leków lub jako narzędzie diagnostyczne w badaniach in vitro.
  - na przykład nowe biomarkery diagnostyczne w postaci agresywnej, przewlekłej białaczki limfatycznej



Utility of validated metabolites as biomarkers of aggressive state of CLL

Metabolite	AUC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Acetylcarnitine	0.695	43.2	93.0	86.1	62.1
Butyrylcarnitine	0.548	10.8	98.0	84.4	52.4
Hexanoylcarnitine	0.690	27.0	96.0	87.1	56.8
Octanoylcarnitine	0.651	29.7	95.0	85.6	57.5
Decanoylcarnitine	0.662	27.0	94.0	81.8	56.3
Palmitoylcarnitine	0.719	40.5	94.0	87.1	61.2
Dodecanamide	0.497	8.1	100.0	100.0	52.1
Hexadecanamide	0.516	5.4	100.0	100.0	51.4
Oleamide	0.600	18.9	96.0	82.5	54.2
Linoleamide	0.672	16.2	98.0	89.0	53.9
Acylcarnitines <sup>a</sup>	0.743	32.4	95.0	86.6	58.4
FAA <sup>b</sup>	0.662	13.9	96.0	77.6	52.7
Acylcarnitines and FAA	0.750	54.0	89.0	83.1	65.9

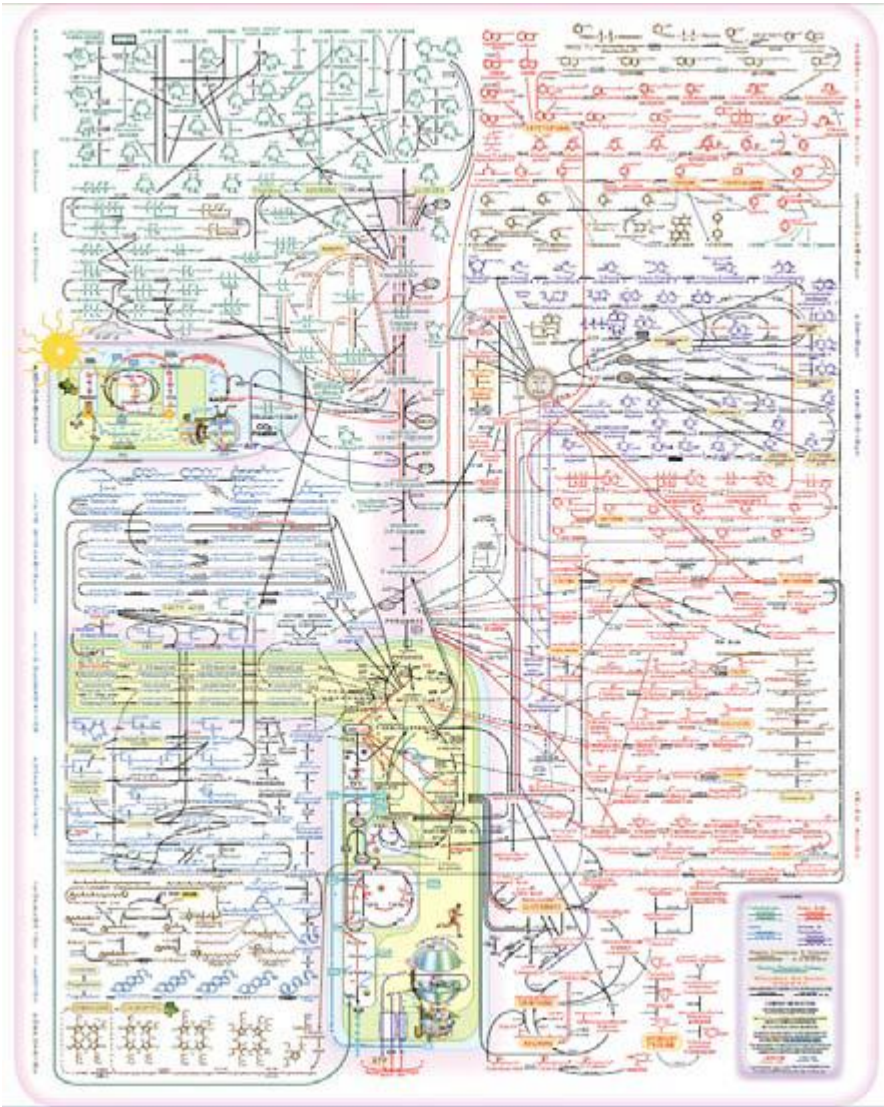
## Do czego może służyć metabolomika.....

- poszukiwanie różnic metabolicznych pomiędzy grupami (chory vs kontrola; przed vs po leczeniu; jeden stan w porównaniu z innym)
- identyfikacja związków, które są znaczące i zaproponowanie mechanizmów
- zdobycie informacji na temat fenotypu
- obserwowanie efektów leczenia
- odnalezienie nowych celów terapeutycznych

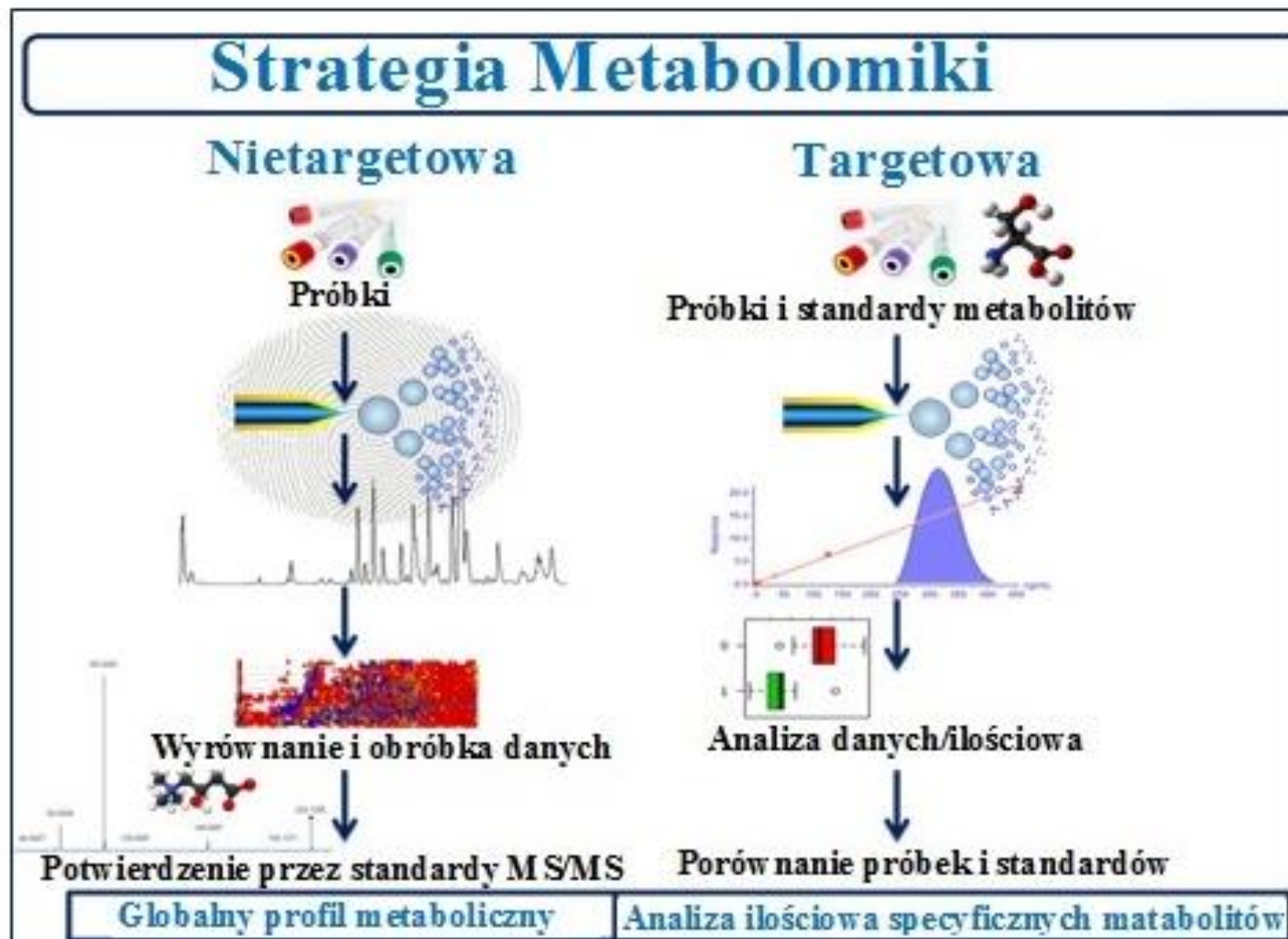
## Czym nie jest metabolomika.....

- metodą, która ukazuje co dzieje się z metabolitami lub lekami
- metodą do oznaczeń ilościowych
- użycie prostego zestawu kitu w celu ilościowego określenia grupy metabolitów (wymaga NMR, MS ...)
- możliwość oznaczeń bez równoczesnego porównania próbek

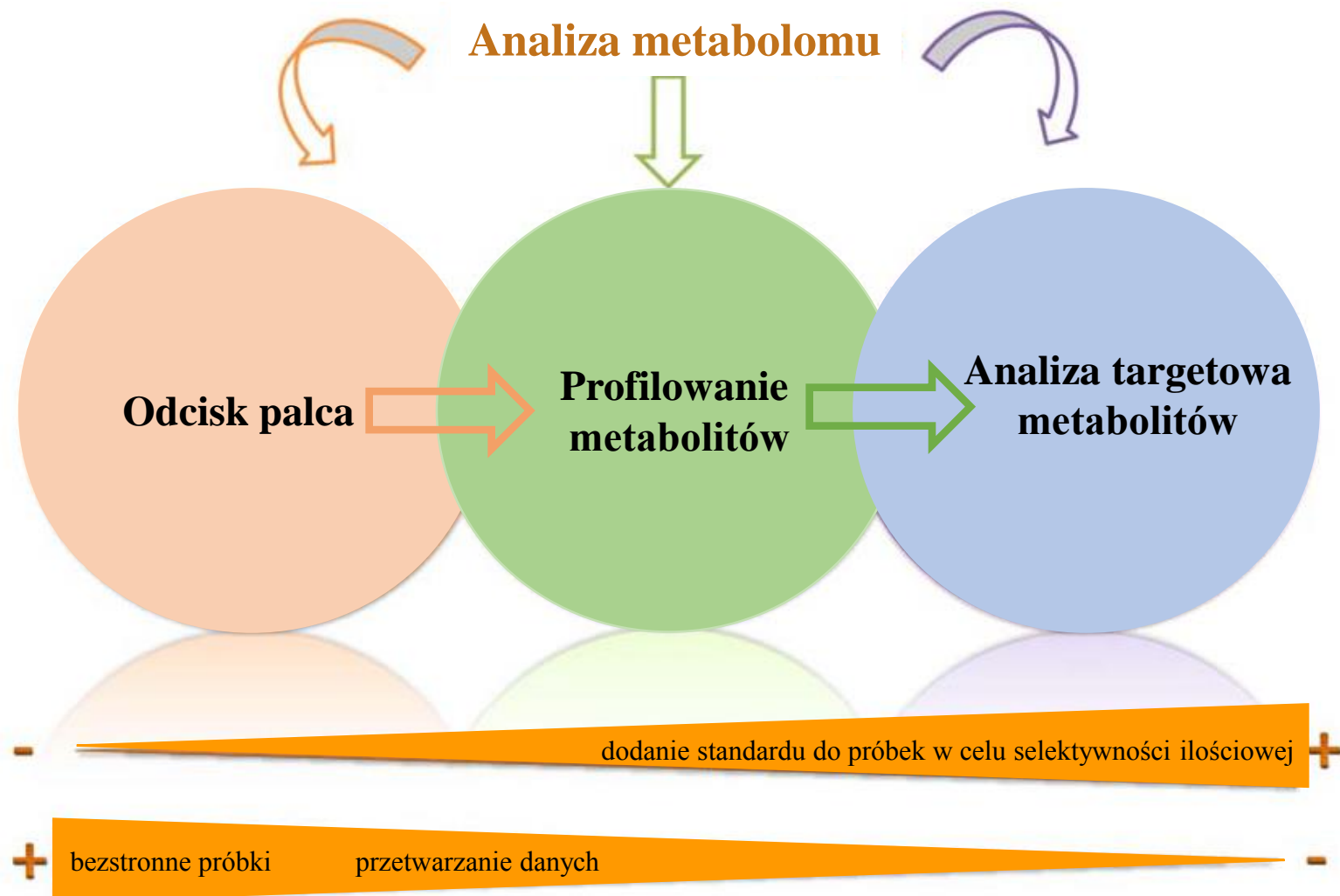
# Definicja metabolizmu



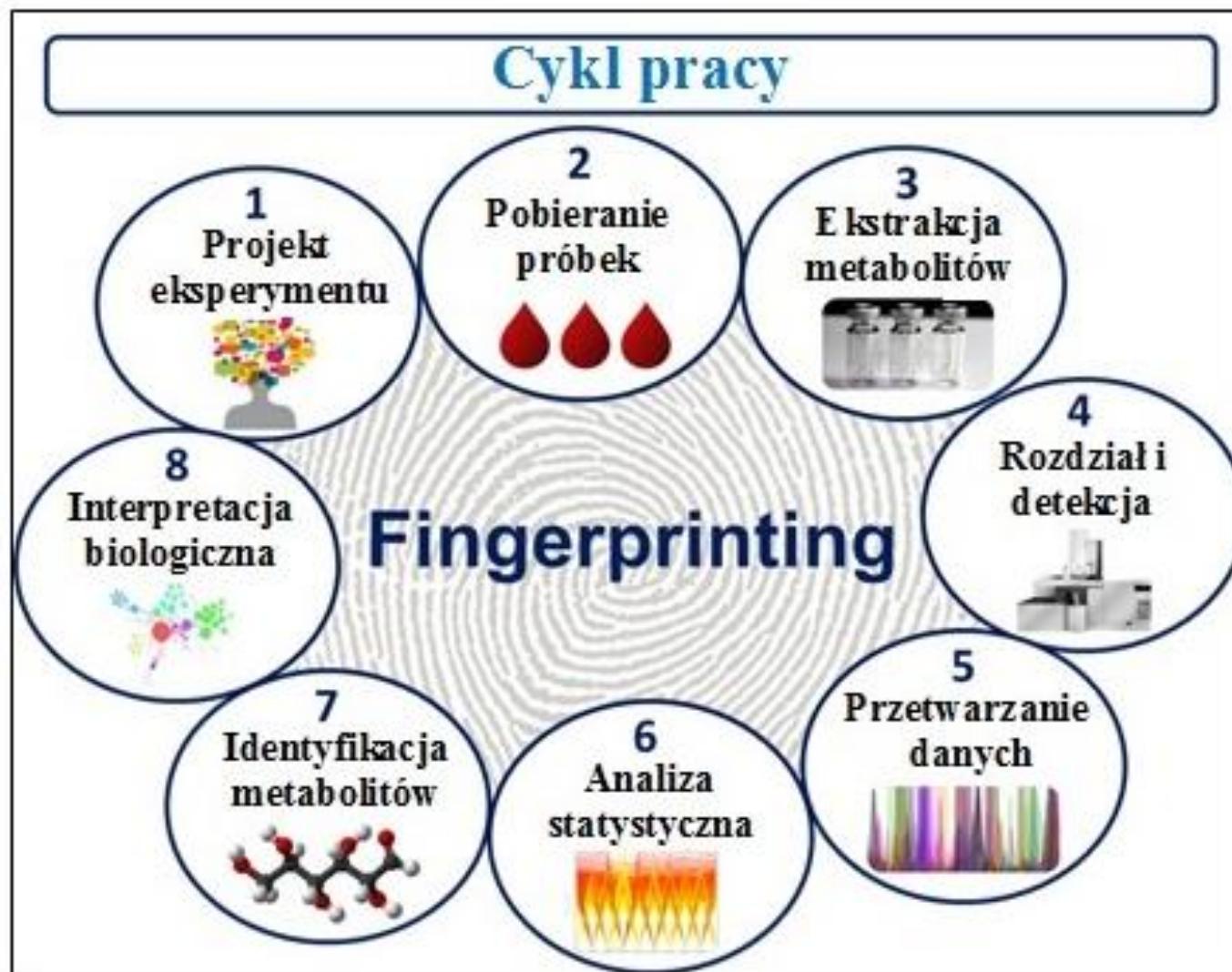
Grupa (bio)chemicznych procesów wewnątrz organelli, komórki, tkanki, narządów lub organizmu, niezbędnych do przeżycia



# Trzy sposoby badań metabolomicznych







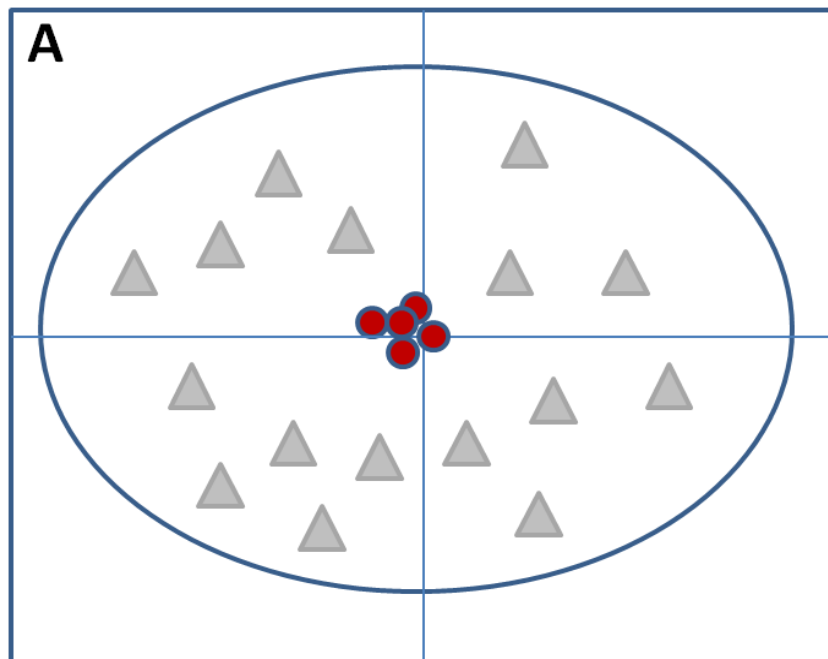
- GC/MS
  - związki polarne o małej masie cząsteczkowej
  - głównie rozpuszczalne w wodzie (niektóre hydrofobowe)
  - obróbka próbki: Derywatyzacja
  - powtarzalna fragmentacja - bazy danych
- NMR
  - rozpuszczalne w wodzie
  - praktycznie bez obróbki próbki
  - wysoki LOD
- LC/MS
  - od małych do dużych (<1500 Da) od średnio polarnych do niepolarnych metabolitów
- CE/MS
  - związki średnio polarne o małej masie cząsteczkowej aminokwasy, acylokarnityny, poliaminy itp.
  - brak derywatyzacji



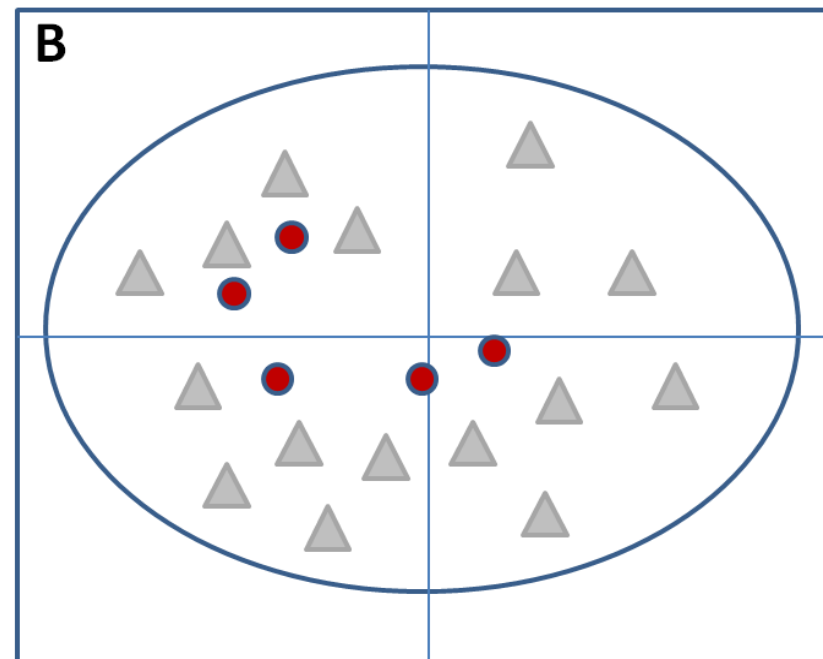
# Platformy analityczne MS w metabolomice

Technika analizy	Aplikacyjność	Zalety	Wady
<b>GC-MS</b>	Rozdział, identyfikacja oraz analiza ilościowa lotnych i termicznie stabilnych metabolitów, mniej polarnych.	Wysoka rozdzielczość chromatograficzna, dostępność bibliotek wielkoformatowych do identyfikacji metabolitów.	Niezdolność do analizowania metabolitów termo-labilnych i wysokocząsteczkowych, wymagana derywatywacja, nielotnych metabolitów.
<b>LC-MS</b>	Rozdział, identyfikacja oraz analiza ilościowa szerokiej grupy metabolitów, w zależności od typu kolumny i fazy ruchomej.	Wysoka czułość, duża pojemność próbki, derywatywacja nie jest wymagana, zdolność do analizy związków termo-labilnych.	Ograniczona dostępności bibliotek komercyjnych, eluentów LC, efekt matrycy, ograniczony potencjał identyfikacyjny, chyba że, stosowana jest technika MS-MS.
<b>CE-MS</b>	Rozdział, identyfikacja oraz analiza ilościowa polarnych i zjonizowanych metabolitów, przy użyciu zredukowanej objętości próbki.	Wysoka rozdzielczość i szybka analiza, użyteczna metoda dla złożonych próbek biologicznych, nawet dla małych objętości próbek.	Ograniczona dostępności bibliotek komercyjnych., niekompatybilność buforów, granice detekcji, ograniczony potencjał identyfikacyjny, chyba że, stosowana jest technika MS-MS.

## Kontrola jakości i procedura zapewnienia jakości w metabolomice

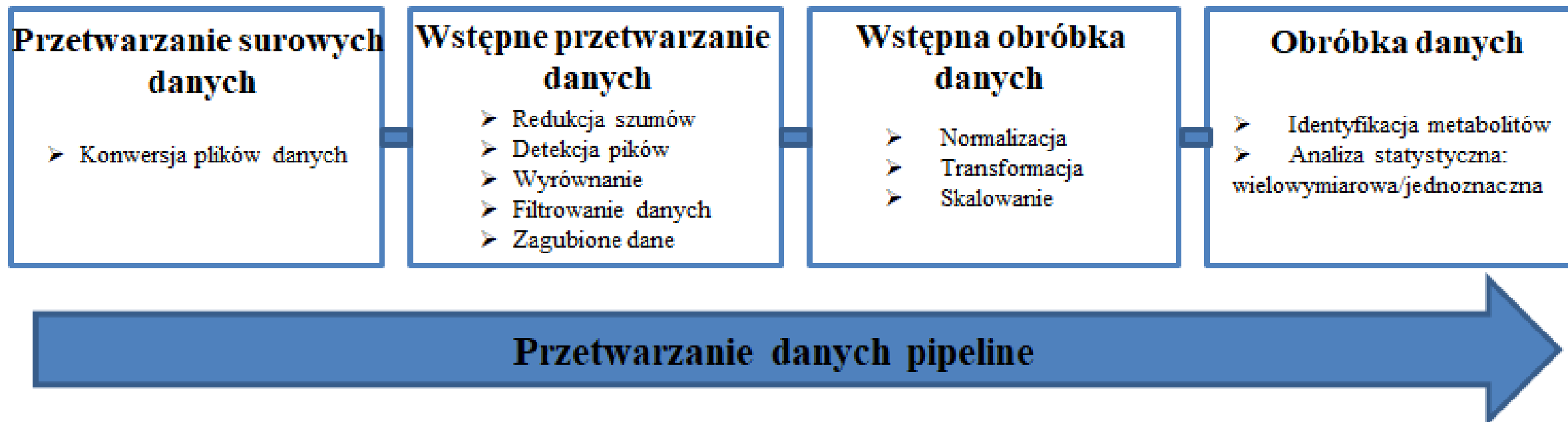


**A:Qcs (czerwone kropki) sklastrowane razem**



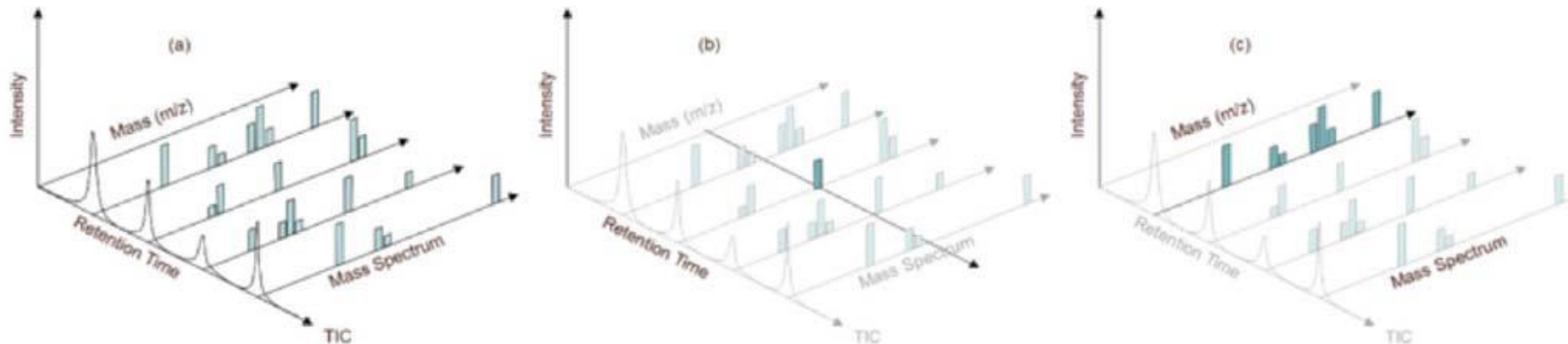
**B:Qcs (czerwone kropki) rozrzucone**

## Analiza danych w metabolomice: Przetwarzanie sygnałów



# Technika analityczna: GC/MS

- Chromatografia gazowa sprzężona ze spektrometrią mas
- Złoty standard
  - bardzo czuła i powtarzalna metoda
  - Informacje: analiza jakościowa i ilościowa
  - biblioteki widm w celu identyfikacji metabolitów
  - 10-20% znanych związków można analizować za pomocą GC
- Wysoka przydatność w metabolomice

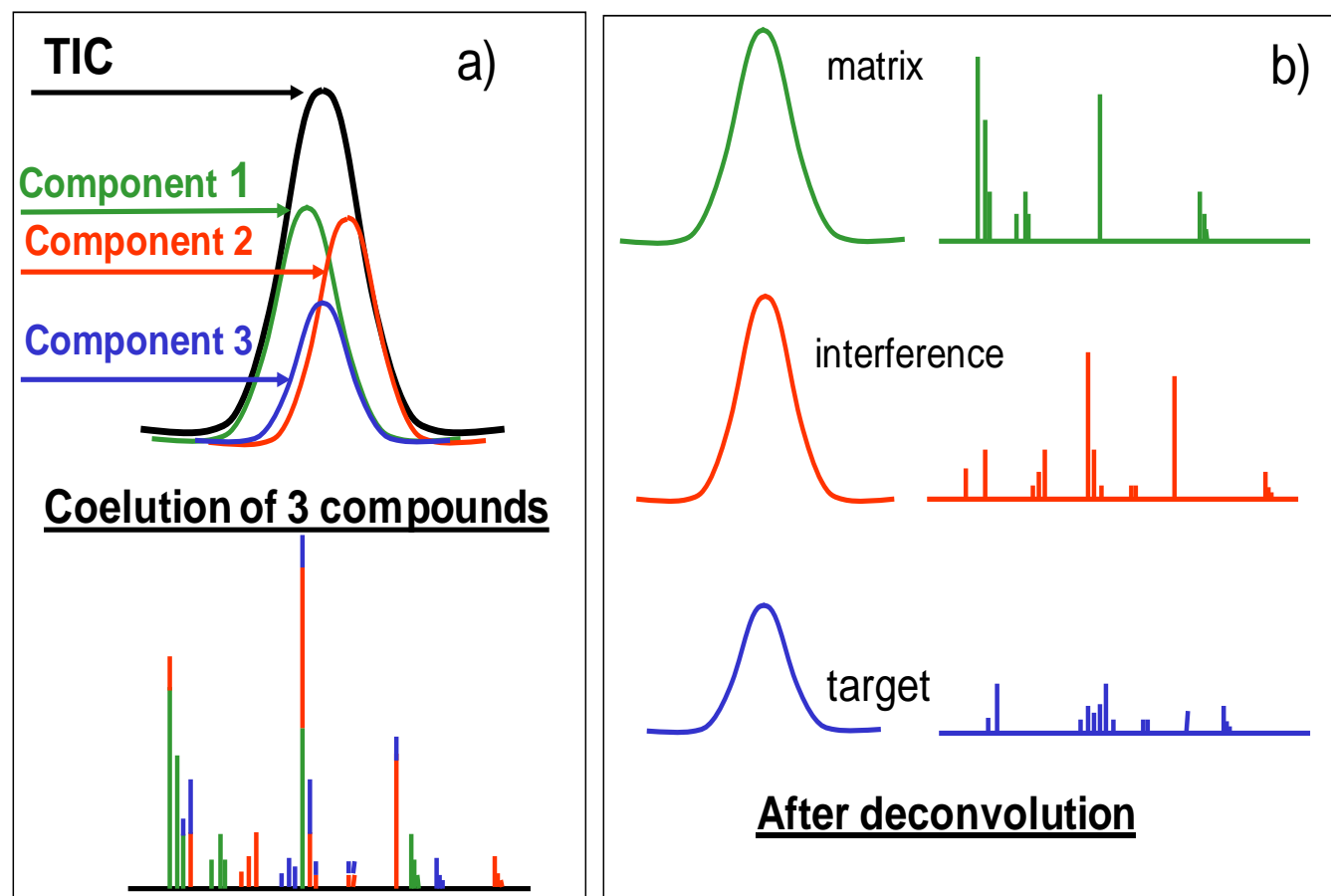


(a) dane 3D GC/MS,

(b) chromatogram wybranego jonu

(c) w pojedynczym punkcie danych w czasie uzyskujemy pojedyncze widmo masowe.

# Dekonwolucja



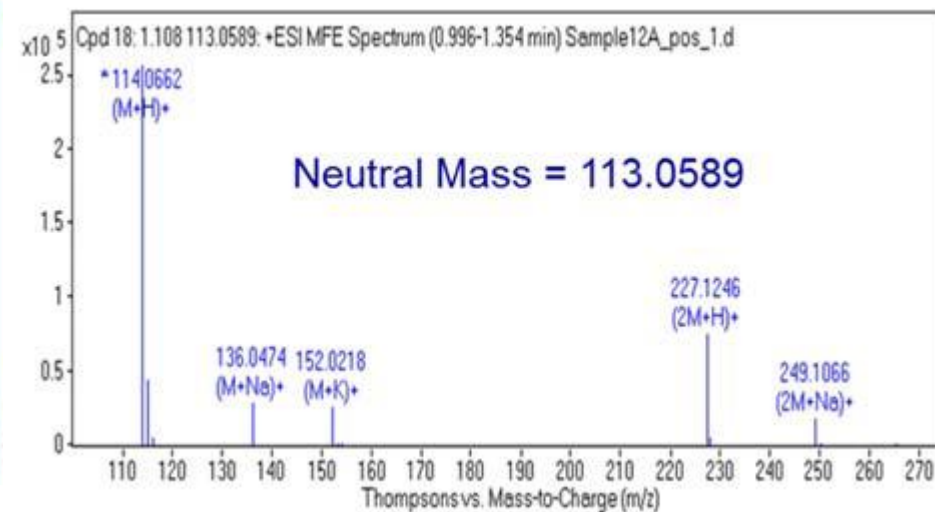
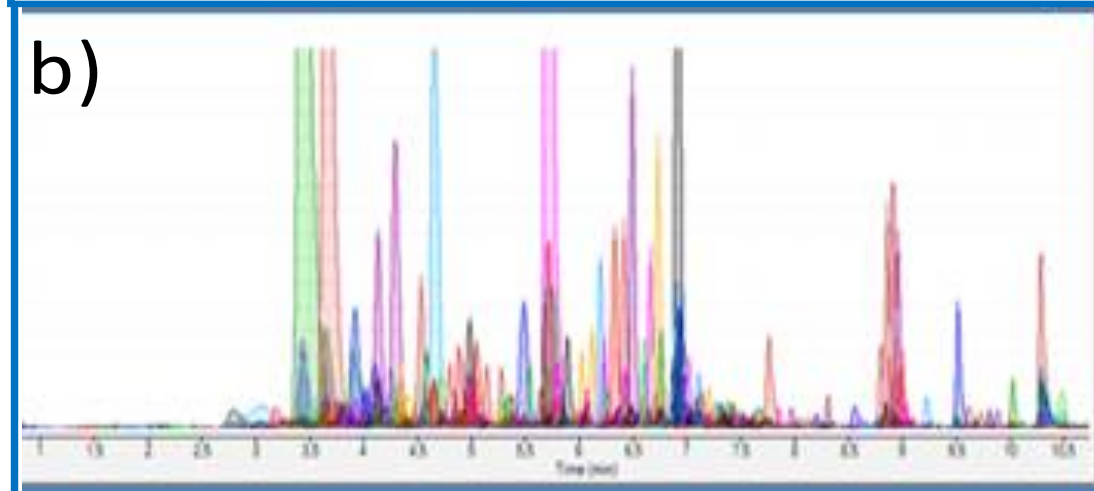
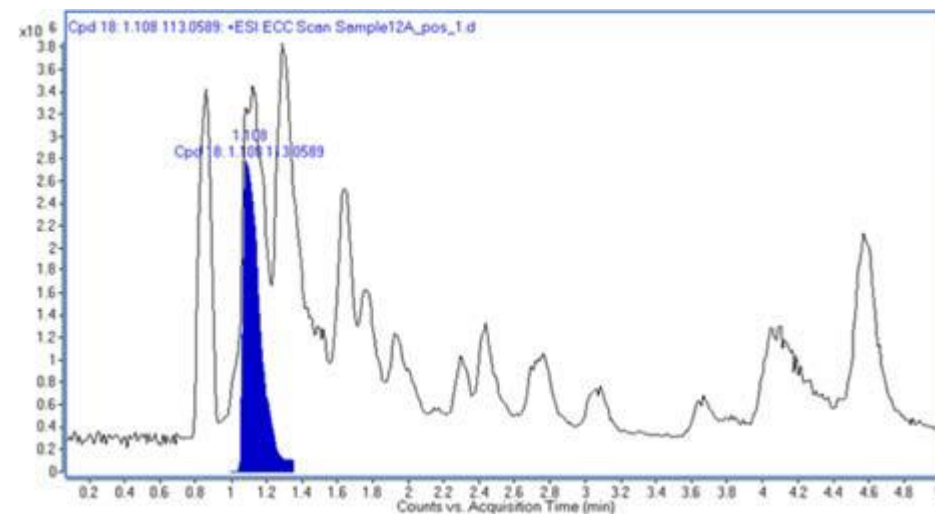
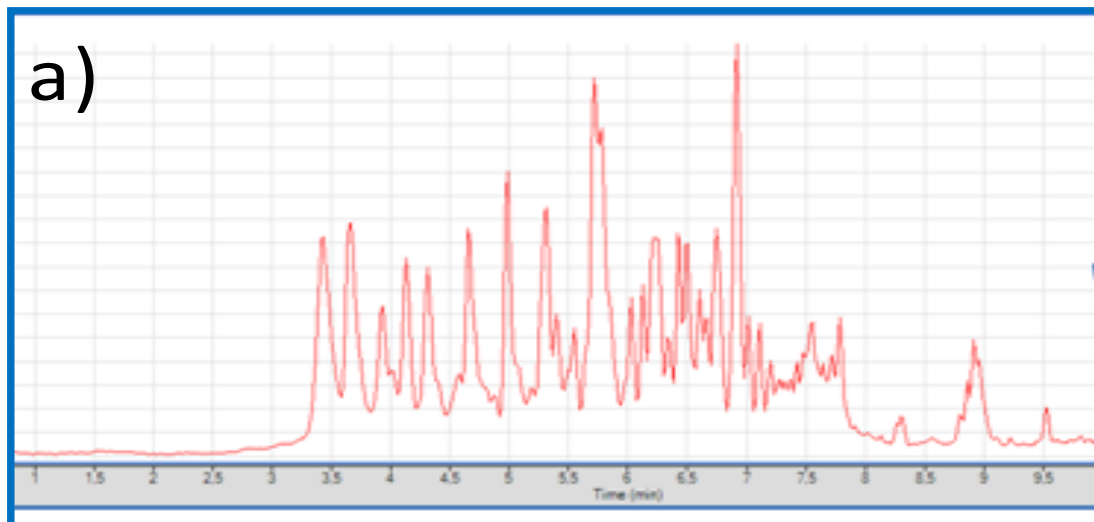
- a) przed i b) po procesie dekonwolucji  
(z <https://www.agilent.com/cs/library/Support/Documents/f05017.pdf>).

## Dekonwolucja widm w LC-ESI-MS i CE-ESI-MS

- Metody oparte na analizie pików
- Molecular Feature Extractor (Agilent) oceniają dokładność pomiarów masy w celu grupowania jonów związanych przez stan powłoki ładunku, rozkład izotopowy oraz możliwe relacje chemiczne przy określaniu, czy różne jony mają tę samą cechę metaboliczną.
- Można rozważyć, czy jony pokrewne tworzą addukty: protonowe, sodowe, potasowe i amonowe w trybie pozytywnej jonizacji lub tracą proton, tworząc addukty mrówczanowe itp. w trybie jonizacji ujemnej



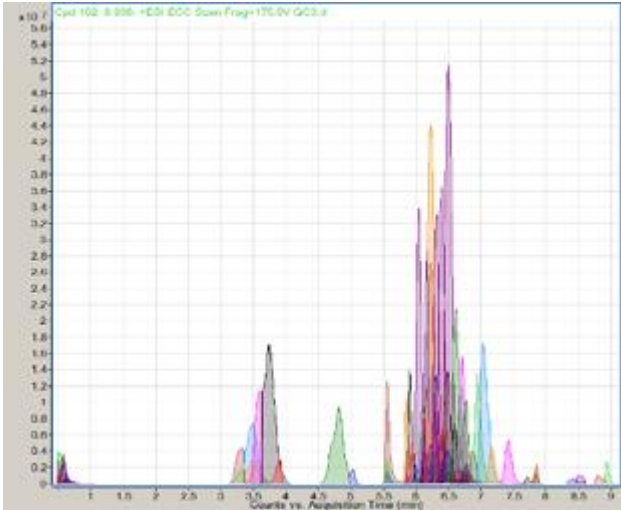
## Po dekonwolucji



a) całkowity prąd jonów

b) chromatogramy dla każdego pojedynczego związku otrzymanego po dekonwolucji

# Chromatogram lub lista cech?



Automatically Show Columns

Show/Hide	Saturated	RT	m/z	Mass	Polarity	Ions	Height	Area	Vol	Quality Score
<input checked="" type="checkbox"/>	S	0.514	280.0923	279.085	Positive	5	2740215		14234373	100
<input checked="" type="checkbox"/>	S	0.517	203.0526	202.0454	Positive	6	4045612		20286412	80
<input checked="" type="checkbox"/>		0.526	140.0682	139.0609	Positive	3	1530642		7519633	100
<input checked="" type="checkbox"/>		0.529	136.0482	135.041	Positive	2	1187260		6016909	100
<input checked="" type="checkbox"/>	S	0.57	162.1126	161.1053	Positive	7	2758926		19465836	100
<input checked="" type="checkbox"/>	S	0.57	304.2998	303.2926	Positive	4	3021606		14360408	100
<input checked="" type="checkbox"/>		0.58	114.0664	113.0591	Positive	4	549599		6985240	80
<input checked="" type="checkbox"/>		0.614	175.1192	174.1119	Positive	3	760396		3860452	91.6
<input checked="" type="checkbox"/>		0.625	156.0768	155.0696	Positive	7	1055485		6005360	100
<input checked="" type="checkbox"/>		0.646	170.0927	169.0854	Positive	4	604901		3355886	100
<input checked="" type="checkbox"/>	S	3.29	520.3389	519.3316	Positive	7	3243984		36560884	80.2
<input checked="" type="checkbox"/>		3.298	544.3388	543.3315	Positive	4	1339590		12508261	87
<input checked="" type="checkbox"/>	S	3.483	520.3389	519.3316	Positive	11	5382675		86088232	85.2
<input checked="" type="checkbox"/>	S	3.5	544.3389	543.3316	Positive	7	2216628		36501788	100
<input checked="" type="checkbox"/>	S	3.573	496.3389	495.3316	Positive	6	8281163		80624664	86.1
<input checked="" type="checkbox"/>	S	3.746	496.3388	495.3316	Positive	13	11742118		200075408	80
<input checked="" type="checkbox"/>	S	3.89	522.3545	521.3472	Positive	8	2204035		17763280	87
<input checked="" type="checkbox"/>	S	4.801	524.3702	523.3629	Positive	12	6862920		111854136	100
<input checked="" type="checkbox"/>		5.027	524.3702	523.3629	Positive	6	1308318		8748089	87
<input checked="" type="checkbox"/>		5.559	163.0393	162.032	Positive	4	1349239		5379156	100
<input checked="" type="checkbox"/>	S	5.561	391.2839	390.2767	Positive	15	7367381		51796644	100
<input checked="" type="checkbox"/>	S	5.561	149.0235	148.0162	Positive	4	2915086		12886771	100
<input checked="" type="checkbox"/>		5.597	338.3418	337.3345	Positive	9	787867		4058574	100
<input checked="" type="checkbox"/>	S	5.849	675.5425	674.5353	Positive	7	6821314		39496860	100
<input checked="" type="checkbox"/>		5.857	627.5339	626.5266	Positive	3	712763		3673685	100
<input checked="" type="checkbox"/>	S	5.904	701.5583	700.551	Positive	8	8615328		59468044	100
<input checked="" type="checkbox"/>	S	5.941	689.558	688.5507	Positive	8	3178377		18867332	100
<input checked="" type="checkbox"/>		5.991	754.5372	753.5299	Positive	4	737180		4127601	100
<input checked="" type="checkbox"/>		5.995	730.5360	729.5287	Positive	7	1506513		8184120	100

Chromatogram Results MS Spectrum

# Procesowanie danych

- **Wyrównanie**

- przesunięcia pików wzdłuż osi RT
- dwie grupy:
  - dane wyrównywane przed detekcją pików
  - metoda wyrównywania pików, gdy wykryte widmo naszego pików wyrównywane jest względem innych próbek.
- oprogramowanie:
  - MetaboAnalyst (metaboanalyst.ca)
    - mzmine and mzmine2 (<http://mzmine.sourceforge.net/>)
    - metAlign
    - BinBase (fiehnlab.ucdavis.edu)
    - xcms and xcms2 (Scripps)
    - metaXCMS (Scripps)
    - XCMS Online (Scripps)

- **Utracone dane**

- problemy w dalszej analizie
- różne strategie
  - zastąpienie przez wartość minimalną, przez średnią/medianę, metoda k -najbliższy sąsiad (kNN), probabilistyczna PCA (PPCA), metoda Bayesa PCA (BPCA), rozkład według wartości osobliwych (SVD) .....

- **Filtracja**

- zmienne o bardzo małych wartościach - wykrywane za pomocą średniej lub mediany
- zmienne, które są prawie stałe - wykrywane za pomocą odchylenia standardowego (SD)
- zmienne, które wykazują niską powtarzalność - wykrywane przy użyciu próbki QC

# Wstępna obróbka danych

- **Normalizacja**

- Normalizacja specyficzna dla próbki (tj. masa, objętość)
- Normalizacja za pomocą sumy lub mediany
- Normalizacja za pomocą próbki odniesienia
- Normalizacja przez spulowaną próbkę pochodzącą od grupy kontrolnej
- Normalizacja za pomocą funkcji odniesienia
- Normalizacja kwantyli

- **Transformacja danych**

- Transformacja logarytmiczna
- Transformacja cube root

- **Skalowanie danych**

- Średnia centrowana
- Automatyczne skalowanie (średnio-centrowane i podzielone przez odchylenie standardowe każdej zmiennej)
- Skalowanie Pareto (średnio-centrowane i podzielone przez pierwiastek kwadratowy standardowego odchylenia każdej zmiennej)
- Skalowanie zakresu (średnio-centrowane i podzielone przez zakres każdej zmiennej)

# Statystyka dla metabolomiki

## CELE:

- wykryć różnice pomiędzy grupami na poziomie chemicznym
- uszeregować związki według względnego znaczenia dla różnicowania próbek

## ZMIENNE

- **zmienna zależna**: reprezentuje wydajność lub efekt, lub jest testowana w celu sprawdzenia, czy istnieje efekt, np. : ilość metabolitu
- **zmienna niezależna**: reprezentuje moc wejścia lub przyczyny, lub testowane są, w celu zbadania, czy są jakieś przyczyny, np. : warunki leczenia w obrębie eksperymentu

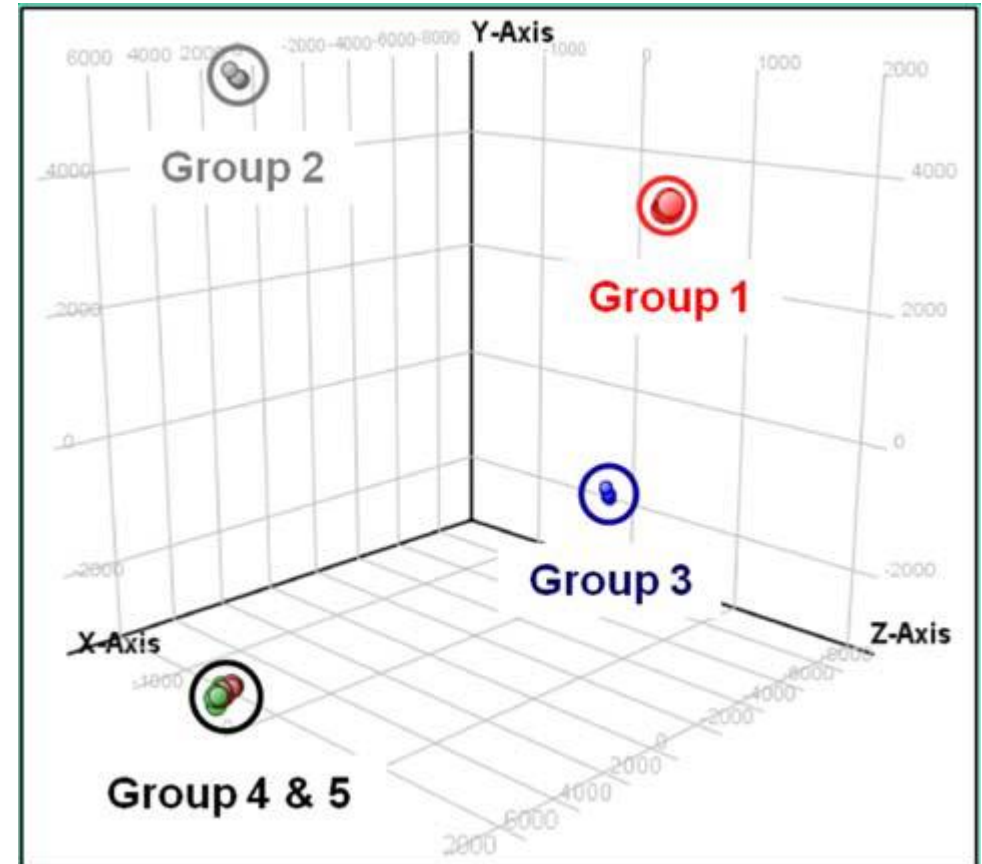
## TYPY

- **Jednowymiarowa analiza UVA:**
  - Rozkład normalny: *t*-test Studenta, ANOVA,
  - Rozkład nienormalny: Mann-Whitney U-Test, Kruskal-Wallis
- **Wielowymiarowa analiza MVA:** PCA, PLS-DA, OPLSDA



# PCA

- wykorzystywane jako narzędzie w analizie danych eksploracyjnych
- każda kropka graficznie przedstawia każdą zmierzoną próbkę
- algorytm nie ma wiedzy na temat przynależności próbek do grupy -analiza bez nadzoru
- pierwszy główny składnik wyjaśnia w najbardziej istotny sposób wariancję
- związki wskazują wpływ próbki w analizie
- każda kropka jest sumą związków dla próbki
- szczelność klasteringu ma wpływ na wariancję próbek



# Przewidywanie klas

algorytm wykorzystujący macierzyste dane w celu przewidywania wyników przyszłych obserwacji

- algorytm ma wiedzę na temat grup do których przynależą próbki - analiza nadzorowana
- Wspólne algorytmy
  - **Analiza dyskryminacyjna cząstkowych najmniejszych kwadratów (PLS-DA)**
  - Maszyna wektorów nośnych
  - Drzewo decyzyjne
  - Naïve Bayes
  - Sieć neuronowa

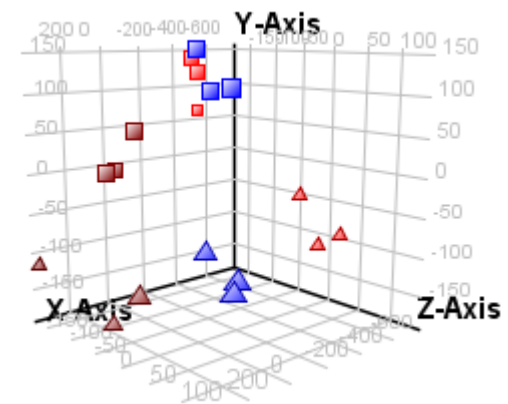
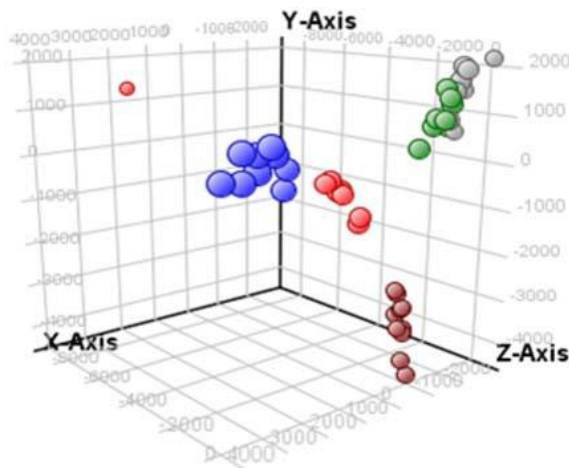


# Przewidywanie klas: PLS-DA

Analiza metodą cząstkowych najmniejszych kwadratów - Analiza dyskryminacyjna  
Projekcja na strukturach opóźnionych - Analiza dyskryminacyjna

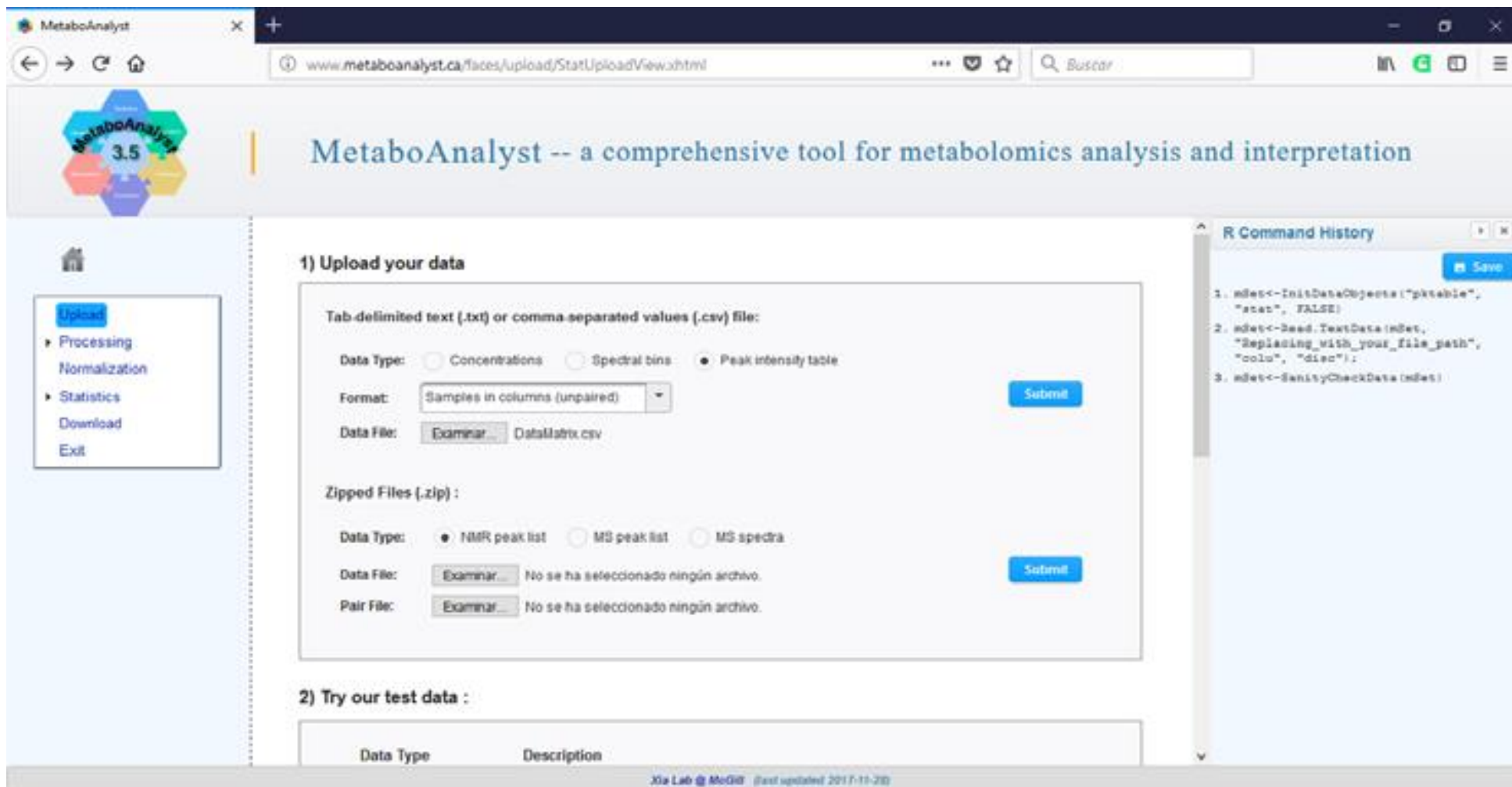
Metoda statystyczna, która ma pewien związek z analizą głównych składowych (PCA), ale jest analizą nadzorowaną [supervised analysis]

- tworzy model regresji liniowej, wyświetlając przewidywane i obserwowane zmienne w nowej przestrzeni
- dobrze dopasowana, gdy istnieje więcej predyktorów (związków) niż obserwacji (próbek)
- każdy związek ma t-wynik, który reprezentuje jego wpływ na prognozowanie
- wartość ufności prognozowania jest przypisywana, gdy model jest uruchamiany





# Jednowymiarowa i wielowymiarowa analiza statystyczna



The screenshot displays the MetaboAnalyst 3.5 web interface. The browser address bar shows the URL `www.metaboanalyst.ca/faces/upload/StatUploadView.xhtml`. The page title is "MetaboAnalyst -- a comprehensive tool for metabolomics analysis and interpretation".

**1) Upload your data**

Tab-delimited text (.txt) or comma-separated values (.csv) file:

Data Type:  Concentrations  Spectral bins  Peak intensity table

Format:

Data File:  DataMatrix.csv

**Zipped Files (.zip) :**

Data Type:  NMR peak list  MS peak list  MS spectra

Data File:  No se ha seleccionado ningún archivo.

Pair File:  No se ha seleccionado ningún archivo.

**2) Try our test data :**

Data Type	Description
-----------	-------------

**R Command History**

```
1. mdet<-InitDataObjects("pktable",  
"acet", FALSE)  
2. mdet<-Read.TextData(mdet,  
"Replacing_with_your_file_path",  
"col", "desc")  
3. mdet<-DensityCheckData(mdet)
```

Xia Lab @ McGill (last updated 2017-11-28)

## Przewidywanie klas: Walidacja modelu

ocenia dokładność tworzonej reguły predykcji, wskazuje na nadmierne dopasowane modele

### Zostaw jedną

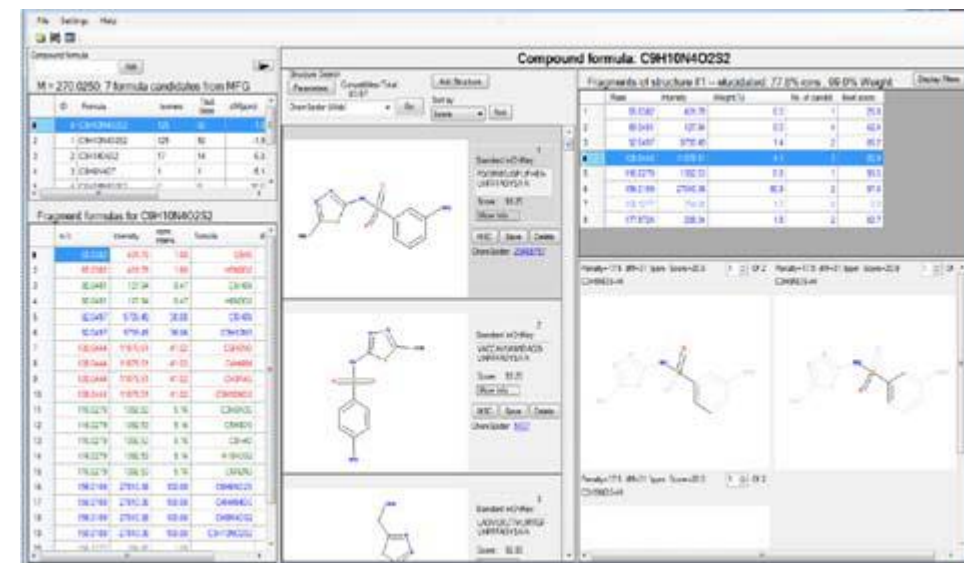
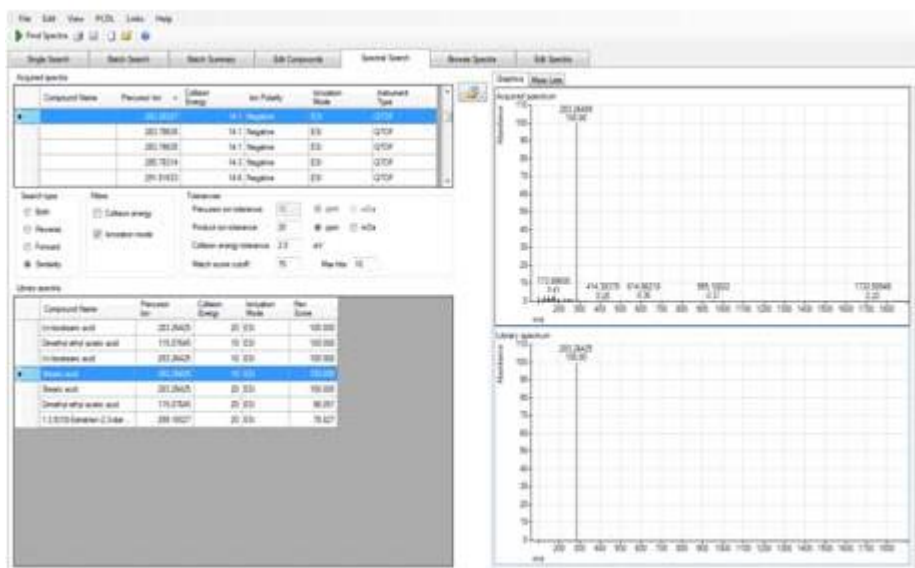
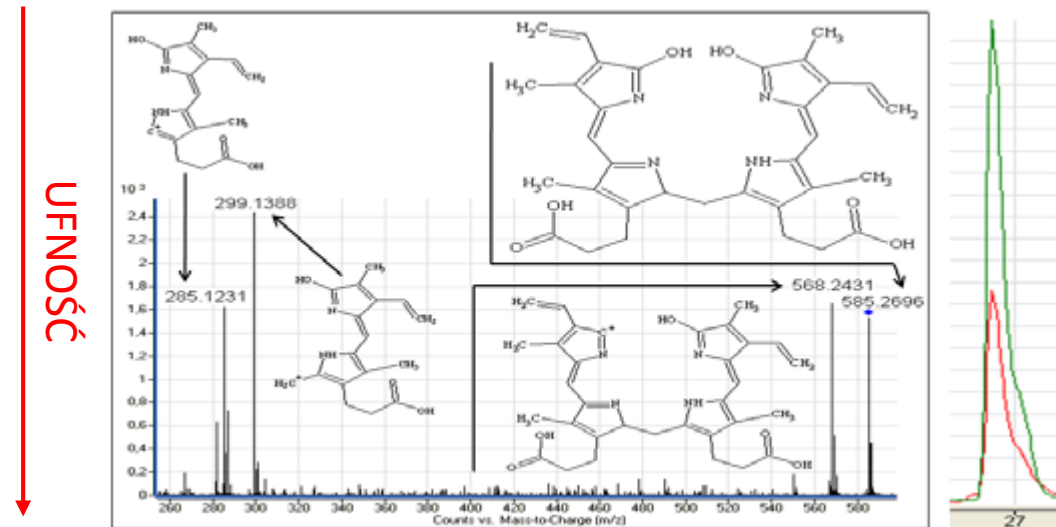
- wszystkie próbki w zbiorze treningowym, z wyjątkiem jednej używane są w celu zbudowania reguły przewidywania
- za pomocą tej reguły przewidywana jest klasa próbki, która została pominięta
- próbka zostaje zwrócona do zestawu treningowego, podczas gdy inna próbka jest pominięta i prognozowana reguła budowana jest z pozostałych próbek
- proces ten powtarza się, aż każda próbka w zbiorze treningowym zostanie przewidziana dokładnie raz
- liczba poprawnych i niepoprawnych prognoz jest następnie sprawdzana w celu określenia wskaźnika sukcesu

### N-krotnie

1. próbki w zbiorze treningowym są losowo dzielone na  $N$  równą podzbiorom, zachowując relatywną częstość klas
2. podzbiory  $N-1$  są następnie łączone dla treningu, a pozostały zestaw służy do testowania
3. powtórzyc krok 2 z pominięciem każdej grupy
4. powtórz krok 1, 2, 3  $M$  razy
5. każda próbka zostaje przewidziana  $M$  razy, a większość klas przewidywana jest  $M$  razy zgłaszanych w wynikach weryfikacji

# Identyfikacja

1. Bazy danych dopasowują związki za pomocą dokładnego pomiaru masy
2. Bazy danych dopasowują związki za pomocą **wzorca izotopowego**
3. Bazy danych dopasowują związki za pomocą wzorca izotopowego oraz **czasu retencji**
4. Dopasowanie za pomocą biblioteki MS/MS
5. Dopasowanie za pomocą biblioteki MS/MS oraz czasu retencji



# Klasyfikacja baz danych

- Na podstawie danych spektralnych
  - głównie małe cząsteczki, a nie tylko metabolity
  - NMR
  - MS lub MS/MS
- Na podstawie informacji o związku
  - nazwa związku, struktura, właściwości fizyczne, identyfikacja
- Na podstawie bazy danych szlaków metabolicznych
  - metabolity, ksenobiotyki, białka, szlaki sygnałowe
- kompletna baza danych metabolomicznych
  - kombinacja poprzednich

# Lista baz danych w 2018 roku

Name	URL	Name	URL
ARALIP	<a href="http://aralip.plantbiology.msu.edu/pathways/pathways">http://aralip.plantbiology.msu.edu/pathways/pathways</a>	KEGG	<a href="http://prime.psc.riken.jp/?action=metabolites_index">http://prime.psc.riken.jp/?action=metabolites_index</a>
AtIPD	<a href="http://www.atipd.ethz.ch/">http://www.atipd.ethz.ch/</a>	KEGG Glycan	<a href="http://www.genome.jp/kegg/glycan/">http://www.genome.jp/kegg/glycan/</a>
BiGG	<a href="http://bigg.ucsd.edu/">http://bigg.ucsd.edu/</a>	KNAPSAcK	<a href="http://prime.psc.riken.jp/?action=metabolites_index">http://prime.psc.riken.jp/?action=metabolites_index</a>
BioCyc	<a href="http://biocyc.org/">http://biocyc.org/</a>	LipidMaps	<a href="http://www.lipidmaps.org/">http://www.lipidmaps.org/</a>
BioNumbers	<a href="http://bionumbers.hms.harvard.edu/">http://bionumbers.hms.harvard.edu/</a>	MarkerDB	<a href="http://www.markerdb.ca/users/sign_in">http://www.markerdb.ca/users/sign_in</a>
BML-NMR	<a href="http://www.bml-nmr.org/">http://www.bml-nmr.org/</a>	MassBank	<a href="http://www.massbank.jp/">http://www.massbank.jp/</a>
BioMagResBank	<a href="http://www.bmrb.wisc.edu/metabolomics/">http://www.bmrb.wisc.edu/metabolomics/</a>	MetaboAnalyst	<a href="http://www.metaboanalyst.ca/MetaboAnalyst/">http://www.metaboanalyst.ca/MetaboAnalyst/</a>
BMDB	<a href="http://www.cowmetdb.ca/cgi-bin/browse.cgi">http://www.cowmetdb.ca/cgi-bin/browse.cgi</a>	MetaboLights	<a href="http://www.ebi.ac.uk/metabolights/index">http://www.ebi.ac.uk/metabolights/index</a>
ChEBI	<a href="http://www.ebi.ac.uk/chebi/">http://www.ebi.ac.uk/chebi/</a>	MetaCrop	<a href="http://metacrop.ipk-gatersleben.de/apex/f?p=269:111:">http://metacrop.ipk-gatersleben.de/apex/f?p=269:111:</a>
ChEMBL	<a href="https://www.ebi.ac.uk/chembl/about#">https://www.ebi.ac.uk/chembl/about#</a>	MetaCyc	<a href="http://metacyc.org/">http://metacyc.org/</a>
ChEBI	<a href="http://www.ebi.ac.uk/chebi/">http://www.ebi.ac.uk/chebi/</a>	METAGENE	<a href="http://www.metagene.de/program/a.prg">http://www.metagene.de/program/a.prg</a>
ChemMine	<a href="http://chemminedb.ucr.edu/">http://chemminedb.ucr.edu/</a>	METLIN	<a href="https://metlin.scripps.edu/index.php">https://metlin.scripps.edu/index.php</a>
ChemSpider	<a href="http://www.chemspider.com/">http://www.chemspider.com/</a>	MMCD	<a href="http://mmcd.nmrfam.wisc.edu/">http://mmcd.nmrfam.wisc.edu/</a>
CCD	<a href="http://ccd.chemnetbase.com/intro/index.jsp#about">http://ccd.chemnetbase.com/intro/index.jsp#about</a>	mzCloud	<a href="https://mzcloud.org/">https://mzcloud.org/</a>
CSF Metabolome Database	<a href="http://www.csfmetabolome.ca/">http://www.csfmetabolome.ca/</a>	OMIM	<a href="http://www.ncbi.nlm.nih.gov/omim/">http://www.ncbi.nlm.nih.gov/omim/</a>
CyberCell Database	<a href="http://ccdb.wishartlab.com/CCDB/">http://ccdb.wishartlab.com/CCDB/</a>	OMMBID	<a href="http://ommbid.mhmedical.com/">http://ommbid.mhmedical.com/</a>
DrugBank	<a href="http://www.drugbank.ca/">http://www.drugbank.ca/</a>	Oryzabase	<a href="http://www.shigen.nig.ac.jp/rice/oryzabase/">http://www.shigen.nig.ac.jp/rice/oryzabase/</a>
ECMDB	<a href="http://www.ecmdb.ca/">http://www.ecmdb.ca/</a>	PepBank	<a href="http://pepbank.mgh.harvard.edu/">http://pepbank.mgh.harvard.edu/</a>
ExPaSy Pathways	<a href="http://web.expasy.org/pathways/">http://web.expasy.org/pathways/</a>	PharmGKB	<a href="http://www.pharmgkb.org/">http://www.pharmgkb.org/</a>
Fiehn GC-MS Database	<a href="http://fiehnlab.ucdavis.edu/Metabolite-Library-2007/">http://fiehnlab.ucdavis.edu/Metabolite-Library-2007/</a>	PMN	<a href="http://www.plantcyc.org/">http://www.plantcyc.org/</a>
FooDB	<a href="http://www.foodb.ca">http://www.foodb.ca</a>	PubChem	<a href="http://pubchem.ncbi.nlm.nih.gov/">http://pubchem.ncbi.nlm.nih.gov/</a>
GMDB	<a href="http://gmd.mpimp-golm.mpg.de/">http://gmd.mpimp-golm.mpg.de/</a>	Reactome	<a href="http://www.reactome.org/">http://www.reactome.org/</a>
HMDB	<a href="http://metabolomics.pharm.uconn.edu/iimdb/">http://metabolomics.pharm.uconn.edu/iimdb/</a>	RiceCyc	<a href="http://pathway.gamene.org/gamene/ricecyc.shtml">http://pathway.gamene.org/gamene/ricecyc.shtml</a>
HumanCyc	<a href="http://www.genome.jp/kegg/">http://www.genome.jp/kegg/</a>	Serum Metabolome Database	<a href="http://www.serummetabolome.ca/">http://www.serummetabolome.ca/</a>
IIDMB	<a href="http://www.genome.jp/kegg/glycan/">http://www.genome.jp/kegg/glycan/</a>	SetupX & BinBase	<a href="http://fiehnlab.ucdavis.edu/projects/binbase_setupx">http://fiehnlab.ucdavis.edu/projects/binbase_setupx</a>

CEU

MASS MEDIATOR

Seleccionar archivo Ningún archivo seleccionado

**Experimental Masses (\*):**  
enter significant input masses

**Retention Times:**  
enter significant retention times

**Composite Spectra:**  
enter significant composite spectra

Seleccionar archivo Ningún archivo seleccionado

**All Experimental Masses:**  
enter all input masses

**All Retention Times:**  
enter all retention times

**All Composite Spectra:**  
enter all composite spectra

**Chemical Alphabet (\*):**  
All  
CHNOPS  
CHNOPS + Cl

**Modifiers (\*):**  
None  
NH3  
HCOO  
CH3COO  
HCOONH3  
CH3COONH3

**Databases (\*):**  
 All except MINE  
 All (Including In Silico Compounds)  
 Kegg  
 HMDB  
 LipidMaps  
 Metlin  
 MINE (Only In Silico Compounds)

**Metabolites (\*):**  
All except peptides  
Only lipids  
All including peptides

**Input Masses Mode (\*):**  
Neutral Masses  
m/z Masses

**Ionization Mode (\*):**  
Neutral  
Positive Mode  
Negative Mode  
calculation of new m/z from neutral mass based on selected adducts

**Adducts (\*):**  
 All  
 M+H  
 M+2H  
 M+Na  
 M+K  
 M+NH4

- Stosowanie adnotacji metabolitów.
- Wyszukuje ujednoliconych związków z różnych źródeł.
- Stosuje wiedzę na podstawie danych podanych przez użytkownika.
- Pomocny w identyfikacji utlenionych lipidów.
- <http://ceumass.eps.uspceu.es/mediator>

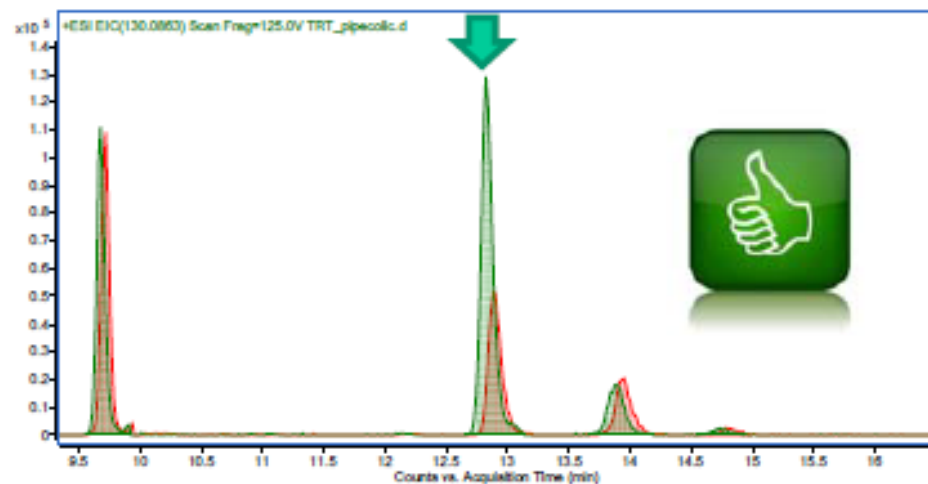
# CEU



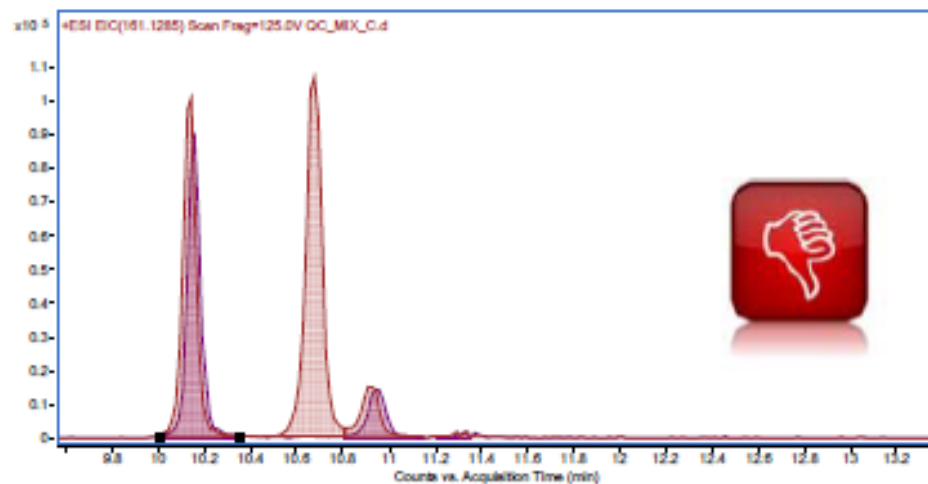
# MASS MEDIATOR

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
LIST OF COMPOUNDS	Experimental mass	Identifier	Adduct	PPM Error	Molecular Weight	Name	Formula	CAS	Kegg	HMDB	LipidMaps	Metlin	PubChem	InChIKey	Pathways
1	399.3367	17732	M-H	5	399.3349	L-palmitoylcarnitine	C23H45NO4				LMFA07070079			XOMRRQXKHMVMOC-NRFANRHFSA-N	
2	399.3367	17751	M-H	5	399.3349	O-palmitoylcarnitine	C23H45NO4				LMFA07070098			XOMRRQXKHMVMOC-UHFFFAOYSA-N	
3	399.3367	17657	M-H	5	399.3349	Palmitoylcarnitine	C23H45NO4	2364-67-2	C02390	HMDB0000222	LMFA07070004	36667	11953816	XOMRRQXKHMVMOC-OAQYLSRUS	Fatty acid Metabolism Fatty acid
4	399.3367	0	M-Na	0	0	No compounds found for experimental mass 399.3367 and adduct: I									
5	399.3367	13442	M-NH4	5	382.3083	methyl 9-butylperoxy-10,12-octadecadienoate	C23H42O4				LMFA01040036	74461		XVZNKFOXXKMPCH-IZAPIVEJSA-N	
6	399.3367	13443	M-NH4	5	382.3083	methyl 13-butylperoxy-9,11-octadecadienoate	C23H42O4				LMFA01040037	74462		GUWYTNFTPALBDJ-AEPWDTSSA-N	
7	399.3367	95769	M-NH4	5	382.3083	Lepidumterpenyl ester	C23H42O4			HMDB0036885		31868	11794320	PASMASQJCDKBJK-UHFFFAOYSA-N	
8	399.3367	53861	M-NH4	5	382.3083	MG(0:0/20:2(1I2,14Z)/0:0)	C23H42O4			HMDB0011544		62328	53480984	PMJSUEZTCFTBMD-HZJYTRNSA-N	
9	399.3367	52646	M-NH4	5	382.3083	MG(20:2(1I2,14Z)/0:0/0:0)	C23H42O4			HMDB0011574		62356	53480983	QRBGFYBOCBYOSN-KDTZXJSHSA-N	
10	399.3367	80631	M-NH4	5	382.3083	Persenone B	C23H42O4			HMDB0035955		91145		NLXNQLZUOMHEHB-ISLYRVAYSA-N	
11	399.3367	0	M-H-H2O	0	0	No compounds found for experimental mass 399.3367 and adduct: I									
12	421.3169	16368	M-H	6	421.3192	Gamma-linolenyl carnitine	C25H43NO4			HMDB0006318	LMFA07010893	58389	53477819	VDPMHVILMYSAGX-BAHSRKMSSA-N	
13	421.3169	96332	M-H	6	421.3192	Alpha-linolenyl carnitine	C25H43NO4			HMDB0006319		58390	53477821	DFVGGGHKDAHYIU-UHMZJXMFSA-N	
14	421.3169	126612	M-H	9	421.3205	AGELASINE	C26H39N5					43731			
15	421.3169	138401	M-H	9	421.313	Latanoprost ethyl amide-d4	C25H35D4NO4					96571			
16	421.3169	17732	M-Na	0	399.3349	L-palmitoylcarnitine	C23H45NO4				LMFA07070079			XOMRRQXKHMVMOC-NRFANRHFSA-N	
17	421.3169	17751	M-Na	0	399.3349	O-palmitoylcarnitine	C23H45NO4				LMFA07070098			XOMRRQXKHMVMOC-UHFFFAOYSA-N	
18	421.3169	17657	M-Na	0	399.3349	Palmitoylcarnitine	C23H45NO4	2364-67-2	C02390	HMDB0000222	LMFA07070004	36667	11953816	XOMRRQXKHMVMOC-OAQYLSRUS	Fatty acid Metabolism Fatty acid
19	421.3169	1294	M-NH4	6	404.2927	1alpha,25-dihydroxy-21-nor-22-oxavitamin D3 / 1alpha,25-dihydroxy-2-	C25H40O4				LMST03020029	41970		AOIMQZGQKVVZTNV-OQGZJSIESA-N	
20	421.3169	1295	M-NH4	6	404.2927	1alpha,25-dihydroxy-24-nor-22-oxavitamin D3 / 1alpha,25-dihydroxy-2-	C25H40O4				LMST03020030	41971		UMRLCGLUMINBCT-NPNXQCSOSA-N	
21	421.3169	2650	M-NH4	6	404.2927	7b-Hydroxy-3-oxo-5b-cholan-24-oate	C25H40O4				LMST04070028	57939		XHRLTYUHWGHDJ-JFRFJXPMSEA-N	
22	421.3169	117990	M-NH4	6	404.2927	Androstane-3,17-diol dipropionate,5alpha-Androstane-3alpha,17beta-	C25H40O4	4350-14-5	C14624			70213	134572	FWAYUWSHYLKUEY-LFYWYRBCSA-N	
23	421.3169	86607	M-NH4	6	404.2927	11'-Carboxy-gamma-chromanol	C25H40O4			HMDB0012517			53481453	ITULCXNDMOKAH-YYUODDRSA-N	
24	421.3169	89679	M-NH4	6	404.2927	MG(0:0/22:5(7Z,10Z,13Z,16Z,19Z)/0:0)	C25H40O4			HMDB0011556		62339	53480971	LRBJLYDLUADN-JLNMKQISIA-N	
25	421.3169	50261	M-NH4	6	404.2927	MG(22:5(4Z,7Z,10Z,13Z,16Z)/0:0/0:0)	C25H40O4			HMDB0011595		62367	53480993	HDIQCISTZKHUDD-AJWITYRPSA-N	
26	421.3169	105373	M-NH4	6	404.2927	MG(0:0/22:5(4Z,7Z,10Z,13Z,16Z)/0:0)	C25H40O4			HMDB0011555		62338	53480970	NPZWSBAEZZLYQU-WMPRHZDHSAN	
27	421.3169	56824	M-NH4	6	404.2927	MG(22:5(7Z,10Z,13Z,16Z,19Z)/0:0/0:0)	C25H40O4			HMDB0011586		62368	53480994	IDSLCYURGAOAO-YAWQMFZPFA-N	
28	421.3169	17695	M-H-H2O	5	439.3298	3-hydroxylinoleoylcarnitine	C25H45NO5				LMFA07070042			WQYXCASYXUFNSHUI-JQPVESSA-N	
29	315.2424	17712	M-H	5	315.241	Decanoylcarnitine	C17H33NO4			HMDB0000651	LMFA07070059		10245190	LZOSYCMHQXPBFU-UHFFFAOYSA-N	
30	315.2424	126591	M-H	5	315.241	L-Hexanoylcarnitine n-butyl ester	C17H33NO4					3549			
31	315.2424	17659	M-H	5	315.241	O-decanoyl-L-carnitine	C17H33NO4	3992-45-8	C03239	HMDB0062631	LMFA07070006	36669	11953821	LZOSYCMHQXPBFU-OAHLLOKOSA-N	
32	315.2424	0	M-Na	0	0	No compounds found for experimental mass 315.2424 and adduct: M									
33	315.2424	14277	M-NH4	5	298.2144	8E-Heptadecenedioic acid	C17H30O4				LMFA01170053	74925		VDTSYDDUGXHLDO-OWQJBTEDSA-N	
34	315.2424	111819	M-NH4	5	298.2144	Plakortin acid	C17H30O4			C17158		71590	10402441	ZCLJFHUIADAYRQ-CMDGGGBGSA-N	
35	315.2424	0	M-H-H2O	0	0	No compounds found for experimental mass 315.2424 and adduct: M									
36	337.2234	0	M-H	0	0	No compounds found for experimental mass 337.2234 and adduct: I									
37	337.2234	17712	M-Na	2	315.241	Decanoylcarnitine	C17H33NO4			HMDB0000651	LMFA07070059		10245190	LZOSYCMHQXPBFU-UHFFFAOYSA-N	
38	337.2234	126591	M-Na	2	315.241	L-Hexanoylcarnitine n-butyl ester	C17H33NO4					3549			
39	337.2234	17659	M-Na	2	315.241	O-decanoyl-L-carnitine	C17H33NO4	3992-45-8	C03239	HMDB0062631	LMFA07070006	36669	11953821	LZOSYCMHQXPBFU-OAHLLOKOSA-N	
40	337.2234	1060	M-NH4	6	320.1988	testolic acid	C19H28O4			C01618	LMST02020081	57817	439534	KMUJIFRPXRPT-DZBHQSCQSA-N	
41	337.2234	116962	M-NH4	6	320.1988	10beta-Hydroxy-6beta-isobutyrylfluranorempophilane	C19H28O4			C09685		57884	442377	W'BNAQLGGJMUDJ-BIGGFVEDSA-N	
42	337.2234	130809	M-NH4	6	320.1988	(+)-a-CMEHC	C19H28O4	7083-09-2				44834			
43	337.2234	53386	M-NH4	6	320.1988	[8]-Gingerdione	C19H28O4	77334-06-6		HMDB0039275		93864	14440537	QDSRAFNZQKMHPZ-UHFFFAOYSA-N	
44	337.2234	42685	M-NH4	6	320.1988	5'-Carboxy-alpha-chromanol	C19H28O4			HMDB0012738			53481524	QWPMI VBAEZ-IRMIHMFVZPKHSA-N	

## Potwierdzenie przez dodanie standardu



**kwas pipekolowy**

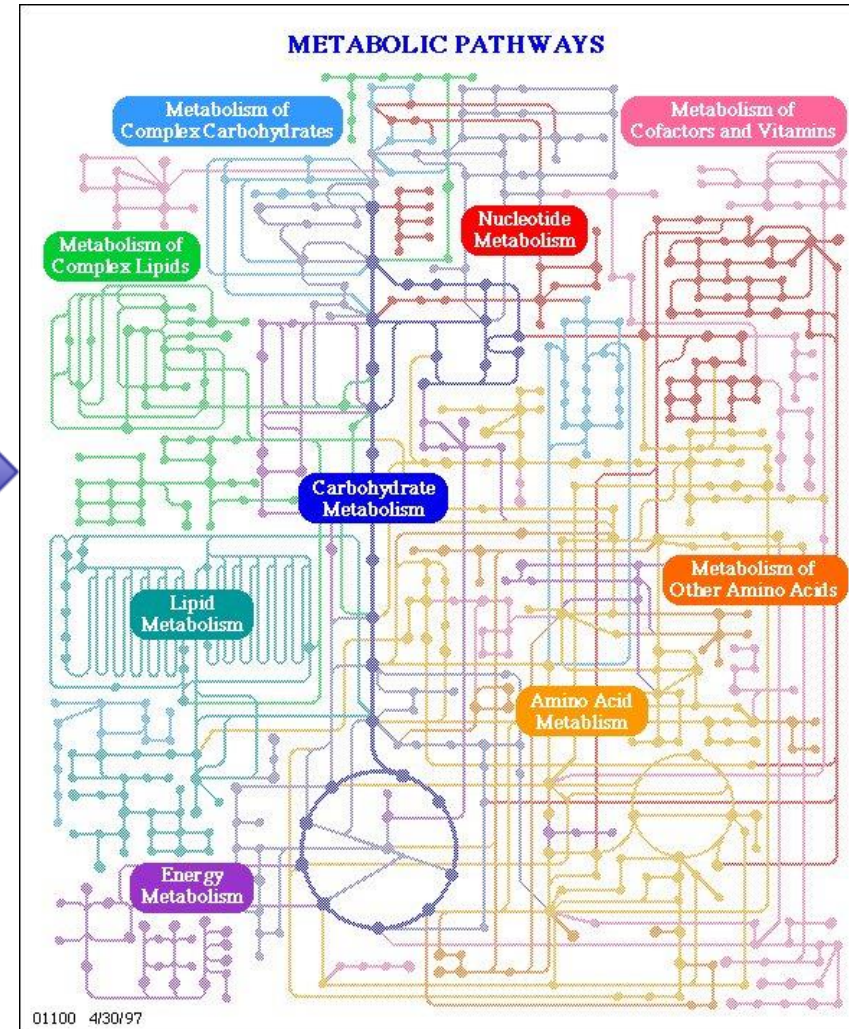


**metylo-lizyna**



# Od list do szlaków biochemicznych związków

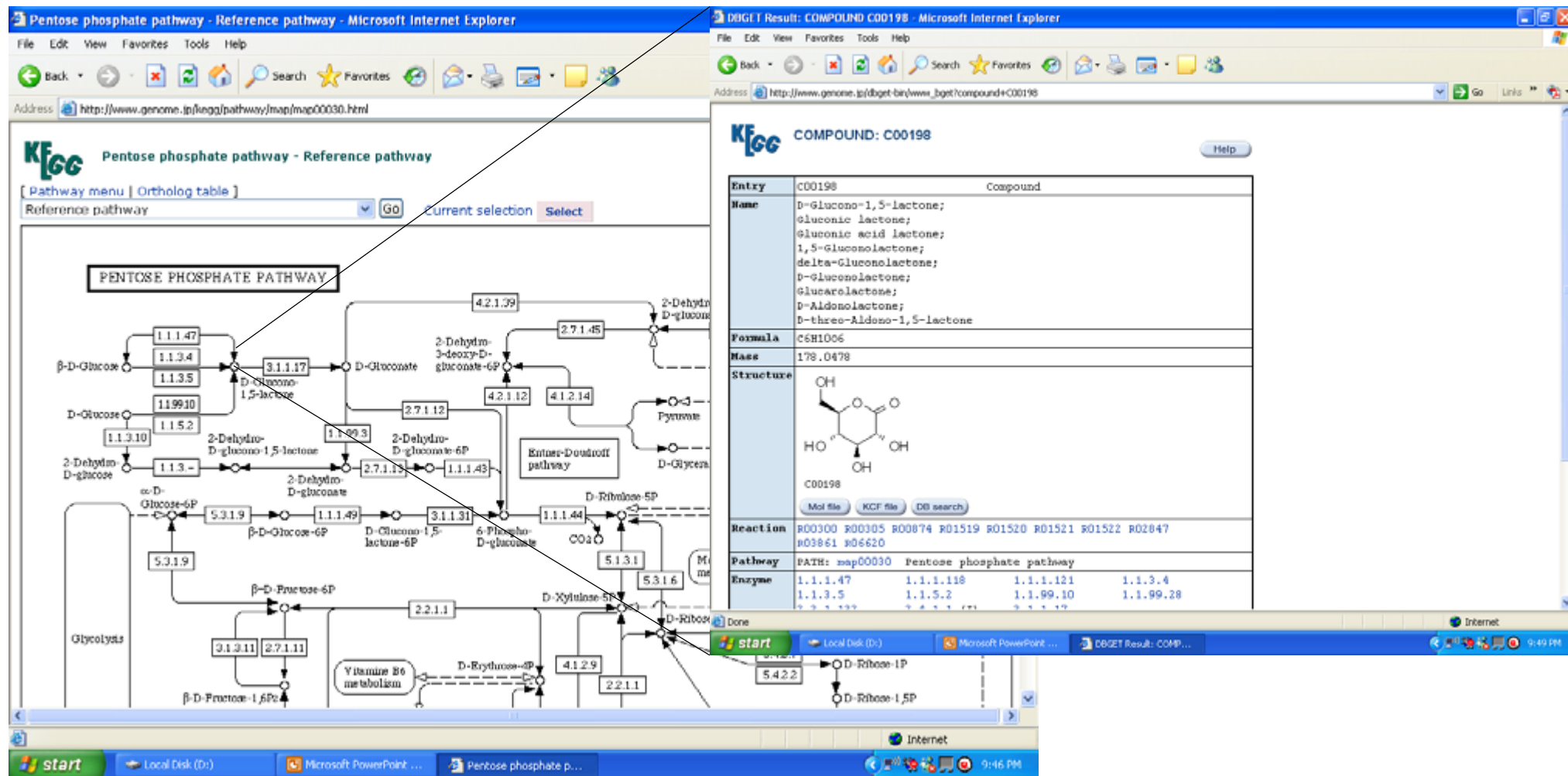
Compound	Retention Time (min)	Conc. in Urine (µM)	Compound	Retention Time (min)	Conc. in Urine (µM)
Dns-o-phospho -L-serine	0.92	<0.L*	Dns-Ile	6.35	25
Dns-o-phospho -L-tyrosine	0.95	<0.L	Dns-β-aminocaproic acid	6.44	0.5
Dns-achosine monophosphate	0.99	<0.L	Dns-pipecolic acid	6.50	0.5
Dns-o-phosphoethanolamine	1.06	16	Dns-Leu	6.54	54
Dns-glucosamine	1.06	22	Dns-cystathionine	6.54	0.3
Dns-o-phospho -L-threonine	1.09	<0.L	Dns-Leu-Pro	6.60	0.4
Dns-6-dimet. hylamine putre	1.20	<0.L	Dns-6-hydroxylysine	6.65	1.6
Dns-3-methyl -histidine	1.22	80	Dns-Cystine	6.73	160
Dns-aurine	1.25	834	Dns-N-norleusine	6.81	0.1
Dns-carnosine	1.34	28	Dns-5-hydroxydopamine	7.17	<0.L
Dns-Arg	1.53	36	Dns-dimethylamine	7.33	293
Dns-Asn	1.55	133	Dns-6-HIAA	7.46	18
Dns-hypoxanthine	1.58	10	Dns-umbelliferone	7.47	1.9
Dns-homocarnosine	1.61	3.9	Dns-2,3 -diaminopropanoic acid	7.63	<0.L
Dns-guanine	1.62	<0.L	Dns-L-ornithine	7.70	15
Dns-Gln	1.72	633	Dns-4-acetylamidophenol	7.73	51
Dns-atlantoin	1.83	3.8	Dns-proline	7.73	8.9
Dns-L-ornithine	1.87	2.9	Dns-homocysteine	7.76	3.3
Dns-1 (or 3 -)-methylhistamine	1.94	1.9	Dns-acetaminophen	7.97	62
Dns-adenosine	2.06	2.6	Dns-Phe-Phe	8.03	0.4
Dns-methylguanidine	2.20	<0.L	Dns-6-methoxy xycaloxyloic acid	8.04	2.1
Dns-βer	2.24	511	Dns-Lys	8.16	184
Dns-aspartic acid amide	2.44	26	Dns-aniline	8.17	<0.L
Dns-4-hydroxy -proline	2.56	2.3	Dns-leu -Phe	8.22	0.3
Dns-Glu	2.57	21	Dns-His	8.35	1550
Dns-Asp	2.60	90	Dns-4-thiastine	8.37	<0.L
Dns-Thr	3.03	157	Dns-benzylamine	8.38	<0.L
Dns-ephedrine	3.05	<0.L	Dns-1-ephedrine	8.50	0.6
Dns-ethanolamine	3.11	471	Dns-tryptamine	8.63	0.4
Dns-aminoadipic acid	3.17	70	Dns-pyridoxamine	8.94	<0.L
Dns-Gly	3.43	2510	Dns-2-methyl -benzylamine	9.34	<0.L
Dns-Ala	3.88	593	Dns-6-hydroxytryptophan	9.25	0.12
Dns-aminolevulinic acid	3.97	30	Dns-1,3 -diaminopropane	9.44	0.23
Dns-r-amino -butyric acid	3.98	4.6	Dns-putrescine	9.60	0.5
Dns-p-amino -hippuric acid	3.98	2.9	Dns-1,2 -diaminopropane	9.66	0.1
Dns-5-hydroxymethyluracil	4.58	1.9	Dns-lysineamide	9.79	29
Dns-tryptophanamide	4.70	5.5	Dns-dopamine	10.08	140
Dns-boguanine	4.75	<0.L	Dns-oadaverine	10.08	0.08
Dns-6-aminopenicillanic acid	4.79	1.6	Dns-histamine	10.19	0.4
Dns-carosidine	4.81	7.2	Dns-β-methoxy -tyramine	10.19	9.2
Dns-β-amino -isobutyrate	4.81	85	Dns-Tyr	10.28	321
Dns-2-aminobutyric acid	4.91	17	Dns-cysteamine	10.44	<0.L



## Bazy danych szlaków biochemicznych

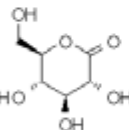
- Bogate źródło danych biologicznych, które dotyczą metabolitów genów, białek, chorób, sygnalizacyjnych zdarzeń i procesów
- Zapewnienie różnych narzędzi umożliwiających wizualizację i mapowanie genów/metabolitów
- Często obejmują wiele gatunków
- KEGG ([www.genome.jp/kegg/](http://www.genome.jp/kegg/)), BioCyc / MetaCyc (<https://biocyc.org/>), SMPDB ([www.smpdb.ca](http://www.smpdb.ca)), Reactome ([www.reactome.org](http://www.reactome.org)), WikiPathways (<http://www.wikipathways.org>) ...
- "Ściśle mówiąc, można argumentować, że szlaki nie istnieją ... są tylko sieci. "(WikiPathways.org)

# KEGG – Encyklopedia Kyoto Genów i Genomów



The image displays two browser windows from the KEGG database. The left window shows the 'Pentose phosphate pathway - Reference pathway' map, which is a complex network of biochemical reactions involving various sugars and enzymes. A box labeled 'PENTOSE PHOSPHATE PATHWAY' highlights a specific section of the map. The right window shows the 'DBGET Result: COMPOUND C00198' page, which provides detailed information about the compound D-Glucono-1,5-lactone.

**COMPOUND: C00198**

Entry	C00198	Compound
Name	D-Glucono-1,5-lactone; gluconic lactone; gluconic acid lactone; 1,5-Gluconolactone; delta-Gluconolactone; D-Gluconolactone; Gluconolactone; D-Aldonolactone; D-three-Aldono-1,5-lactone	
Formula	C6H10O6	
Mass	178.0478	
Structure	 C00198 <a href="#">Mol file</a> <a href="#">KCF file</a> <a href="#">DB search</a>	
Reaction	R00300 R00305 R00874 R01519 R01520 R01521 R01522 R02847 R03861 R06620	
Pathway	PATH: map00030 Pentose phosphate pathway	
Enzyme	1.1.1.47 1.1.1.118 1.1.1.121 1.1.3.4 1.1.3.5 1.1.5.2 1.1.99.10 1.1.99.28 1.1.1.117 1.1.1.119 1.1.1.120 1.1.1.122	

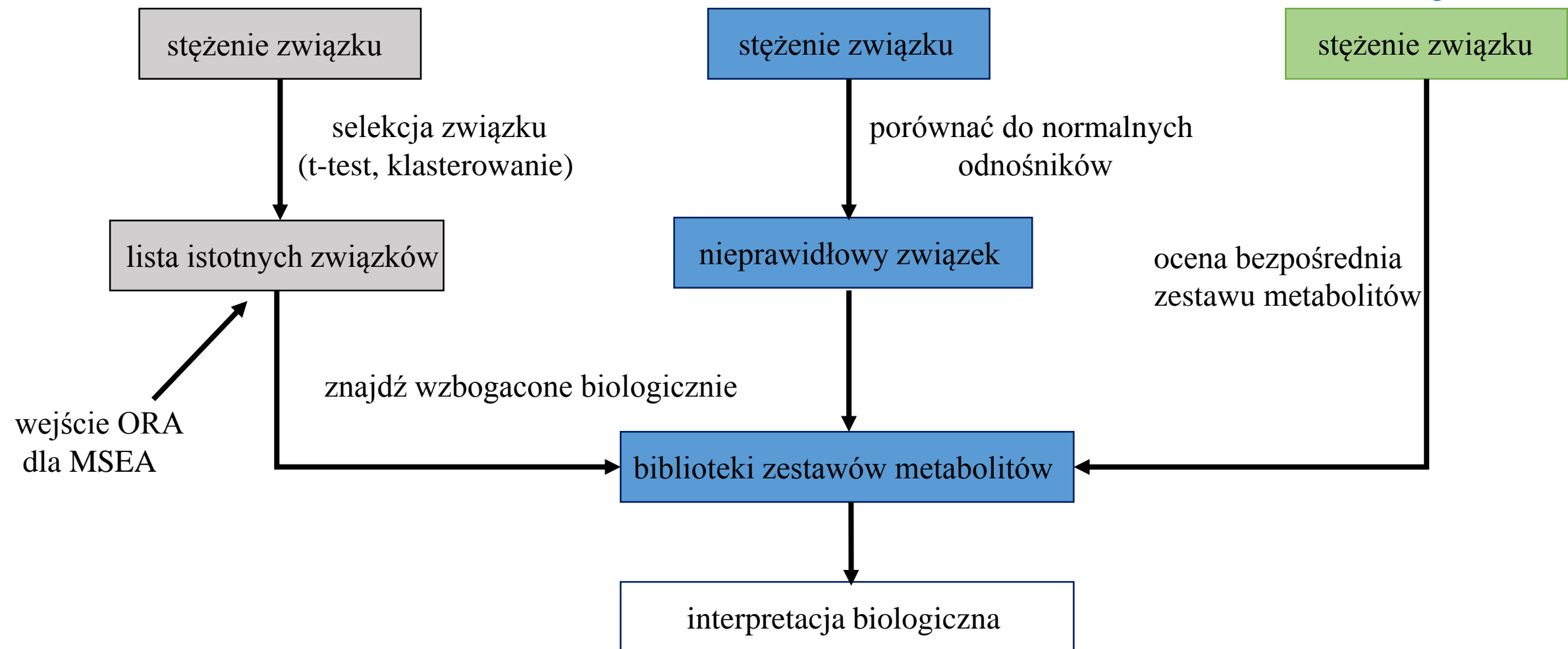
<http://www.genome.jp/kegg/>

# Analiza wzbogacania, podejście MSEA

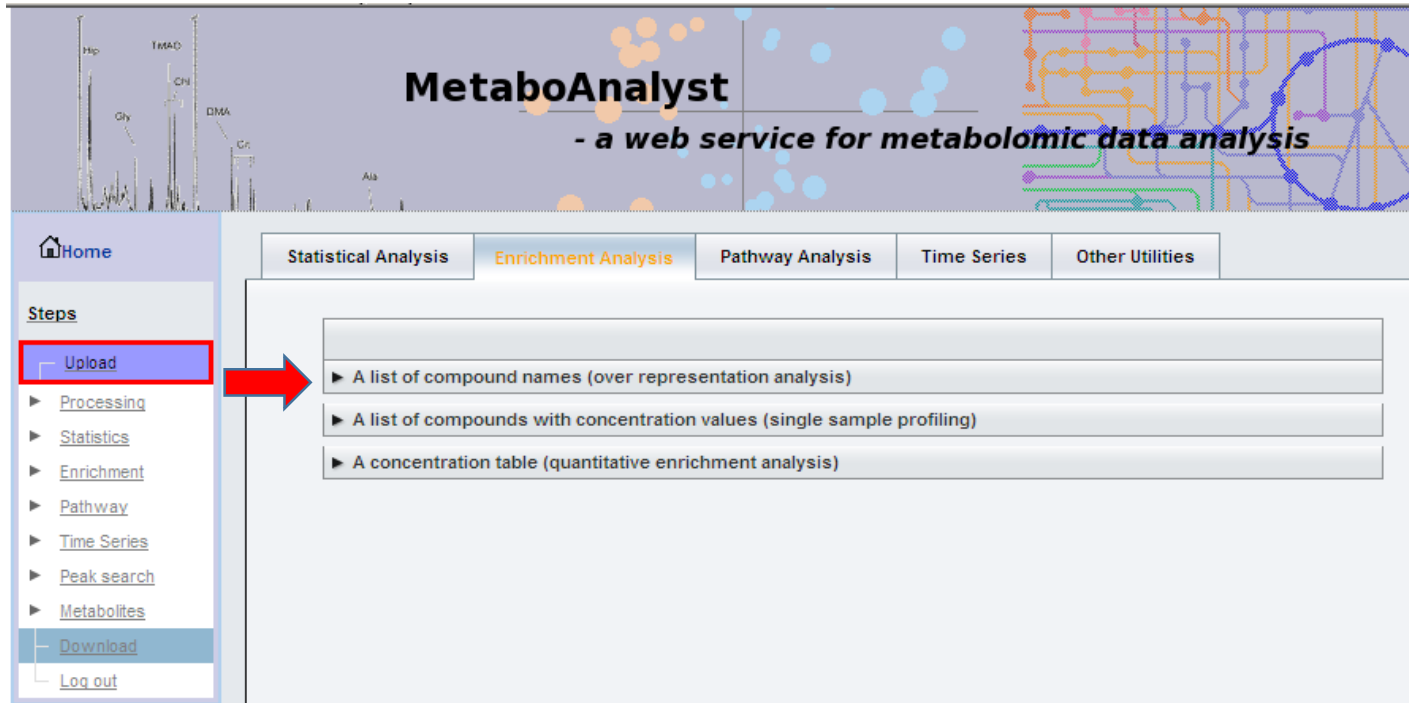
## Analiza reprezentatywna

## Profilowanie pojedynczych próbek

## Analiza ilościowego wzbogacania



# Rozpocznij od listy związków



**MetaboAnalyst**  
- a web service for metabolomic data analysis

Home | Statistical Analysis | **Enrichment Analysis** | Pathway Analysis | Time Series | Other Utilities

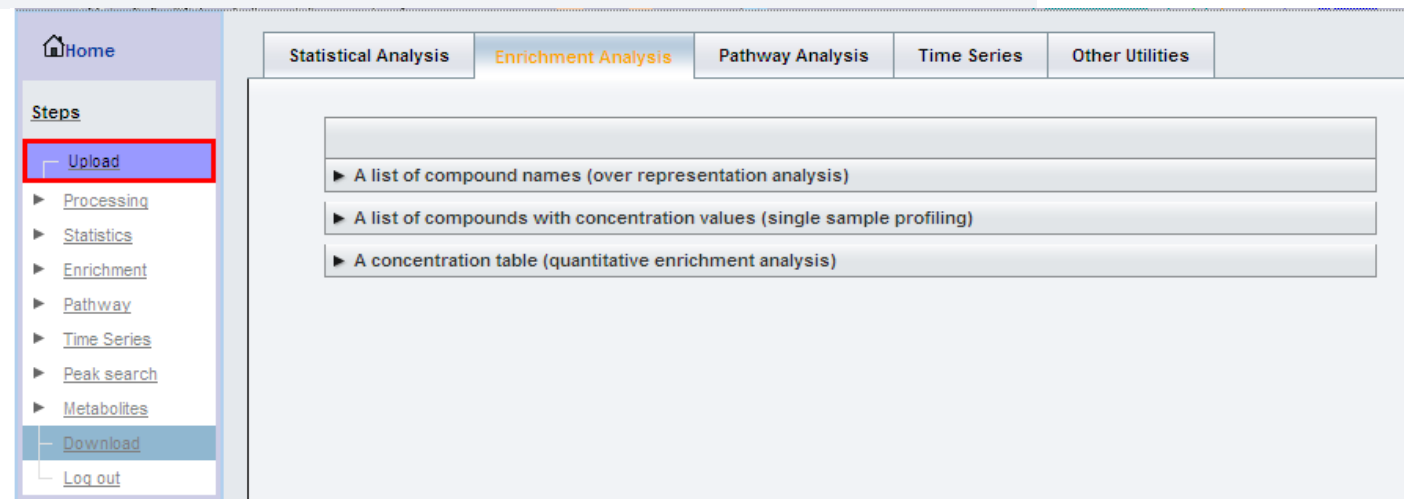
**Steps**

- Upload**
- Processing
- Statistics
- Enrichment
- Pathway
- Time Series
- Peak search
- Metabolites
- Download
- Log out

▶ A list of compound names (over representation analysis)

▶ A list of compounds with concentration values (single sample profiling)

▶ A concentration table (quantitative enrichment analysis)



Home | Statistical Analysis | **Enrichment Analysis** | Pathway Analysis | Time Series | Other Utilities

**Steps**

- Upload**
- Processing
- Statistics
- Enrichment
- Pathway
- Time Series
- Peak search
- Metabolites
- Download
- Log out

▶ A list of compound names (over representation analysis)

▶ A list of compounds with concentration values (single sample profiling)

▶ A concentration table (quantitative enrichment analysis)






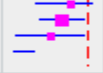



Home


Steps

- Upload
- Processing
  - Pre-process
  - Name check
  - Conc. check**
  - Data check
  - Missing value
  - Data filter
  - Data editor
  - Color picker
  - Normalization
- Statistics
- Enrichment
- Pathway
- Time Series
- Peak search
- Metabolites
- Download
- Log out

## Comparison with Reference Concentration

Note: *reference concentrations* are in the form of **mean(min - max)** format. In cases where the ranges were not reported in the original literature, the min and max were calculated using the 95% confidence intervals. In the *Comparison* column, **H, M, L** means **higher, medium (within range), lower** compared to the reference concentrations. Click the **Image Icon** link to see a graphical summary for the comparisons.

Compound	Concentration	Reference Concentrations	Comparison	Detail	Include
L-Isoleucine	0.34	1.579 (0.789 - 2.368); 0.94 (0.27 - 1.61); 3.75 (1 - 6.5); 3 (1.5 - 4.5); 1.8 (0.8 - 2.8)	M		<input type="checkbox"/>
Fumaric acid	0.47	10.4 (2.8 - 53.7); 0.5 (0.1 - 1.7); 1 (0 - 2); 0.95 (0.02 - 1.88); 0.8 (0.1 - 1.7); 10.7 (0.1 - 28.2); 4.8 (0 - 35.2); 5 (1 - 33.5)	M		<input type="checkbox"/>
Acetone	0.58	4.2 (0.98 - 15.3); 0.92 (0.2 - 2.8); 320 (103 - 1290); 20 (2 - 180); 15.3 (2 - 120)	M		<input type="checkbox"/>
Succinic acid	9.4	14.4 (9.5 - 19.3); 3.8 (1.25 - 6.7); 12.6 (0.47 - 24.73); 14.48 (11.28 - 17.68); 9.9 (4.9 - 14.9); 39 (37 - 41); 197.2 (29.4 - 486.2); 185.4 (6 - 342.6); 7.7 (1.9 - 20); 11.6 (4 - 27.3); 8.25 (0.5 - 16)	M		<input type="checkbox"/>
1-Methylhistidine	9.6	2.3 (0 - 7.4); 33.6 (0 - 70); 28.1 (0 - 59.9); 30 (0 - 73); 45.5 (3.9 - 87.1); 1.3 (0 - 4.06); 4.6 (1.9 - 7.3); 46.1 (0 - 99.6); 15.9 (0 - 35.4)	M		<input type="checkbox"/>
L-Asparagine	19.62	35 (16.4 - 57.2); 9.211 (3.289 - 15.1); 0.96 (0.31 - 1.61); 10 (4.6 - 16.32)	M		<input type="checkbox"/>
3-Methylhistidine	9.7	42.76 (19.92 - 65.6); 15.1 (3.9 - 26.3); 12.5 (8.3 - 16.7)	M		<input type="checkbox"/>
L-Threonine	93.19	36.2 (10.82 - 61.58); 12.7 (4.934 - 20.4); 1 (0.16 - 2.4); 4.9 (2.4 - 7.4); 16 (7 - 25); 18 (8.4 - 27.6)	H		<input checked="" type="checkbox"/>
Creatine	720	46 (9 - 135); 113 (0 - 654); 26 (5 - 95); 167 (124 - 210); 212 (0 - 5000); 450 (0 - 10000)	M		<input type="checkbox"/>



Home

Statistical Analysis **Enrichment Analysis** Pathway Analysis Time Series Other Utilities

Steps

- Upload**
- Processing
- Statistics
- Enrichment
- Pathway
- Time Series
- Peak search
- Metabolites
- Download
- Log out

► A list of compound names (over representation analysis)

► A list of compounds with concentration values (single sample profiling)

▼ A concentration table (quantitative enrichment analysis)

### Upload your concentration data (.csv)

[Format](#)

**Compound Label Type:**

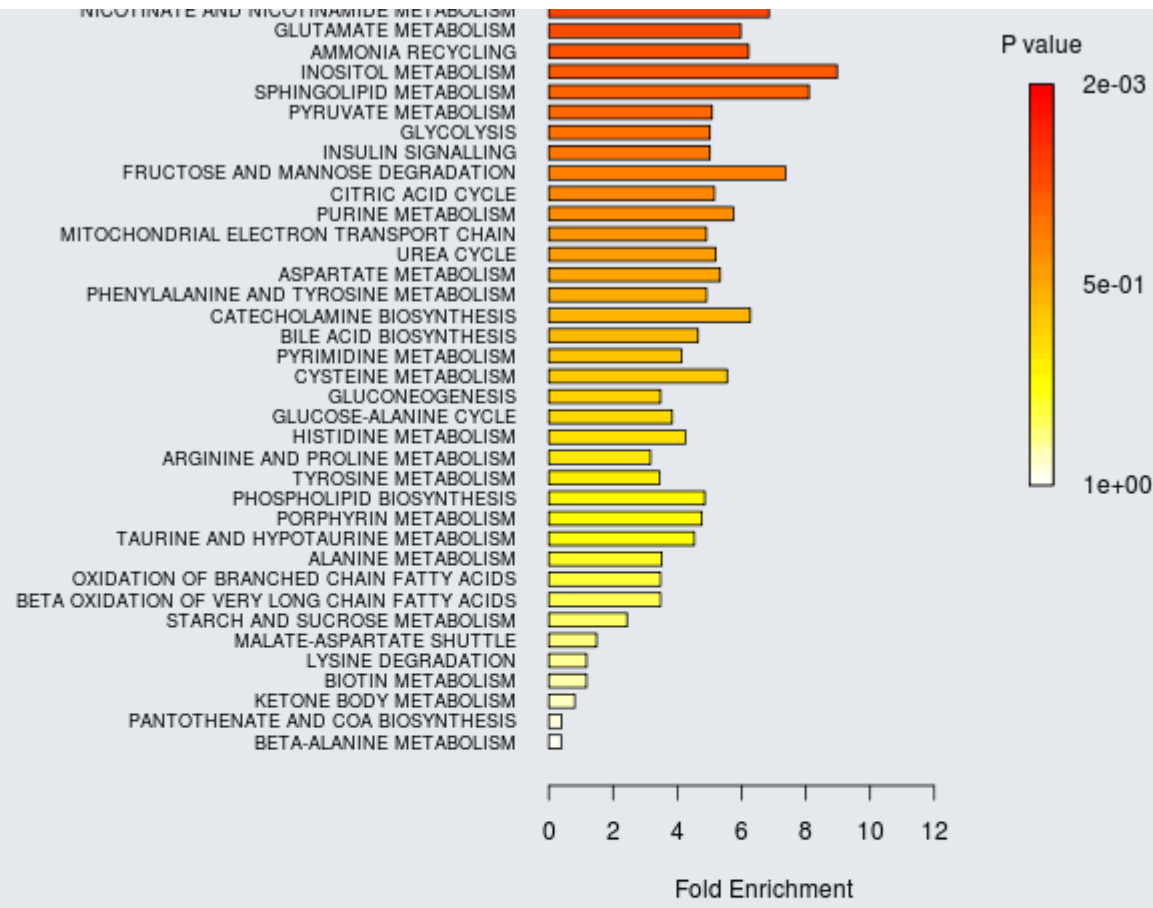
**Phenotype Label:**





---

**Try our test data:**

Data	Compound	Phenotype	Description
<input checked="" type="radio"/> <a href="#">Data 1</a>	Common name	Discrete	Urinary metabolite concentrations from 77 cancer patients measured by 1H NMR. Phenotype: <b>N</b> - cachexic; <b>Y</b> - control
<input type="radio"/> <a href="#">Data 2</a>	PubChem CID	Continuous	Urinary metabolite concentrations from 97 cancer patients measured by 1H NMR. Phenotype: <b>muscle gain</b> (percentage within 300 days, negative values indicate muscle loss)

- ▶ [Time Series](#)
- ▶ [Peak search](#)
- ▶ [Metabolites](#)
- └─ [Download](#)
- └─ [Log out](#)



Metabolite Set	Total	Hit	Statistic	Expected	P Value	Holm P	FDR	Details
TRYPTOPHAN METABOLISM	34	2	15.088	1.3158	5.3712E-5	0.0024707	0.0020529	
PROPANOATE METABOLISM	18	1	17.695	1.3158	1.3942E-4	0.0062741	0.0020529	
BETAINE METABOLISM	10	2	14.311	1.3158	1.4515E-4	0.0063865	0.0020529	
METHIONINE METABOLISM	24	4	11.386	1.3158	1.7852E-4	0.0076762	0.0020529	



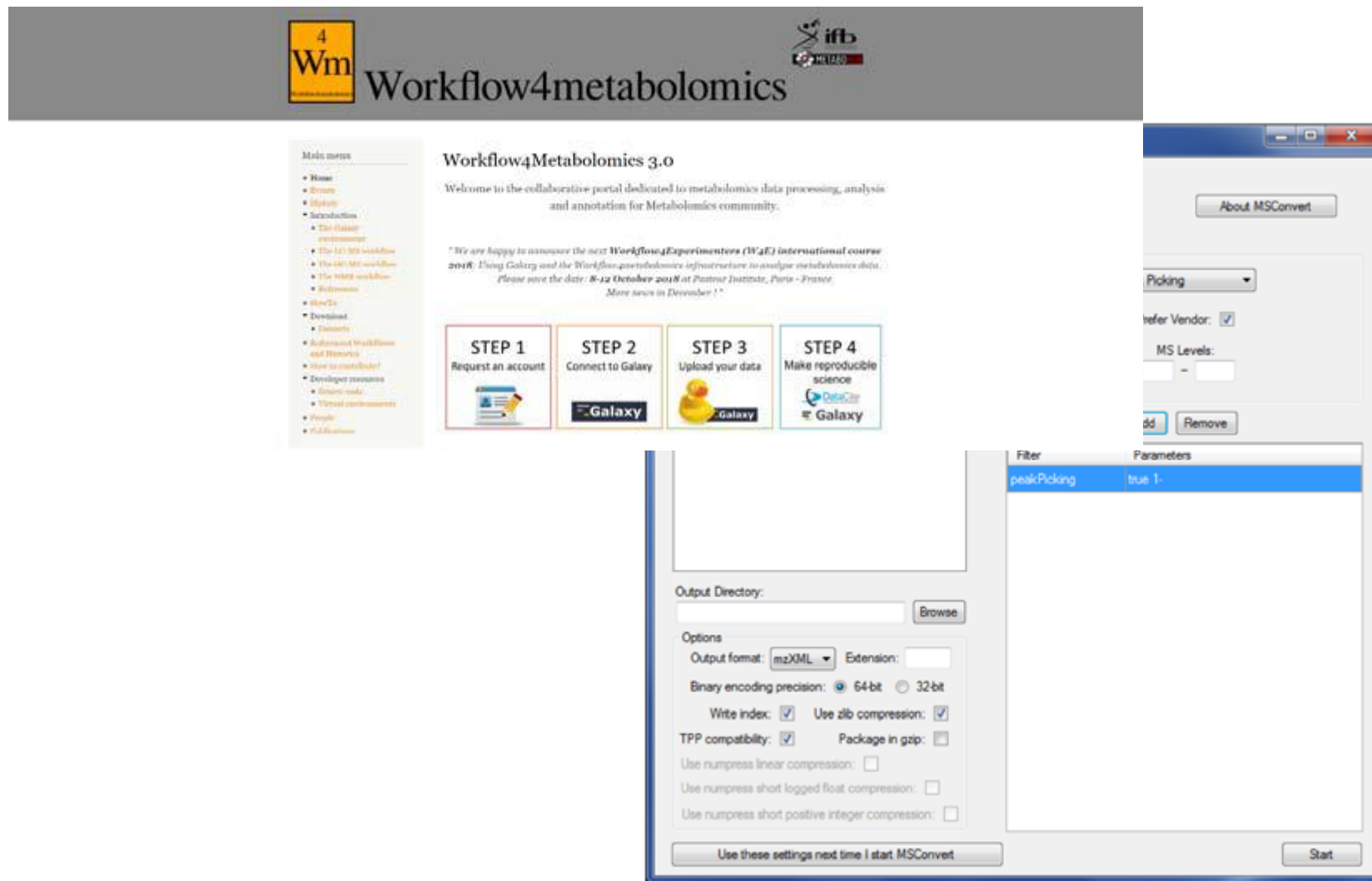


# Metaboanalitik

## Analiza szlaku metabolicznego (MetPA)

- Cel: rozszerzenie i wzmocnienie analizy wzbogacania zestawu metabolitów dla szlaków przez
  - Biorąc pod uwagę struktury szlaku
  - Wizualizację dynamiczną szlaku
- Obecnie obejmuje ~ 1500 szlaków 17 organizmów (w oparciu o KEGG)

# Zajęcia praktyczne. Wizualizacja\_1



The image displays the Workflow4Metabolomics 3.0 website and its MSConvert interface. The website header features the '4 Wm' logo and the 'ifb' logo. The main content area includes a 'Main menu' on the left, a welcome message, and a four-step process: STEP 1 (Request an account), STEP 2 (Connect to Galaxy), STEP 3 (Upload your data), and STEP 4 (Make reproducible science). The MSConvert interface is shown in the foreground, featuring a 'Filter' section with 'peak-Picking' selected, an 'Output Directory' field, and various options for output format (mzXML), binary encoding precision (64-bit), and compression settings. A 'Start' button is visible at the bottom right of the MSConvert window.

**4 Wm** Workflow4metabolomics

**ifb**

## Workflow4Metabolomics 3.0

Welcome to the collaborative portal dedicated to metabolomics data processing, analysis and annotation for Metabolomics community.

*"We are happy to announce the next **Workflow4Experimenters (W4E) international course 2018: Using Galaxy and the Workflow4metabolomics infrastructure to analyse metabolomics data. Please save the date: 8-12 October 2018 at Pasteur Institute, Paris - France. More news in December!"***

**STEP 1**  
Request an account

**STEP 2**  
Connect to Galaxy

**STEP 3**  
Upload your data

**STEP 4**  
Make reproducible science

**MSConvert**

About MSConvert

Filter: peak-Picking

Parameters: true 1-

Output Directory: [ ] Browse

Options

Output format: mzXML Extension: [ ]

Binary encoding precision:  64-bit  32-bit

Write index:  Use zlib compression:

TPP compatibility:  Package in gzip:

Use numpress linear compression:

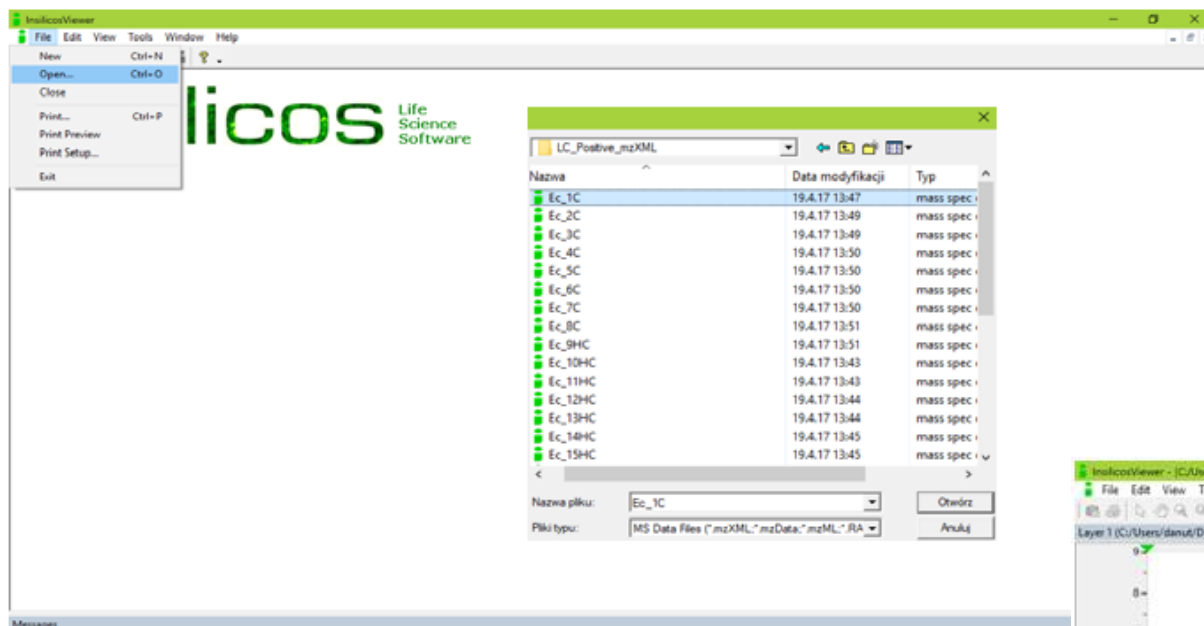
Use numpress short logged float compression:

Use numpress short positive integer compression:

Use these settings next time I start MSConvert

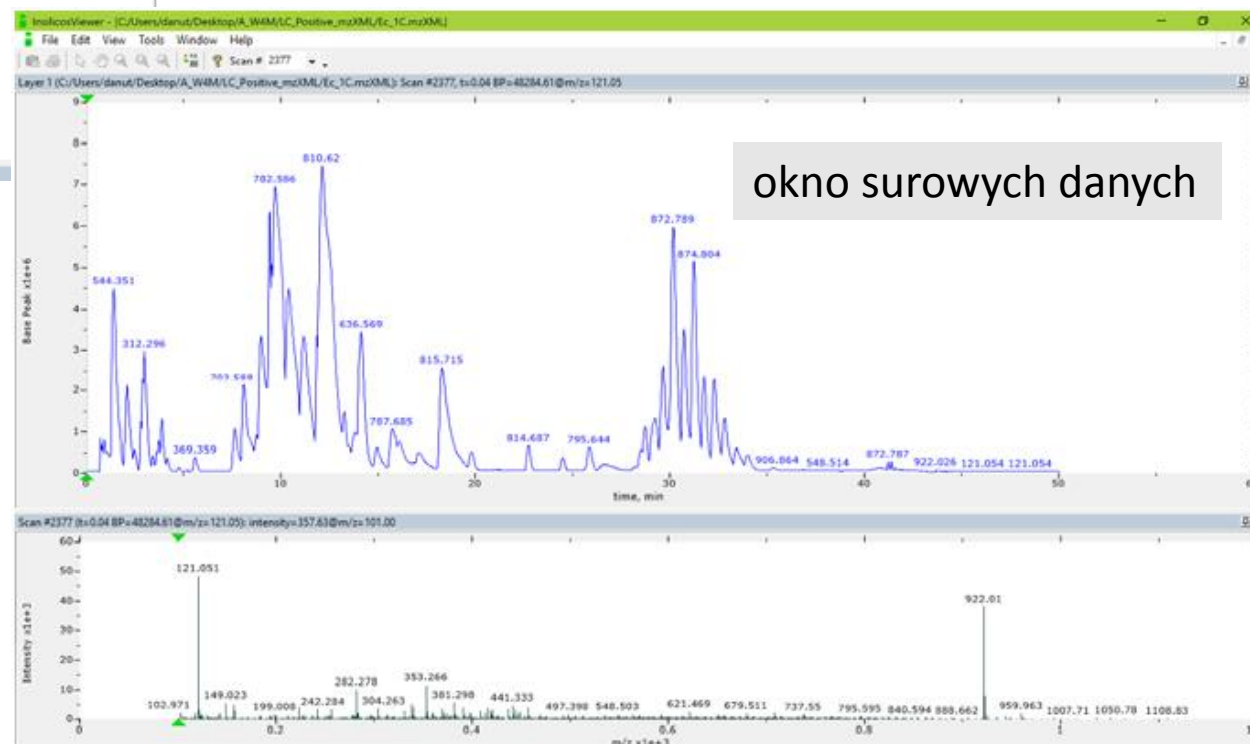
Start

# Zajęcia praktyczne. Wizualizacja\_2



The screenshot shows the licos Life Science Software interface. A menu is open, displaying options: New (Ctrl+N), Open... (Ctrl+O), Close, Print... (Ctrl+P), Print Preview, Print Setup..., and Exit. The main window displays a list of files under the name 'LC\_Positive\_mzXML'. The files are listed in a table with columns for 'Nazwa', 'Data modyfikacji', and 'Typ'. The file 'Ec\_1C' is selected.

Nazwa	Data modyfikacji	Typ
Ec_1C	19.4.17 13:47	mass spec
Ec_2C	19.4.17 13:49	mass spec
Ec_3C	19.4.17 13:49	mass spec
Ec_4C	19.4.17 13:50	mass spec
Ec_5C	19.4.17 13:50	mass spec
Ec_6C	19.4.17 13:50	mass spec
Ec_7C	19.4.17 13:50	mass spec
Ec_8C	19.4.17 13:51	mass spec
Ec_9HC	19.4.17 13:51	mass spec
Ec_10HC	19.4.17 13:43	mass spec
Ec_11HC	19.4.17 13:43	mass spec
Ec_12HC	19.4.17 13:44	mass spec
Ec_13HC	19.4.17 13:44	mass spec
Ec_14HC	19.4.17 13:45	mass spec
Ec_15HC	19.4.17 13:45	mass spec



# Zajęcia praktyczne. Wizualizacja\_3





## Workflow4metabolomics



### Main menu

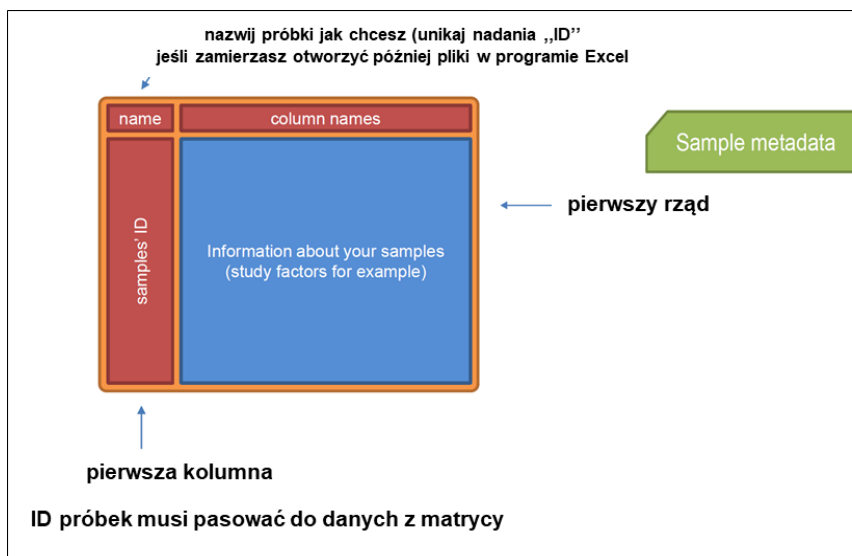
- Home
- Search
- Help
- Subscriptions
  - The Workflow4metabolomics
  - The W4M workflow
  - The W4M workflow
  - The W4M workflow
  - References
- How to
  - Download
  - Install
  - Automated Workflow and Metrics
  - How to create data?
- Developer resources
  - Getting ready
  - Virtual environments
- People
  - FAQ/Feedback

### Workflow4Metabolomics 3.0

Welcome to the collaborative portal dedicated to metabolomics data processing, analysis and annotation for Metabolomics community.


*"We are happy to announce the next **Workflow4Experimenters (W4E) International course 2018** using Galaxy and the Workflow4metabolomics infrastructure to analyse metabolomics data. Please save the date: **8-12 October 2018** at Pasteur Institute, Paris - France. More news in December !"*

<b>STEP 1</b> Request an account 	<b>STEP 2</b> Connect to Galaxy 	<b>STEP 3</b> Upload your data 	<b>STEP 4</b> Make reproducible science 
--	---	--	---



# Zajęcia praktyczne. Wizualizacja\_5

sampleName	class	polarity	sampleType	batch	injectionOrder	diet
QC	one	positive	pool	B1	1	NA
C1	one	positive	sample	B1	7	C
HC3	one	positive	sample	B1	10	HC
BL	one	positive	blank	B1	12	NA
...	...	...	...	...	...	...



Galaxy / 4 / Metabolomics

Tools



Download from web or upload from disk

Workflow4metabolomics

Download from web or upload from disk

Regular Compose Collection

You added 1 file(s) to the queue. Add more files or click "Start" to proceed.

Name	Size	Type	Genome	Settings	Status
 samples.zip	19.6 MB	Auto-detect	unspecified (?)		

Upload configuration

- Convert spaces to tabs
- Use POSIX standard

Type (set all): Auto-detect Genome (set all): unspecified (?)

Choose local file  Choose FTP file  Paste/Fetch data

# Zajęcia praktyczne. Wizualizacja\_6



Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Help User Using 7.6 GB

Tools

Download from URL or upload files from disk

Upload File from your computer

Workflow4metabolomics

LC-MS

Preprocessing  
Normalisation  
Quality Control  
Statistical Analysis  
Annotation

GC-MS

Preprocessing  
Normalisation  
Quality Control  
Statistical Analysis  
Annotation

NMR

Preprocessing  
Normalisation  
Quality Control  
Statistical Analysis  
Annotation

COMPOUND

Data Handling  
Text Manipulation  
Filtering

History

search datasets

Unnamed history (empty)

Download from web or upload from disk

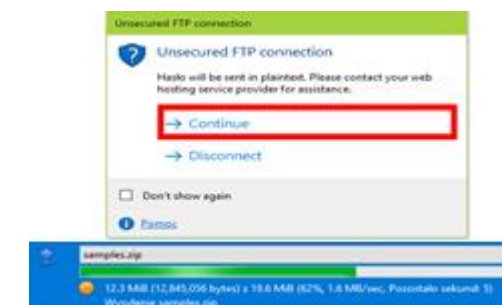
Regular Composite Collection

FTP files

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at **ftp.workflow4metabolomics.org** using your Galaxy credentials (email address and password).

Your FTP directory does not contain any files.

local file **Choose FTP file** Paste/Fetch data Pause Reset Start Close



Unsecured FTP connection

Unsecured FTP connection

Mask will be sent in plaintext. Please contact your web hosting service provider for assistance.

Continue

Disconnect

Don't show again

Files

samples.zip

12.3 MB (12,845,056 bytes) x 19.6 MB (52%, 1.8 MB/sec, Pozostało sekund 3)

Wytykanie samples.zip

# Zajęcia praktyczne. Wizualizacja\_7

## Download from web or upload from disk

Regular Composite Collection

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.

Name	Size	Type	Genome	Settings	Status
samples.zip	19.6 MB				

FTP files

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at <ftp.workflow4metabolomics.org> using your Galaxy credentials (email address and password).

Available files: 1 files 19.6 MB

<input checked="" type="checkbox"/>	Name	Size	Created
<input checked="" type="checkbox"/>	samples.zip	19.6 MB	11/22/2017 05:09:03 PM

Type (set all): unspecified (?)

Choose local file Choose FTP file Paste/Fetch data Pause Reset Start Close

## Download from web or upload from disk

Regular Composite Collection

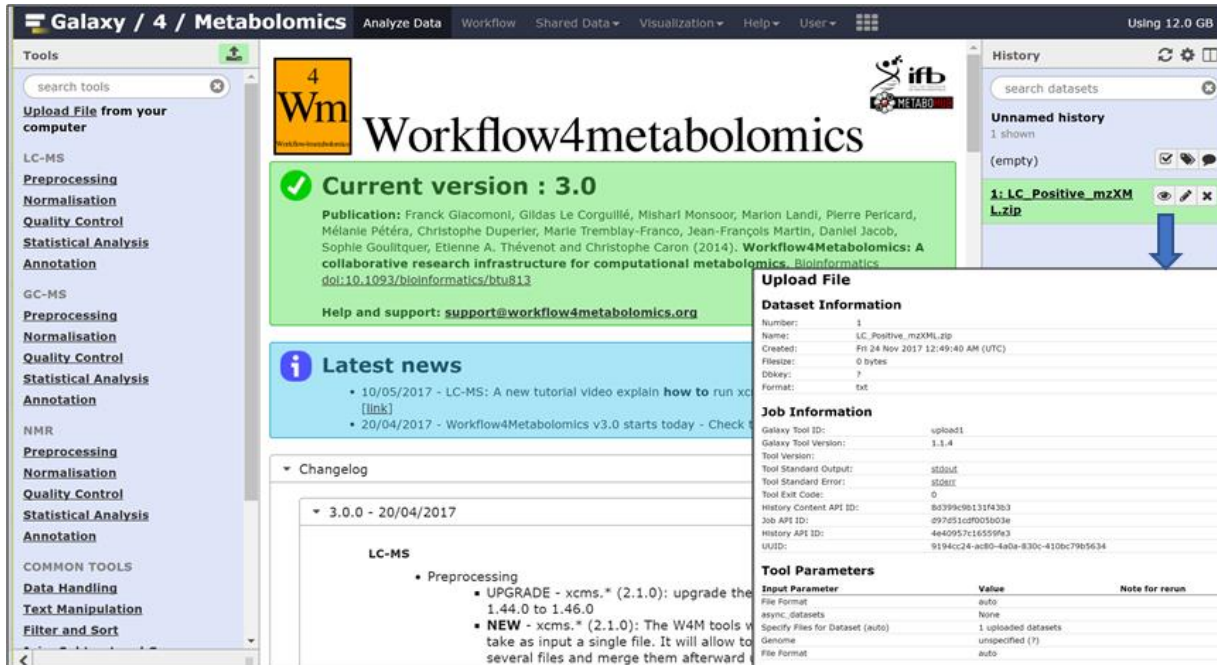
Name	Size	Type	Genome	Settings	Status
LC_Positive_mzXML.zip	3.6 GB	Auto-dete...	unspecified (?)		100% ✓

Type (set all): Auto-detect Genome (set all): unspecified (?)

Choose local file Choose FTP file Paste/Fetch data Pause Reset Start Close



# Zajęcia praktyczne. Wizualizacja\_8



**Galaxy / 4 / Metabolomics** Analyze Data Workflow Shared Data Visualization Help User Using 12.0 GB

**Workflow4metabolomics**

**Current version : 3.0**

**Publication:** Franck Giacomoni, Gildas Le Corguillé, Mishari Monsoor, Marion Landi, Pierre Pericard, Mélanie Pétera, Christophe Duperier, Marie Tremblay-Franco, Jean-François Martin, Daniel Jacob, Sophie Goultquer, Etienne A. Thévenot and Christophe Caron (2014). **Workflow4Metabolomics: A collaborative research infrastructure for computational metabolomics.** *Bioinformatics* doi:10.1093/bioinformatics/btu813

**Help and support:** [support@workflow4metabolomics.org](mailto:support@workflow4metabolomics.org)

**Latest news**

- 10/05/2017 - LC-MS: A new tutorial video explain **how to run** [\[link\]](#)
- 20/04/2017 - Workflow4Metabolomics v3.0 starts today - Check [\[link\]](#)

**Changelog**

- 3.0.0 - 20/04/2017
  - LC-MS
    - Preprocessing
      - UPGRADE - xcms.\* (2.1.0): upgrade the 1.44.0 to 1.46.0
      - NEW** - xcms.\* (2.1.0): The W4M tools v take as input a single file. It will allow to several files and merge them afterward

**Upload File**

**Dataset Information**

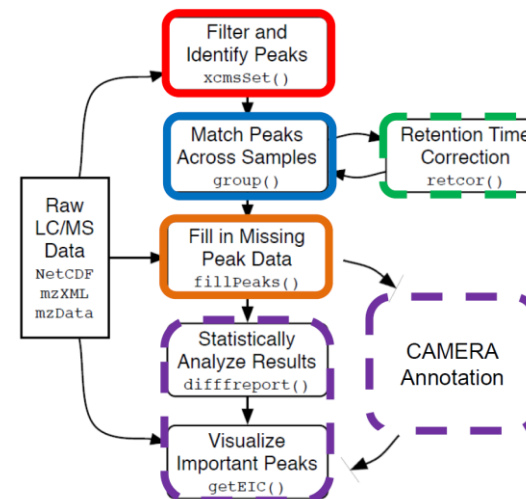
Number:	1
Name:	LC_Positive_mzXML.zip
Created:	Fri 24 Nov 2017 12:49:40 AM (UTC)
Filesize:	0 bytes
Dbkey:	?
Format:	txt

**Job Information**

Galaxy Tool ID:	upload1
Galaxy Tool Version:	1.1.4
Tool Version:	
Tool Standard Output:	STDOUT
Tool Standard Error:	STDERR
Tool Exit Code:	0
History Content API ID:	80399c90131f4303
Job API ID:	097d51e0f0050a3e
History API ID:	4e40957c16550f63
UUID:	9194cc24-ac80-4a0a-830c-410bc7905634

**Tool Parameters**

Input Parameter	Value	Note for rerun
File Format:	auto	
async_datasets:	None	
Specify Files for Dataset (auto):	1 uploaded datasets	
Genome:	unspecified (?)	
File Format:	auto	



- **Extraction**
  - Extraction of ions in each sample independently.
- **Grouping** ⇔ alignment
  - Each ion is aligned across all samples
- **Retention time correction** (optional)
- **Fill peaks**
  - Replace missing data with baseline value
- **Statistics and visualisation** (optionals)
- **CAMERA**
  - For annotation of adducts, neutral loss and isotopes

# Zajęcia praktyczne. Wizualizacja\_9

Galaxy / 4 / Metabolomics

Tools

Upload File from your computer

LC-MS

Processing

xcms.xcmsSet: Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis

xcms.xcmsSet Merger: Merge xcms.xcmsSet xset in one to be used by group

xcms.group: Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.

xcms.retop: Retention Time Correction using retop function from xcms R package

xcms.filterpeaks: Integrate a sample's signal in regions where peak groups are not represented to create new peaks in missing areas

xcms.summary: Create a

1 Job has been successfully added to the queue - resulting in the following datasets:

- 1: LC\_Positive\_mzXML.xset.RData
- 2: LC\_Positive\_mzXML.sampleMetadata.tsv
- 3: LC\_Positive\_mzXML.xset.BPCs\_raw.pdf
- 4: LC\_Positive\_mzXML.xset.TICs\_raw.pdf
- 5: LC\_Positive\_mzXML.xset.log.txt
- 6: LC\_Positive\_mzXML.xset.zip

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

History

search datasets

Unnamed history

6 shown

3.61 GB

- 6: LC\_Positive\_mzXML.xset.log.txt
- 5: LC\_Positive\_mzXML.xset.BPCs\_raw.pdf
- 4: LC\_Positive\_mzXML.xset.TICs\_raw.pdf
- 3: LC\_Positive\_mzXML.sampleMetadata.tsv
- 2: LC\_Positive\_mzXML.xset.RData
- 1: LC\_Positive\_mzXML.zip

sample	1	2	3
sampleMetadata	class	class	polarity
Et_10HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_11HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_12HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_13HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_14HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_15HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_16HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_17HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_18HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_19HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_20HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_21HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_22HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_23HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_24HC	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_2C	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_3C	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_4C	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_5C	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_6C	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_7C	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_8C	LC_Positive_mzXML	LC_Positive_mzXML	positive
Et_9C	LC_Positive_mzXML	LC_Positive_mzXML	positive
QC10	LC_Positive_mzXML	LC_Positive_mzXML	positive
QC11	LC_Positive_mzXML	LC_Positive_mzXML	positive
QC12	LC_Positive_mzXML	LC_Positive_mzXML	positive
QC13	LC_Positive_mzXML	LC_Positive_mzXML	positive
QC14	LC_Positive_mzXML	LC_Positive_mzXML	positive
QC15	LC_Positive_mzXML	LC_Positive_mzXML	positive

6: LC\_Positive\_mzXML.xset.log.txt

5: LC\_Positive\_mzXML.xset.BPCs\_raw.pdf

4: LC\_Positive\_mzXML.xset.TICs\_raw.pdf

3: LC\_Positive\_mzXML.sampleMetadata.tsv

2: LC\_Positive\_mzXML.xset.RData

1: LC\_Positive\_mzXML.zip

1

2

Data processing

# Zajęcia praktyczne. Wizualizacja\_10

xcms.group Grupuje piki w poprzek próbek przy użyciu nakładających się boksów m / z oraz oblicza wygładzone rozkłady pików w czasie chromatograficznym

	pool1B1			pool1B2			pool1B3		
	mz	rt	int	mz	rt	int	mz	rt	int
Independent peak lists	196.0905	66.6	7810936	196.0910	66.7	11733921	196.0902	66.6	7933325
	158.1180	67.4	71736	342.0310	69.0	74594	158.1173	67.4	82969
	342.0308	67.6	202268	267.0581	65.5	260877	342.0308	21.3	2581
	267.0581	65.5	282039	283.0318	65.2	424631	283.0320	65.3	357448
Group ions by m/z	196.0905	66.6	7810936	196.0910	66.7	11733921	196.0902	66.6	7933325
	158.1180	67.4	71736	342.0310	69.0	74594	158.1173	67.4	82969
	342.0308	67.6	202268	267.0581	65.5	260877	342.0308	21.3	2581
	267.0581	65.5	282039	283.0318	65.2	424631	283.0320	65.3	357448
Group ions by RT	196.0905	66.6	7810936	196.0910	66.7	11733921	196.0902	66.6	7933325
	158.1180	67.4	71736				158.1173	67.4	82969
	342.0308	67.6	202268	342.0310	69.0	74594			
	267.0581	65.5	282039	267.0581	65.5	260877			
			283.0318	65.2	424631	283.0320	65.3	357448	
<b>Resulting</b>									
	mz	rt	pool1B1	pool1B2	pool1B3				
	196.0905	66.6	7810936	11733921	7933325				
	158.1176	67.4	71736		82969				
	342.0308	21.3			2581				
	342.0309	68.3	202268	74594					
	267.0581	65.5	282039	260877					
	283.0319	65.2		424631	357448				

<b>Parameter : num + label</b>	<b>Format</b>
Or : RData file   rdata.xcms.raw	
Or : RData file   rdata.xcms.retcor	

**xcms.group Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time. (Galaxy Version 2.1.0)** Versions Options

**xset RData file**

No rdata.xcms.raw, rdata.xcms.group, rdata.xcms.retcor or rdata dataset available. ▼  
output file from another function xcms (xcmsSet, retcor etc.)

**Method to use for grouping**

density ▼  
[method] See the help section below

**Bandwidth**

30  
[bw] bandwidth (standard deviation or half width at half maximum) of gaussian smoothing kernel to apply to the peak density chromatogram

**Minimum fraction of samples necessary**

0.5  
[minfrac] in at least one of the sample groups for it to be a valid group

**Width of overlapping m/z slices**

0.01  
[mzwid] to use for creating peak density chromatograms and grouping peaks across samples

**Advanced options**

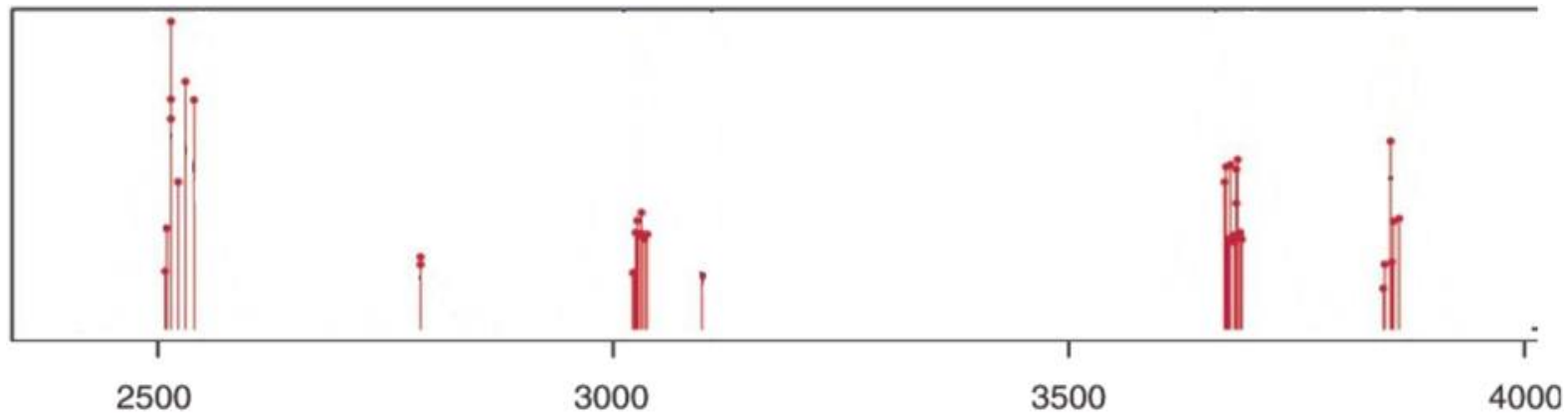
show ▼

**Maximum number of groups to identify in a single m/z slice**

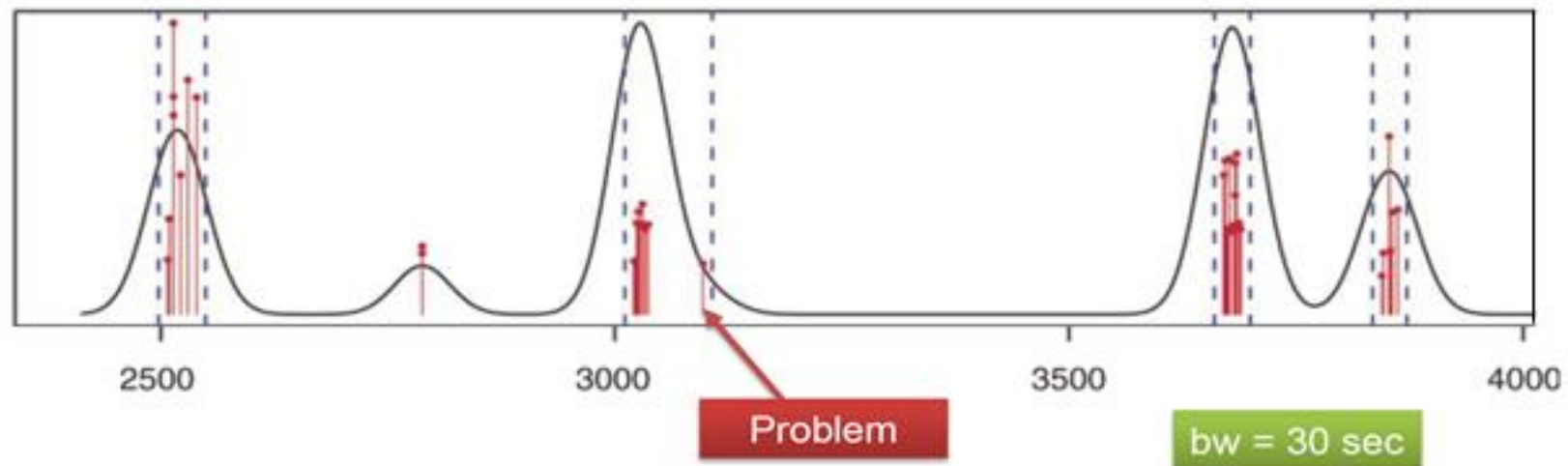
50  
[max]

**Get a Peak List**

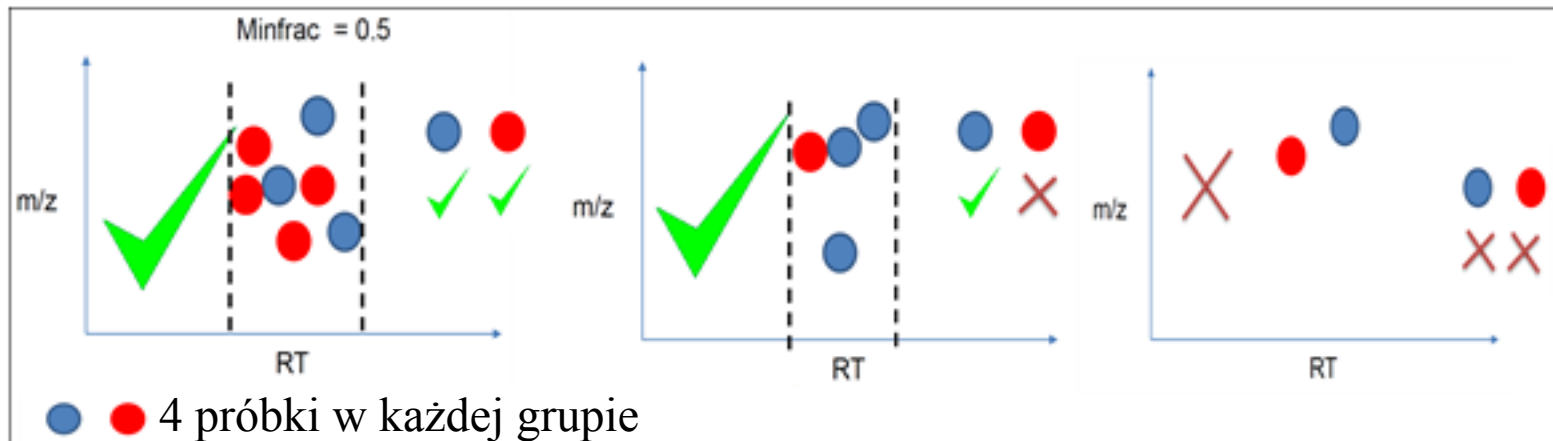
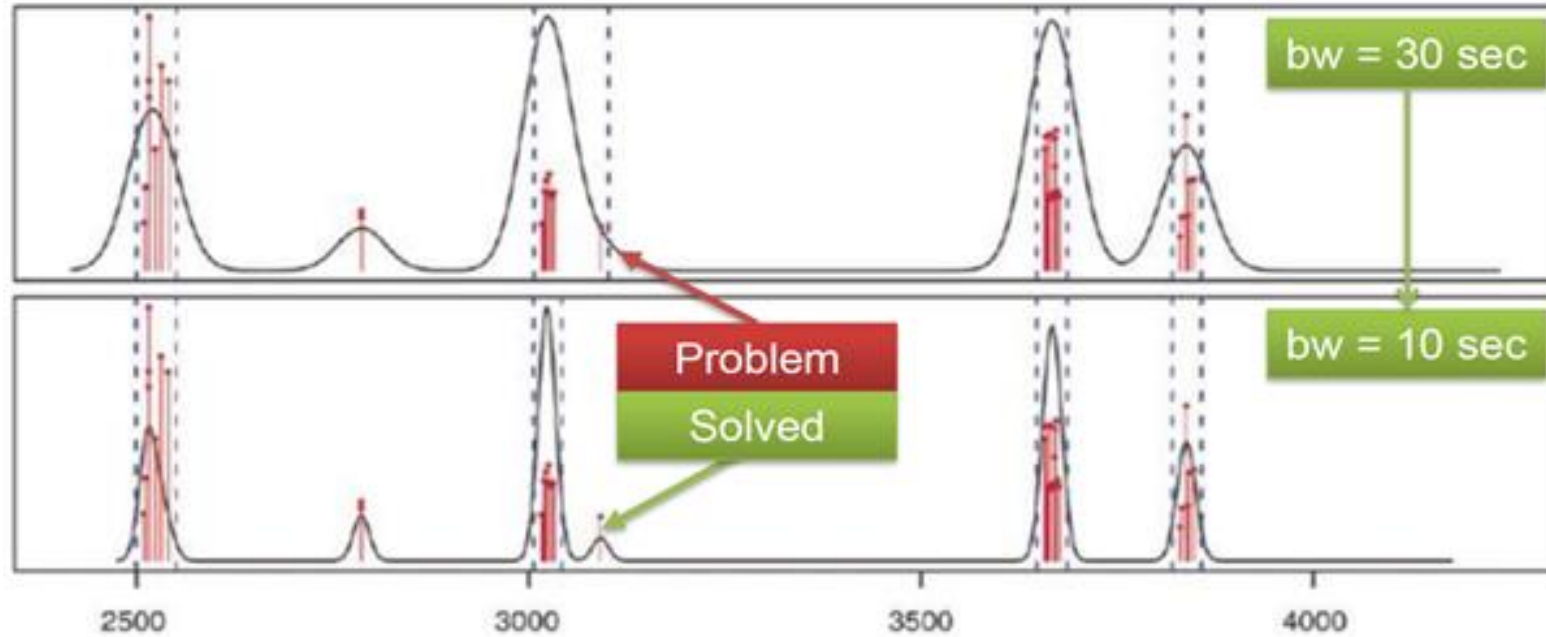
**Grupowanie pików w przedziale mas: 337,975 - 338,225 m/z (mzwid)**



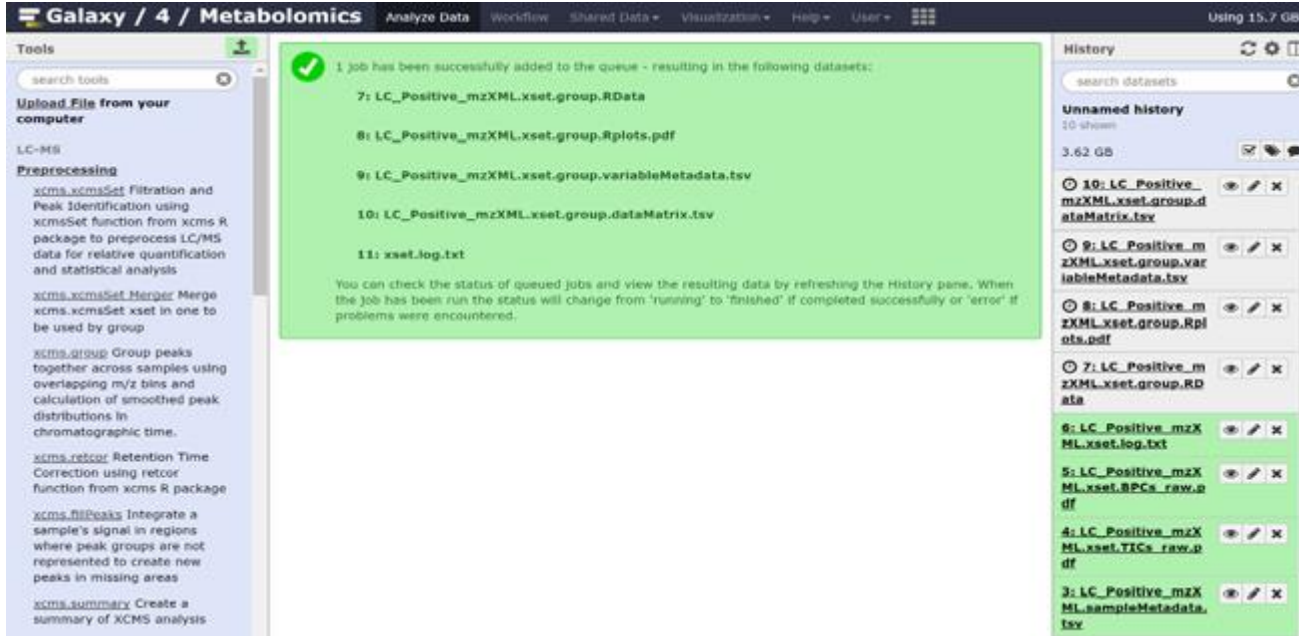
**Grupowanie pików w przedziale mas: 337,975 - 338,225 m/z (mzwid)**



## Grupowanie pików w przedziale mas: 337,975 - 338,225 m/z (mzwid)



# Zajęcia praktyczne. Wizualizacja\_14



**Galaxy / 4 / Metabolomics** Analyze Data Workflow Shared Data Visualization Help User Using 15.7 GB

**Tools**

search tools

**Upload File from your computer**

LC-MS

**Preprocessing**

- xcms.xcmsSet** Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis
- xcms.xcmsSet.Merger** Merge xcms.xcmsSet xset in one to be used by group
- xcms.group** Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.
- xcms.retcor** Retention Time Correction using retcor function from xcms R package
- xcms.fillpeaks** Integrate a sample's signal in regions where peak groups are not represented to create new peaks in missing areas
- xcms.summary** Create a summary of XCMS analysis

**History**

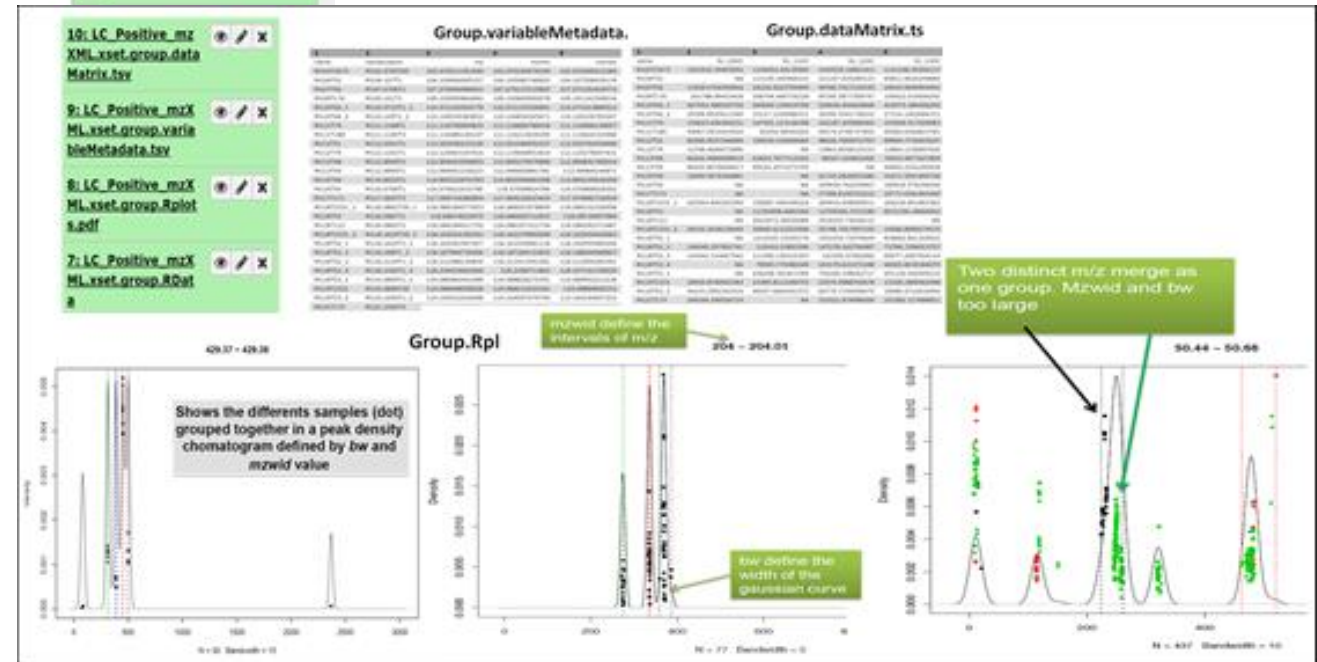
search datasets

Unnamed history  
10 shown  
3.62 GB

1 job has been successfully added to the queue - resulting in the following datasets:

- 7: LC\_Positive\_mzXML.xset.group.RData
- 8: LC\_Positive\_mzXML.xset.group.Rplots.pdf
- 9: LC\_Positive\_mzXML.xset.group.variableMetadata.tsv
- 10: LC\_Positive\_mzXML.xset.group.dataMatrix.tsv
- 11: xset.log.txt

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.



# Zajęcia praktyczne. Wizualizacja\_15

xcms.retcor Korekta czasu retencji za pomocą funkcji rektora z pakietu xcms R.

Parameter : num + label	Format
1 : RData file	rdata.xcms.group

**xcms.retcor Retention Time Correction using retcor function from xcms R** [Versions](#) [Options](#)  
package (Galaxy Version 2.1.0)

**xset RData file**  
   No rdata.xcms.raw, rdata.xcms.group or rdata dataset available.  
output file from another function xcms (xcmsSet, retcor etc.)

**Method to use for retention time correction**  
peakgroups  
[method] See the help section below

**Smooth method**  
loess  
[smooth] either 'loess' for non-linear alignment or 'linear' for linear alignment

**Number of extra peaks to allow in retention time correction correction groups**  
1  
[extra]

**Number of missing samples to allow in retention time correction groups**  
1  
[missing]

**Advanced options**  
hide

[Resubmit your raw dataset or your zip file](#)

liczba dopuszczonych, brakujących, widocznych pików w grupie



# Zajęcia praktyczne. Wizualizacja\_16

**Advanced options**

show

**Degree of smoothing for local polynomial regression fitting**

0.2

[span]

**Family**

gaussian

[family] If gaussian fitting is by least-squares with no outlier removal, and if symmetric a re descending M estimator is used with Tukey's biweight function, allowing outlier removal

**plottype**

deviation

wykres wizualizacji wyniku korekty czasu retencji

[plottype] If deviation plot retention time deviation points and regression fit, and if mdevden also plot peak overall peak density and retention time correction peak density

**Resubmit your raw dataset or your zip file**

Execute

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Help User

Using 15.8 GB

Tools

search tools

Upload File from your computer

LC-MS

**Preprocessing**

[xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis

[xcms.xcmsSet Merger](#) Merge xcms.xcmsSet xset in one to be used by group

[xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.

[xcms.retcor](#) Retention Time Correction using retcor function from xcms R package

[xcms.fillPeaks](#) Integrate a sample's signal in regions where peak groups are not represented to create new peaks in missing areas

[xcms.summary](#) Create a summary of XCMS analysis

CAMERA.annotate CAMERA

1 job has been successfully added to the queue - resulting in the following datasets:

- 12: LC\_Positive\_mzXML.xset.group.retcor.RData
- 13: LC\_Positive\_mzXML.xset.group.retcor.Rplots.pdf
- 14: LC\_Positive\_mzXML.xset.group.retcor.TICs\_corrected.pdf
- 15: LC\_Positive\_mzXML.xset.group.retcor.BPCs\_corrected.pdf
- 16: xset.log.txt

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

History

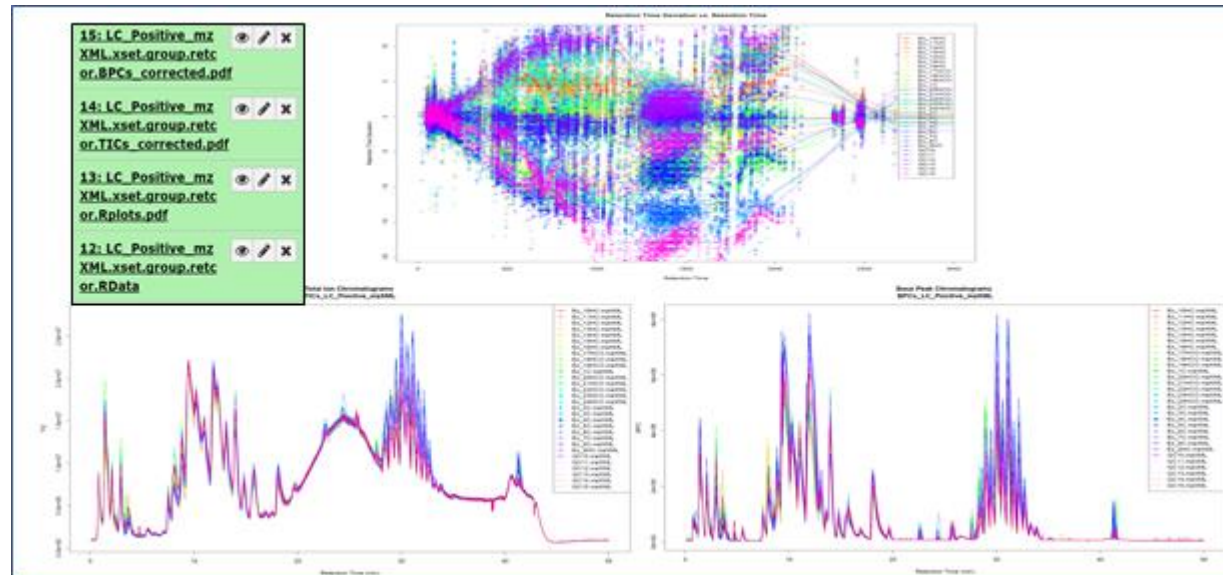
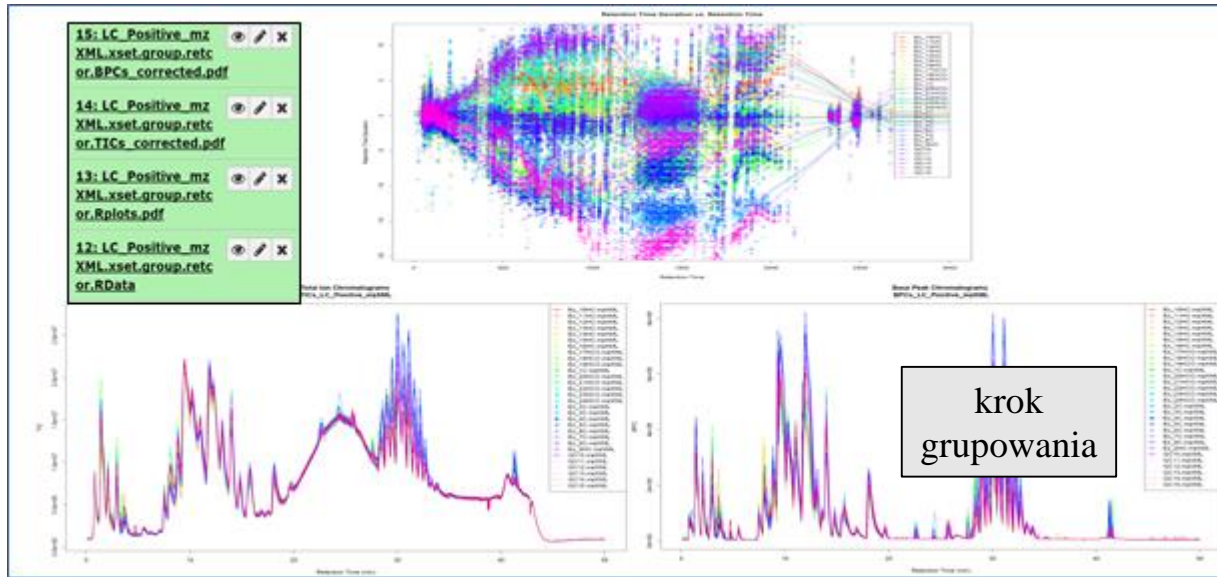
search datasets

Unnamed history

14 shown, 1 hidden

3.71 GB

- 15: LC\_Positive\_mzXML.xset.group.retcor.BPCs\_corrected.pdf
- 14: LC\_Positive\_mzXML.xset.group.retcor.TICs\_corrected.pdf
- 13: LC\_Positive\_mzXML.xset.group.retcor.Rplots.pdf
- 12: LC\_Positive\_mzXML.xset.group.retcor.RData
- 10: LC\_Positive\_mzXML.xset.group.data.Matrix.tsv
- 9: LC\_Positive\_mzXML.xset.group.variableMetadata.tsv
- 8: LC\_Positive\_mzXML.xset.group.Rplots.pdf
- 7: LC\_Positive\_mzXML.xset.group.RData



# Zajęcia praktyczne. Wizualizacja\_18

The screenshot shows the Galaxy interface with a green notification box stating: "1 job has been successfully added to the queue - resulting in the following datasets:". Below this, a list of datasets is shown:

- 17: LC\_Positive\_mzXML.xset.group.retcor.group.RData
- 18: LC\_Positive\_mzXML.xset.group.retcor.group.Rplots.pdf
- 19: LC\_Positive\_mzXML.xset.group.retcor.group.variableMetadata.tsv
- 20: LC\_Positive\_mzXML.xset.group.retcor.group.dataMatrix.tsv
- 21: xset.log.txt

Below the list, a message reads: "You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered."

krok grupowania

xcms.fillPeaks Integruje sygnał próbek w regionach, w których grupy pików nie są reprezentowane, aby utworzyć nowe piki w brakujących obszarach

Parameter : num + label	Format
1 : RData file	rdata.xcms.group

**xcms.fillPeaks Integrate a sample's signal in regions where peak groups are not represented to create new peaks in missing areas (Galaxy Version 2.1.0)** Versions Options

**xset RData file**

No rdata.xcms.group or rdata dataset available.

output file from another xcms function (group)

**Filling method**

chrom

[method] See the help section below

**Get a Peak List**

Yes No

**Resubmit your raw dataset or your zip file**

Execute

# Zajęcia praktyczne. Wizualizacja\_19

**Galaxy / 4 / Metabolomics** Analyze Data Workflow Shared Data Visualization Help User Using 15.9 GB

**Tools**

- search tools
- xcms:xcmsSet function using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis
- xcms:xcmsSet\_Merger Merge xcms:xcmsSet xset in one to be used by group
- xcms:group Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.
- xcms:retcor Retention Time Correction using retcor function from xcms R package
- xcms:fillPeaks Integrate a sample's signal in regions where peak groups are not represented to create new peaks in missing areas
- xcms:summary Create a summary of XCMS analysis
- CAMERA:annotate CAMERA annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition.
- CAMERA:combineXcnosAnnos Wrapper function for the

**History**

search datasets

Unnamed history  
21 shown, 3 hidden  
3.82 GB

- 24: LC\_Positive\_mzXML.xset.group.r etcor.group.fillpeaks.dataMatrix X.tsv
- 23: LC\_Positive\_mzXML.xset.group.r etcor.group.fillpeaks.variableMetadata.tsv
- 22: LC\_Positive\_mzXML.xset.group.r etcor.group.fillPeaks.RData
- 20: LC\_Positive\_mzXML.xset.group.retc or.group.dataMatrix.tsv
- 19: LC\_Positive\_mzXML.xset.group.retc or.group.variableMetadata.tsv
- 18: LC\_Positive\_mzXML.xset.group.retc or.group.Rplots.pdf
- 17: LC\_Positive\_mzXML.xset.group.retc or.group.RData
- 15: LC\_Positive\_mz

**1 job has been successfully added to the queue - resulting in the following datasets:**

- 22: LC\_Positive\_mzXML.xset.group.retcor.group.fillPeaks.RData
- 23: LC\_Positive\_mzXML.xset.group.retcor.group.fillpeaks.variableMetadata.tsv
- 24: LC\_Positive\_mzXML.xset.group.retcor.group.fillpeaks.dataMatrix.tsv
- 25: xset.log.txt

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

variableMetadata						dataMatrix					
1	2	3	4	5		1	2	3	4	5	
name	namecustom	mz	mzmin	mzmax		name	Ec_12nC	Ec_13nC	Ec_12nC	Ec_13nC	
M10372673	M102.9703744.55	102.970311351949	102.970195579509	102.970399312385		M10372673	1061832.30493052	1184403.44135983	1020539.54801412	1121248.09306137	
M104752	M104.10770.86	104.106964995157	104.105982748003	104.107508430139		M104752	116131.202463165	113145.345989103	221247.875289123	85812.9626299884	
M107756	M107.070670.94	107.070666868921	107.070137251825	107.07251958772		M107756	11028.4704300842	24224.0223756465	40340.7417124155	20019.0940904541	
M1097176	M109.10172.93	109.10099584082	109.100845900078	109.10142308016		M1097176	20178.89433418	208794.460755228	95305.5877398747	106222.014504292	
M100786_1	M10.0712271.1_1	110.071223264779	110.071135326801	110.074533889512		M100786_1	447052.480245756	509069.134439706	529016.944020849	419073.389404295	
M100786_2	M10.149371.11_2	110.145035383832	110.144834345671	110.145230793207		M100786_2	29398.8529611569	33127.1524860311	36290.5541728252	27704.1463684311	
M111779	M11.116871.31	111.116796950832	111.116664796058	111.116964539657		M111779	154633.650364231	107591.113126188	162187.205989563	115309.517209983	
M111780	M11.116573	111.116486144107	111.116213639299	111.11664312998		M111780	54067.0514414524	65254.58342024	49274.2740727852	60540.6364823781	
M12751	M12.052170.86	112.052056235529	112.051444952337	112.052702434446		M12751	82566.9537246094	100534.324096069	98626.7065472707	89809.7776453247	
M12779	M12.120171.31	112.120063107016	112.119644951814	112.120279047431		M12779	12748.4660072896	20138.4564137684	13863.0038125153	12884.1150847926	
M133748	M13.899970.81	112.899502356853	112.895275075808	112.89651780014		M133748	66262.9480098419	62663.7677114182	88327.424822492	70933.4877627869	
M13759	M13.999570.98	112.999453182223	112.999065891791	112.99999144871		M13759	96435.8076669617	85020.2573272705	84614.7268862794	90950.2521185818	
M155748	M14.893370.81	114.893323937483	114.892556681068	114.8944315034299		M155748	32890.9970340881	32496.0212502232	41719.2826053286	31671.5001469726	
M161754	M16.070670.9	116.070614307965	116.0708824796	116.070868028302		M161754	172069.77323927	199833.290424495	259939.762530467	189934.579256049	
M177171	M17.069772.85	117.069743980664	117.069232623419	117.070662751624		M177171	25315.1286311528	28091.6590618176	17398.6160722212	19771.6361491645	
M1872331_1	M18.0662770.88_1	118.066184577653	118.066051979809	118.066332328959		M1872331_1	425563.804303394	338087.459169224	394919.658065011	264230.851963383	
M18753_1	M18.086770.88_1	118.086746026075	118.086685712653	118.08738457984		M18753_1	1437074.3398173	11354958.4483364	13709306.7272189	8072194.34844452	
M187112	M18.086271.87	118.086184051792	118.086107212736	118.086292372487		M187112	2734831.17062041	2023073.28526089	2918253.74628112	1595919.90272233	
M1872331_2	M18.162424936784	118.162424936784	118.162275890608	118.162354424063		M1872331_2	29155.2018136445	20060.612223546	25748.7617057155	14566.8990274575	
M18753_2	M18.162570.88_2	118.162541997427	118.161930081116	118.162955469244		M18753_2	1645565.67086	1214220.15291376	1501034.72970629	818662.841329422	
M18752_1	M18.18870.87_1	118.187994792026	118.187799133652	118.188664468827		M18752_1	146040.357902741	115410.21893396	147278.242704987	73784.7296913757	
M18752_2	M18.212470.87_2	118.212388136854	118.211413391281	118.212900282301		M18752_2	143442.314667542	111096.156310303	143350.57062082	69977.2087654114	
M18753_3	M18.236570.87_3	118.236519662066	118.2368718863	118.237341746029		M18753_3	1627977.36488696	78505.770380249	103179.631172188	48305.8035340075	
M19753_1	M19.089570.88_1	119.089442441589	119.08825274391	119.08995212236		M19753_1	798878.512076424	636208.391813786	754206.198042717	451126.926595215	
M1972330	M19.0885738.84	119.088490588528	119.08821520181	119.08889050151		M1972330	24954.8740942383	23399.812346703	23575.9488765678	13338.1884562498	
M19753_2	M19.165670.88_2	119.16562425608	119.164957479794	119.166194857233		M19753_2	54033.2992341819	49257.9664361572	60770.1959596075	35086.4712632904	
M2121773	M21.100272.88	121.100193999284	121.100016027078	121.100258099592		M2121773	240164.240526719	3207661.98720796	331021.874989269	251582.117688811	
M2121779	M21.1001728.98	121.10013672057	121.099952084354	121.100214241705		M2121779	2425847.45517966	174789.838199487	288666.654292461	249454.370647165	
M21217176	M21.124172.94	121.124074566867	121.123811637319	121.124823245796		M21217176	352694.554213724	130687.160363821	426307.452419805	24478.4578494873	

**24: LC\_Positive\_mz**  
**XML.xset.group.ret**  
**or.group.fillpeaks.dataMatrix.ts**

6,929 lines  
 format: **tabular**, database: **2**

**PACKAGE INFO**  
 parallel 3.2.5  
 BiocGenerics 0.16.1  
 Biobase 2.30.0  
 Rcpp 0.12.10  
 mzR 2.4.1  
 xcms 1.46.0  
 snow 0.4.2  
 batch 1.1.4

**ARGUMENTS INFO**  
 xfunction fillPeaks  
 image  
 /work/project/w4m/galaxy4metab  
 /dist/database/files/000/440/datasr

**Exported data**

TID	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	name	Ec_10HC	Ec_11HC	Ec_12HC	Ec_13HC	Ec_14HC	Ec_15HC	Ec_16HC	Ec_17HC	Ec_18HC	Ec_19HC	Ec_20HC	Ec_21HC	Ec_22HC	Ec_23HC	Ec_24HC	Ec_2C	Ec_3C	Ec_4	
2	M0212267	1061832	1144403	1202039	1212348	1475083	1045868	978834	1175380	1011507	1138730	1030791	9712483	1031258	457818	878424	1013329	1088809	1136409	10

xcms.summary Tworzy podsumowanie analizy XCMS

**xcms.summary** Create a summary of XCMS analysis (Galaxy Version 1.0.3) Versions Options

**xset RData file**

output file from another function xcms (xcmsSet, group, retcor, fillpeaks etc.)

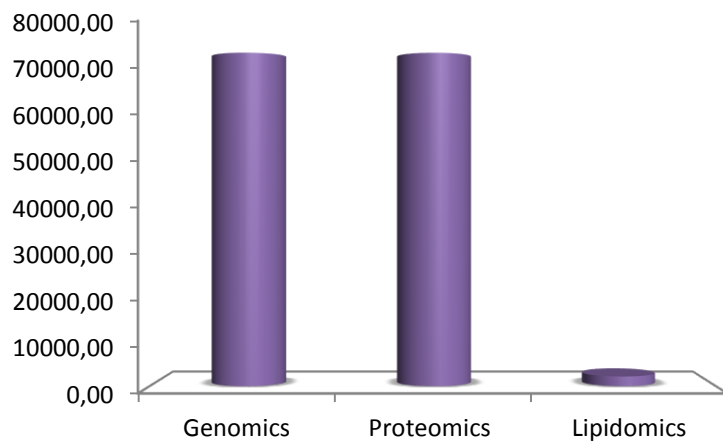
# Lipidomika z wykorzystaniem spektrometrii mas

Elisabete Maciel,  
Eliana Alves,  
Pedro Domingues,  
Rosário Domingues

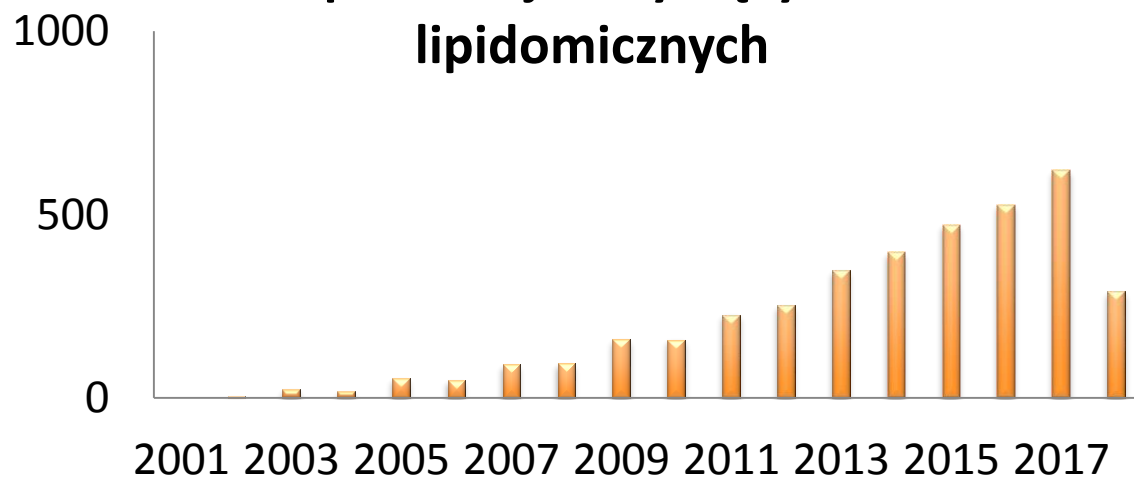


# LIPIDOMIKA

Liczba publikacji z dziedziny nauk omicznych  
(Scopus)



Liczba publikacji dotyczących badań  
lipidomicznych



# Lipidomika

- Pełna charakterystyka związków lipidowych oraz ich biologicznych funkcji w powiązaniu z ekspresją białek warunkujących metabolizm lipidów i ich funkcje, a także ekspresją genów (wg. AOCS Lipids Library)
- Analiza profilu lipidowego w warunkach fizjologicznych oraz jego zmian w stanach patologicznych

## Lipidomika

Ekstrakcja  
lipidów

Analiza  
lipidów

Analiza szlaków  
metabolicznych

Analiza oddziaływań  
lipidy-białka



# Lipidomika

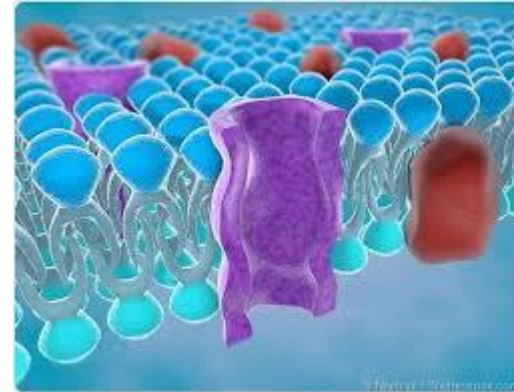
---

- Profilowanie lipidomu komórki
- Lipidy w strukturze błony komórkowej oraz jej dynamika
- Regulacyjna funkcja lipidów (np. jako cząsteczek sygnałowych)
- Integracja omik z wzajemnymi oddziaływaniami składników komórki oraz przemianami zachodzącymi w komórce/organizmie

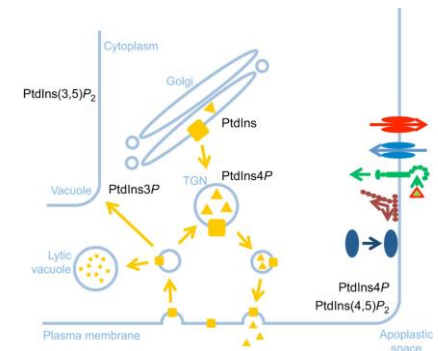
# Dlaczego lipidy są ważne?

**Błony komórkowe**

**Funkcja regulująca w komórce**  
przebieżnictwo sygnałowe,  
hormony, ...



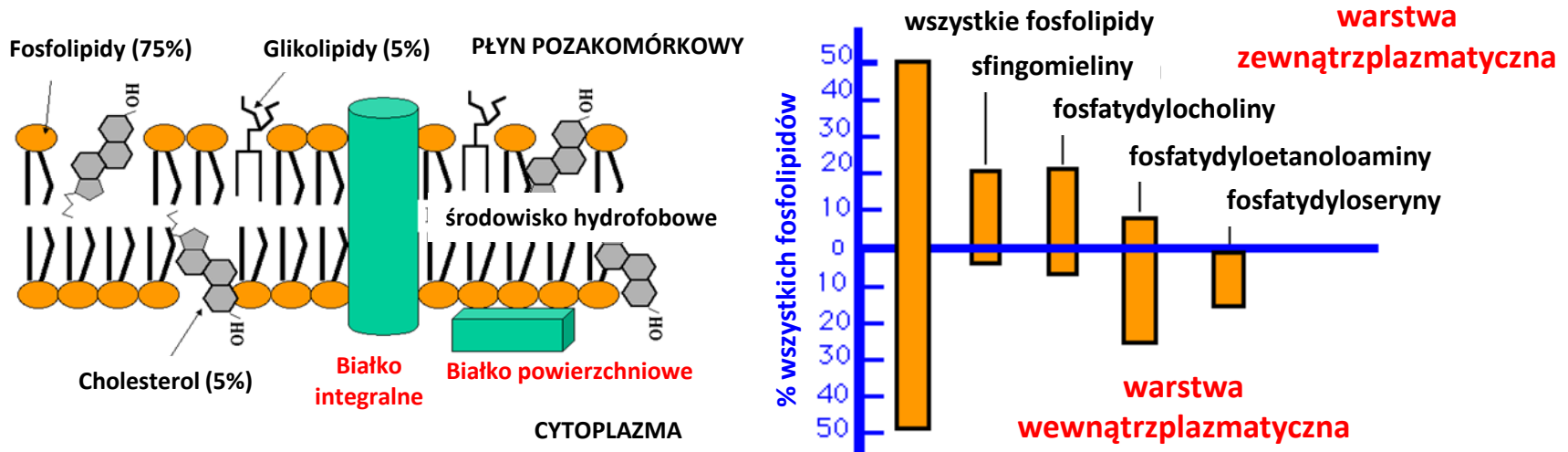
**Metabolizm energetyczny/rezerwa energetyczna**



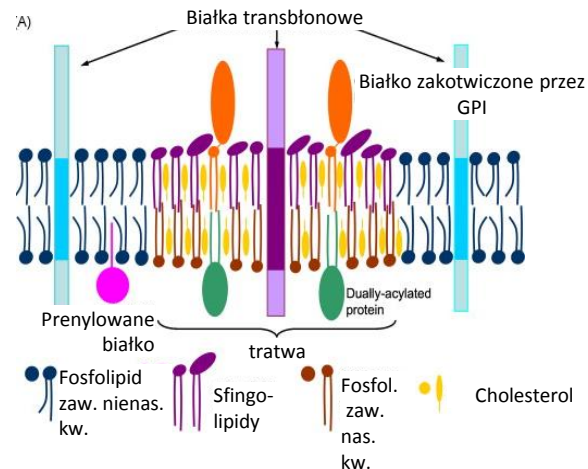
Zaburzenia metabolizmu lipidów i przebieżnictwa sygnałowego z ich udziałem odgrywają kluczową rolę **w fizjologii jak również stanach patologicznych**

# Lipidy – błona komórkowa

## Asymetria błony komórkowej



## Struktura błony tratwy lipidowe



# Profilowanie lipidów w komórce, tkankach i płynach ustrojowych

Każdy rodzaj komórek, tkanka oraz płyn ustrojowy posiada swój charakterystyczny profil lipidowy o określonym składzie związków lipidowych

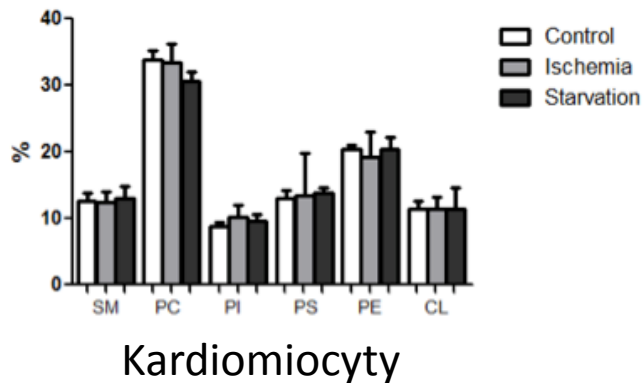
Identyfikacja wszystkich lipidów obecnych w komórce -

**lipidom**

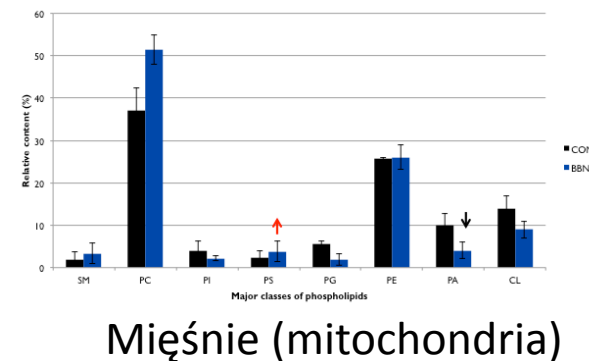
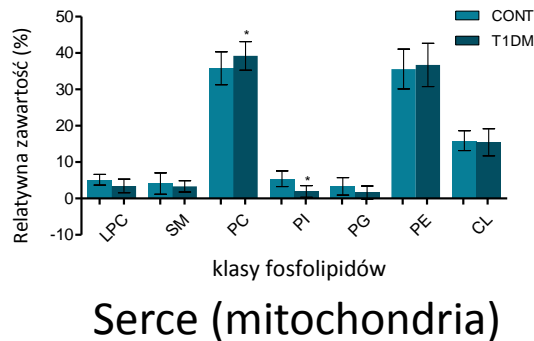
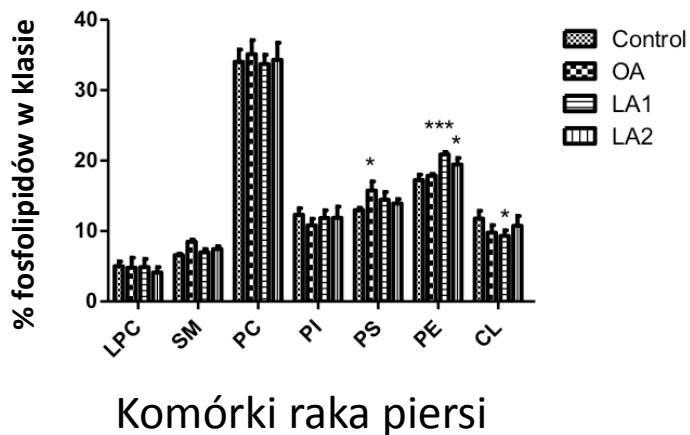
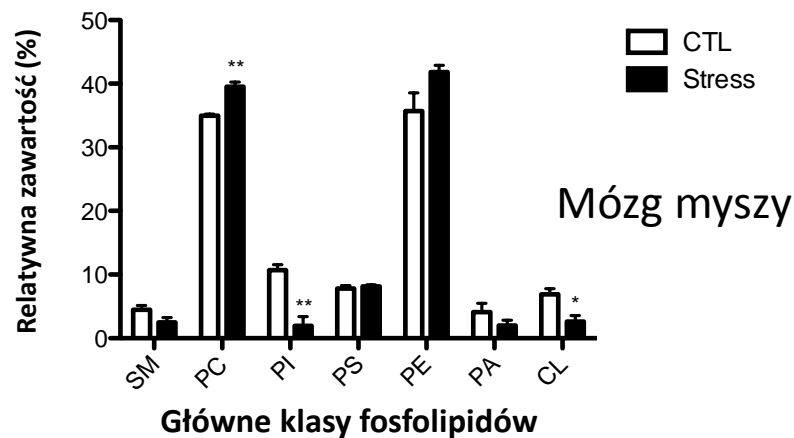
## % lipidów zawierających kwasy tłuszczowe

Glicerolipidy	Tylakoidy chloroplastów	Błona wewnątrzmitochondrialna	Błona komórkowa
<b>MGDG</b>	<b>51%</b>	<b>0</b>	<b>0</b>
<b>DGDG</b>	<b>26</b>	<b>0</b>	<b>0</b>
<b>SQDG</b>	<b>7</b>	<b>0</b>	<b>0</b>
<b>PC</b>	<b>3</b>	<b>27</b>	<b>32</b>
<b>PS</b>	<b>0</b>	<b>25</b>	<b>0</b>
<b>PG</b>	<b>9</b>	<b>0</b>	<b>0</b>
<b>PE</b>	<b>0</b>	<b>29</b>	<b>46</b>
<b>PI</b>	<b>1</b>	<b>0</b>	<b>19</b>
<b>CL</b>	<b>0</b>	<b>20</b>	<b>0</b>

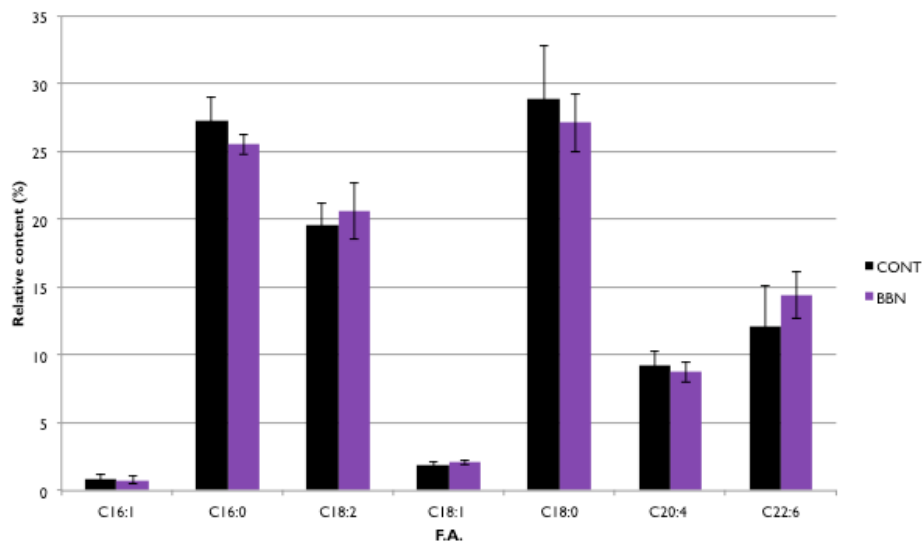
# Profilowanie klas fosfolipidów



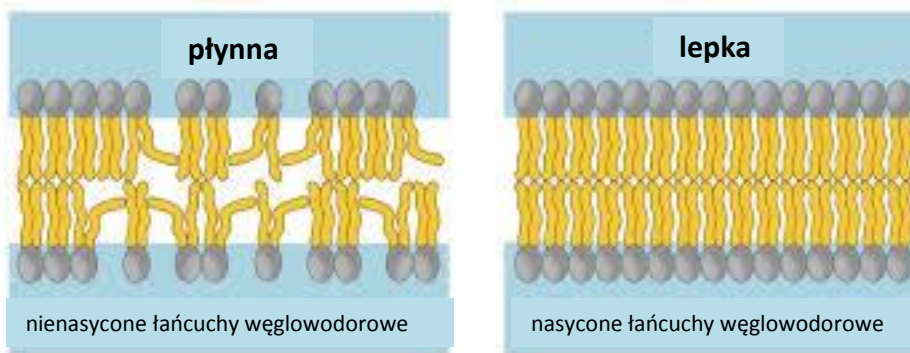
Zawartość fosfolipidów w poszczególnych klasach



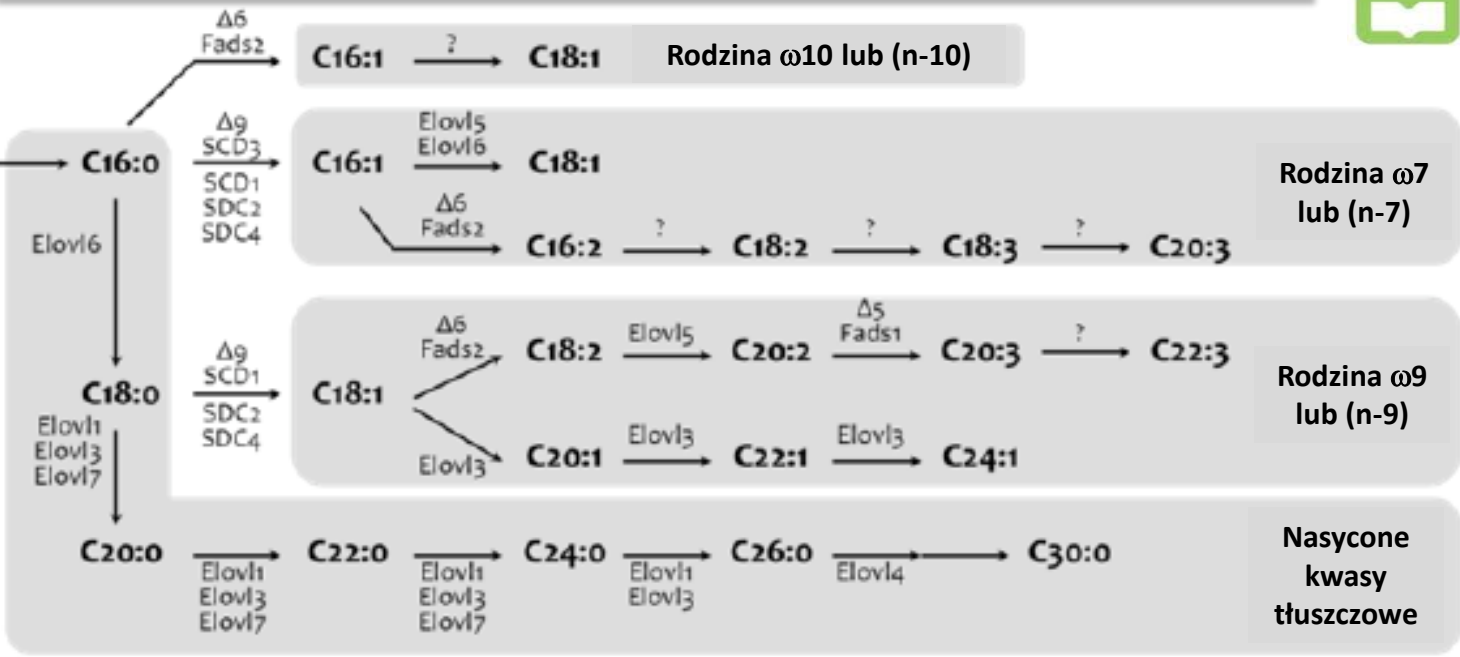
# Profil kwasów tłuszczowych



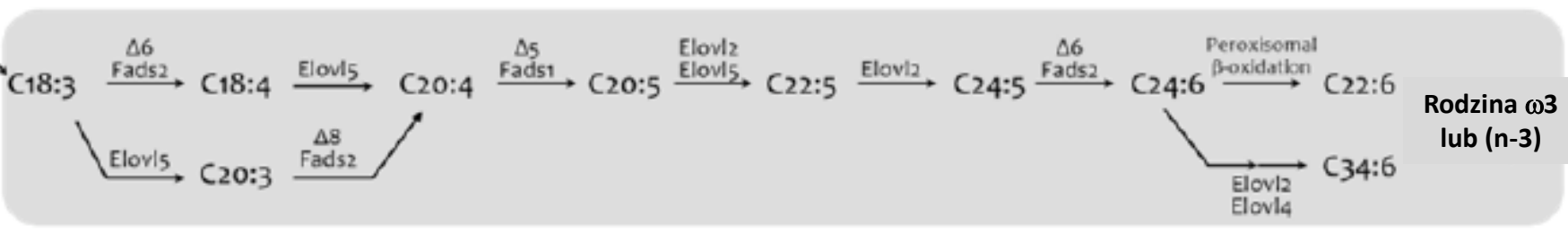
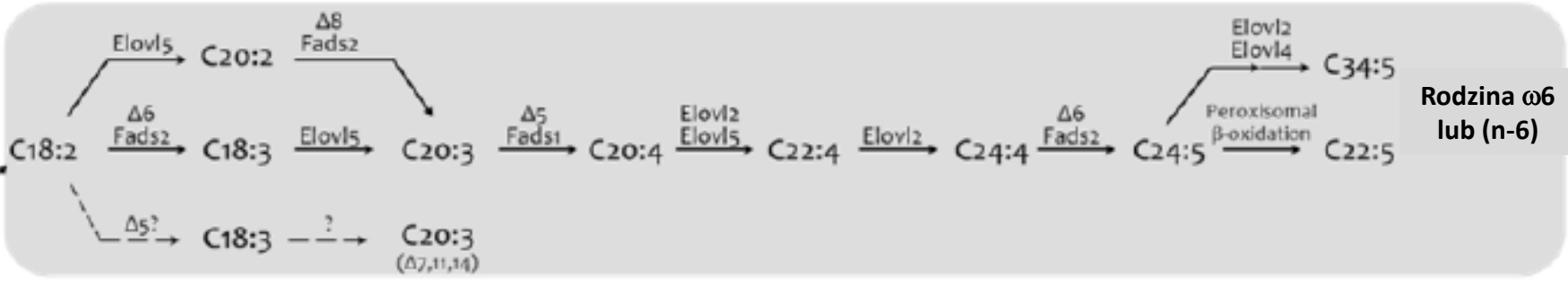
Wpływ rodzaju kwasów tłuszczowych na właściwości błony komórkowej



Syntaza kwasów tłuszczowych



Dieta



# Jakie są największe wyzwania w lipidomice?

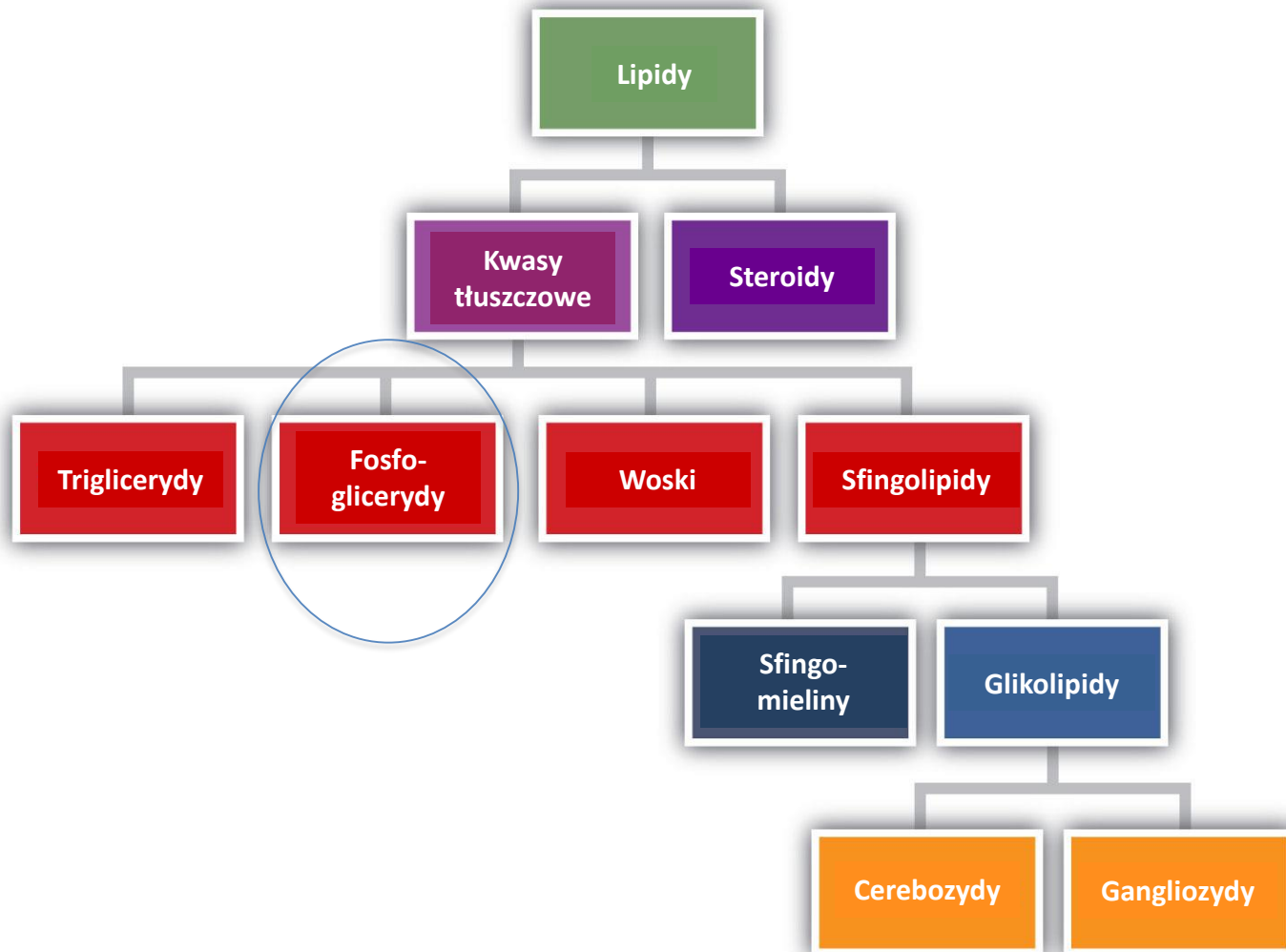
---

AALifeSci 



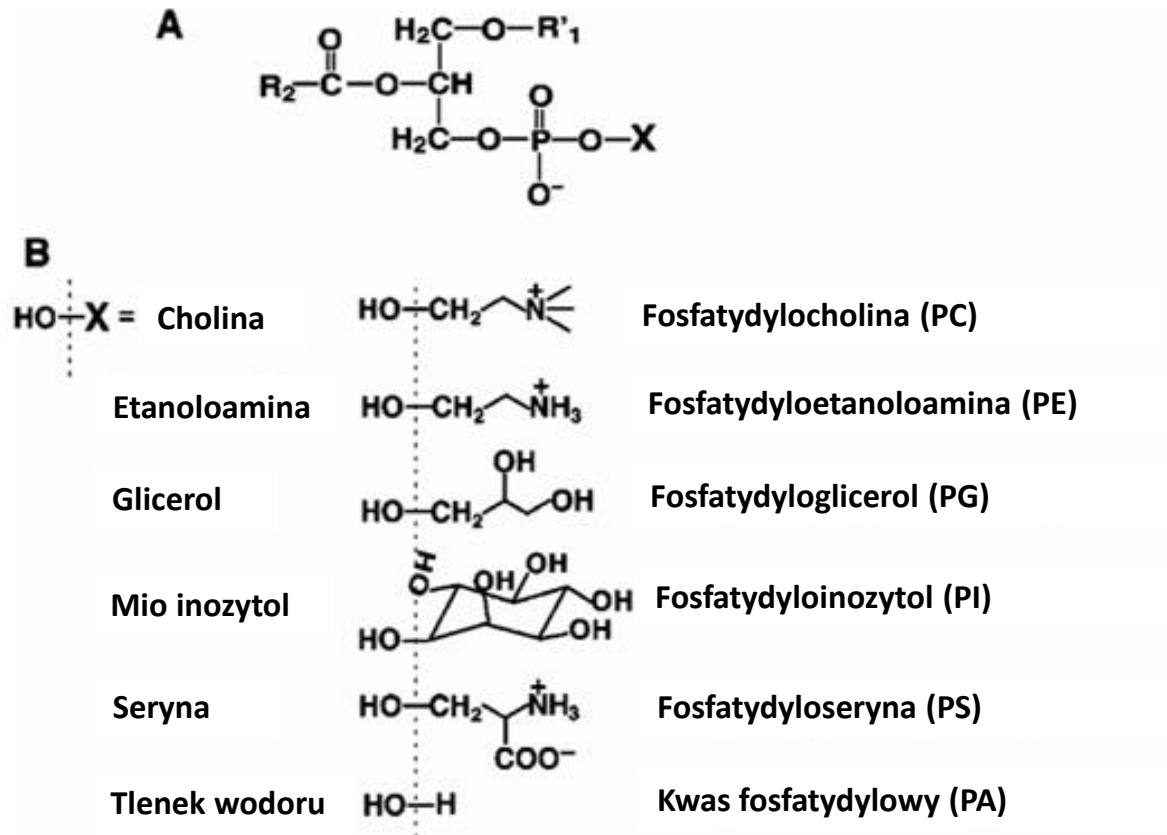


# Strukturalna różnorodność lipidów

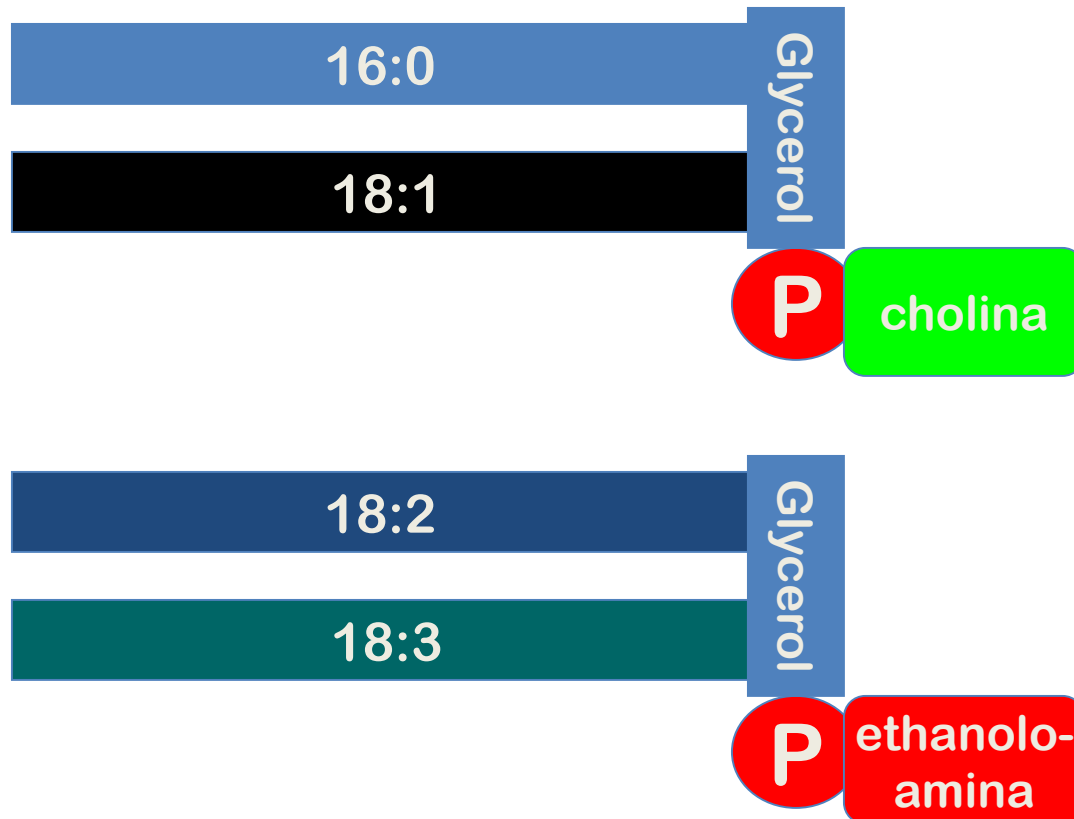


# Klasy fosfolipidów/glicerolipidów

Znane także jako glicerofosfolipidy



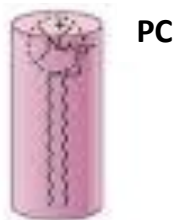
# Klasy fosfolipidów/glicerolipidów



*PC, PE, PS, NPE, PG, CL, PI, PIP, PIP<sub>2</sub>, PIP<sub>3</sub>, LysoPLs.....*

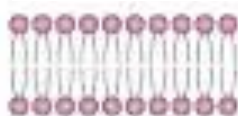
# Fosfolipidy / zawartość kwasów tłuszczowych / a właściwości błony komórkowej

**a** lipidy cylindryczne



PC

dwuwarstwa



**b**

zew.



wew.

lipidy stożkowe

AA



LPC



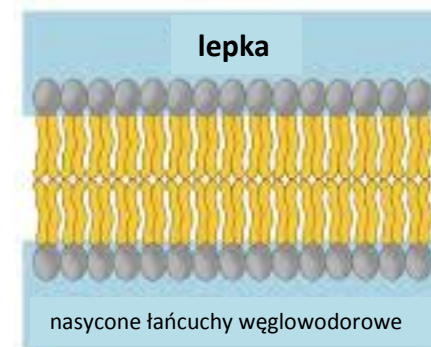
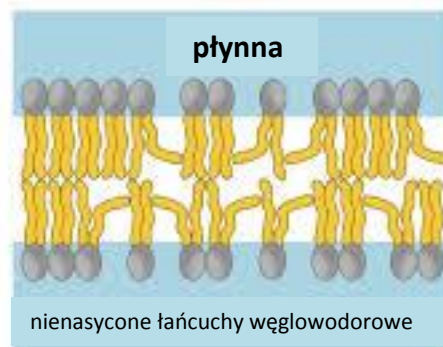
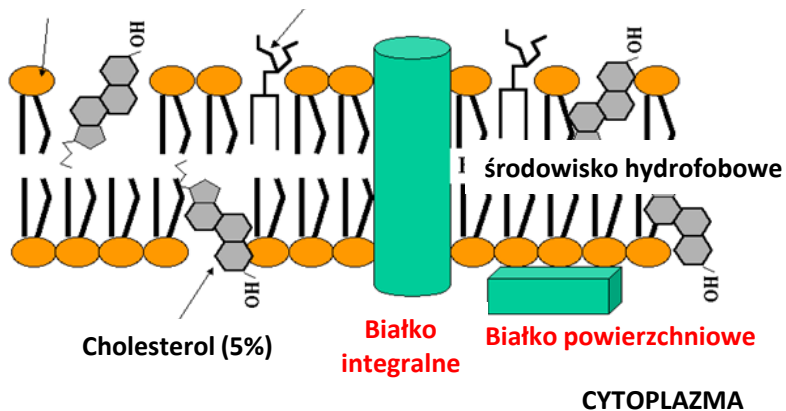
micelle



Fosfolipidy (75%)

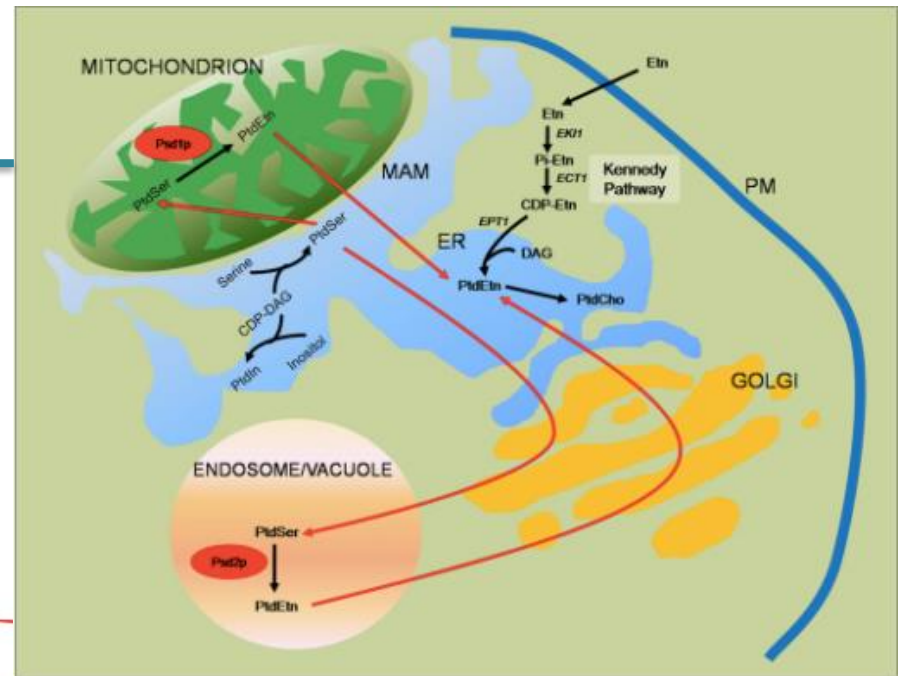
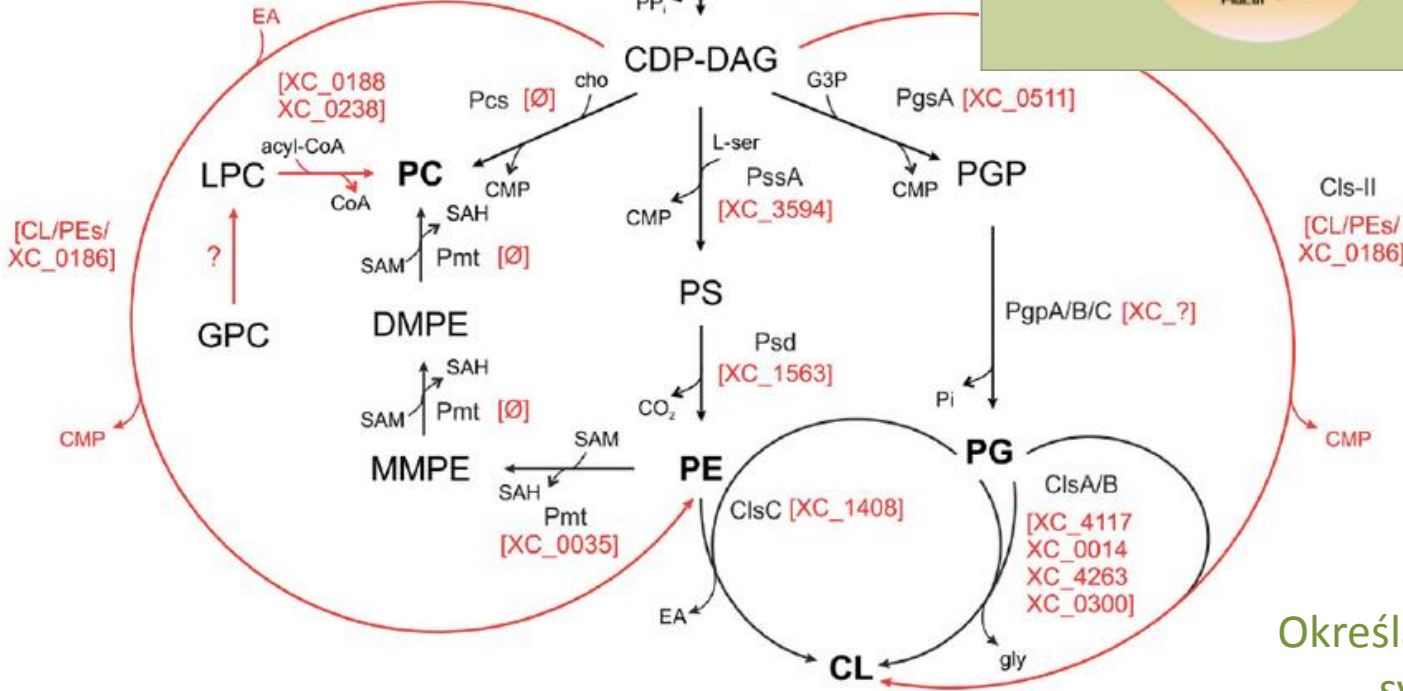
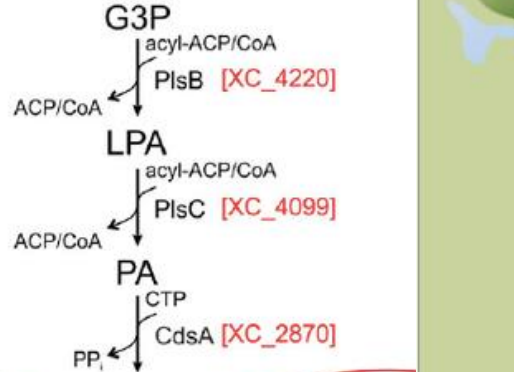
Glikolipidy (5%)

PŁYN POZAKOMÓRKOWY



# Biosynteza fosfolipidów

AAALifeSci



Określa i zmienia specyficzną sygnaturę lipidomu

# Zmiany w lipidomie

---

## Choroby

- upośledzenie szlaków metabolicznych
- Oksydacyjne modyfikacje lipidów

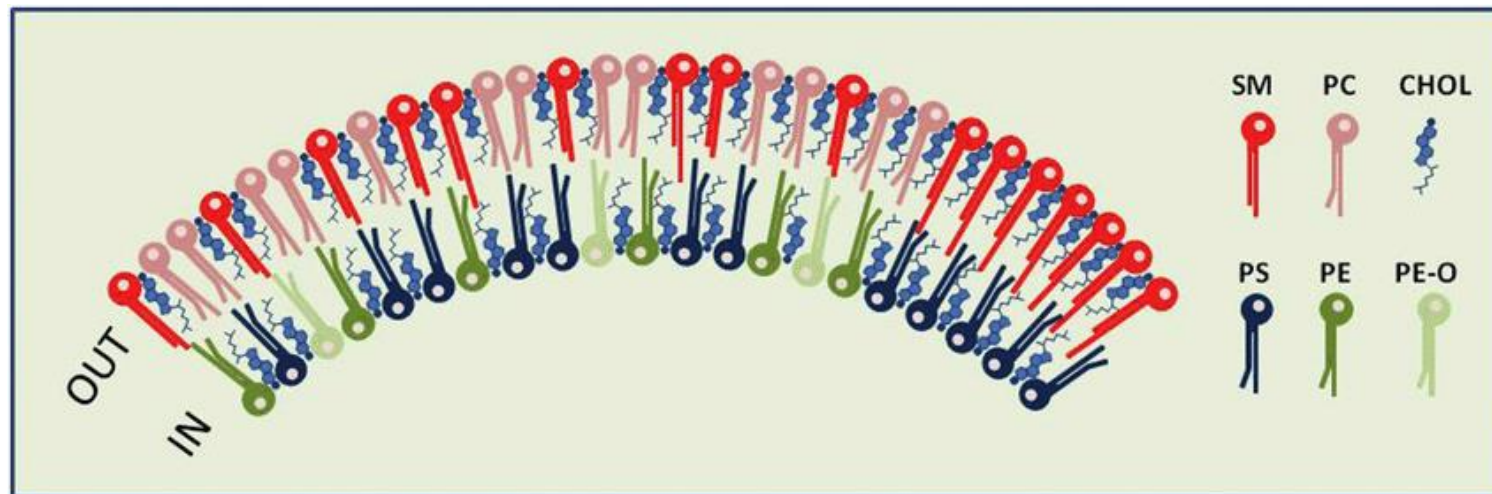
Inne:

- **Dieta** – źródło różnych lipidów

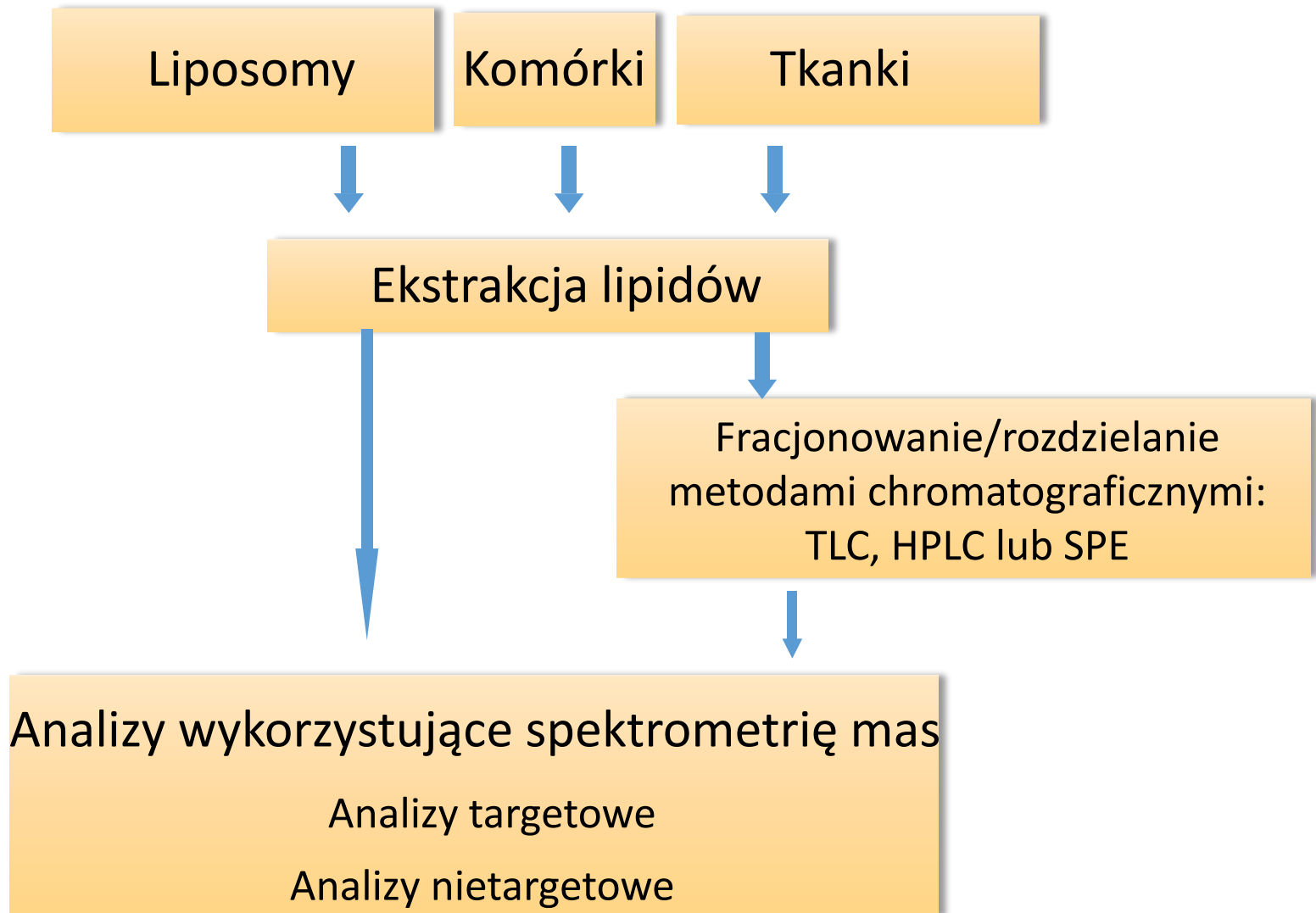
Znaczenie:

- nowe biomarkery
- nowe strategie leczenia
- nowe zastosowania biotechnologiczne

# Lipidomiczne strategie analityczne w celu określania lipidomu

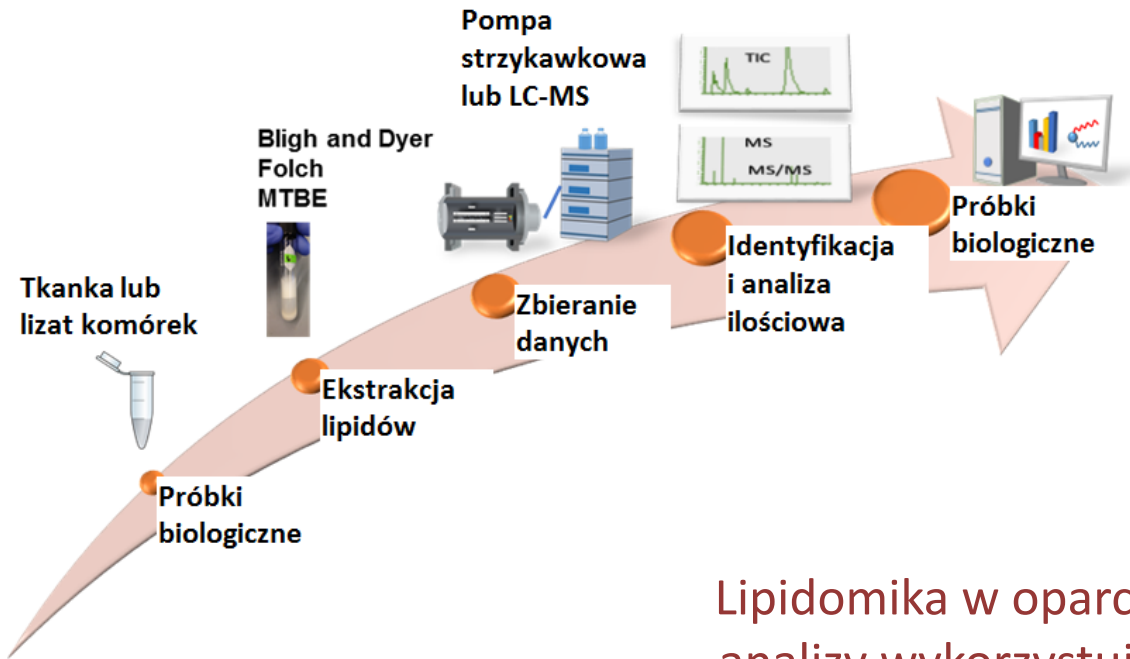


# Podjęcia lipidomiczne





# Procedura analityczna w lipidomice



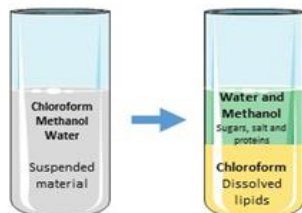
Lipidomika w oparciu o analizy wykorzystujące spektrometrię mas (MS)



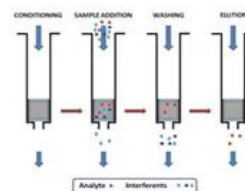
Próbki biologiczne

Ekstrakcja lipidów i frakcjonowanie

Ekstrakcja lipidów methanol/chloroform



Rozdział na frakcje



lub/i

Rozdział lipidów na klasy

Chromatografia cienkowarstwowa



TLC  
2D-TLC

Chromatografia cieczowa



HPLC | RP  
nLC | NP  
UPLC | HILIC

Analiza MS

Spektrometria mas – zbieranie danych

Jonizacja próbeki

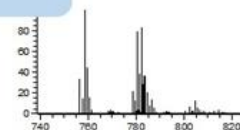
ESI  
MALDI  
APCI/APPI  
SIMS

Analiza jonów

Low mass resolution  
High mass resolution

Detekcja

TIC/RIC  
MS  
MS/MS



AAALifeSci

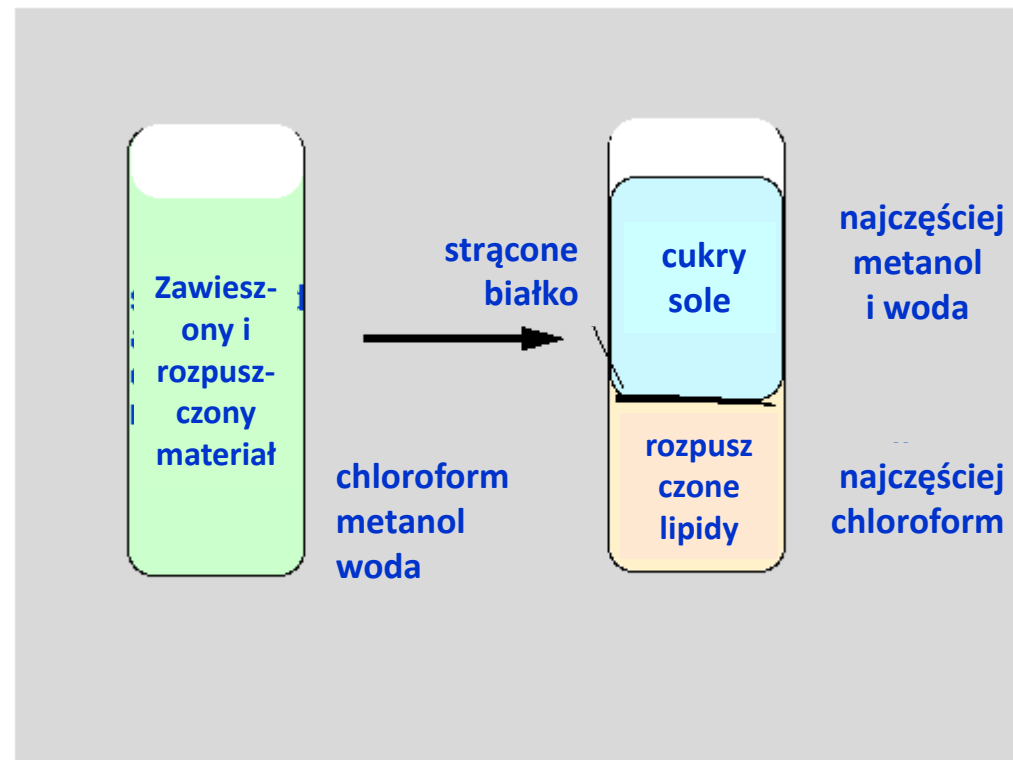


universidade de aveiro

# Ekstrakcja lipidów

Ekstrakcja chemiczna przy wykorzystaniu rozpuszczalników organicznych:

- ❑ Metoda Folcha  
( $\text{CHCl}_3:\text{CH}_3\text{OH}$  2:1)
- ❑ Metoda Bligh i Dyer  
( $\text{CHCl}_3:\text{CH}_3\text{OH}$  1:2)
- ❑ inne

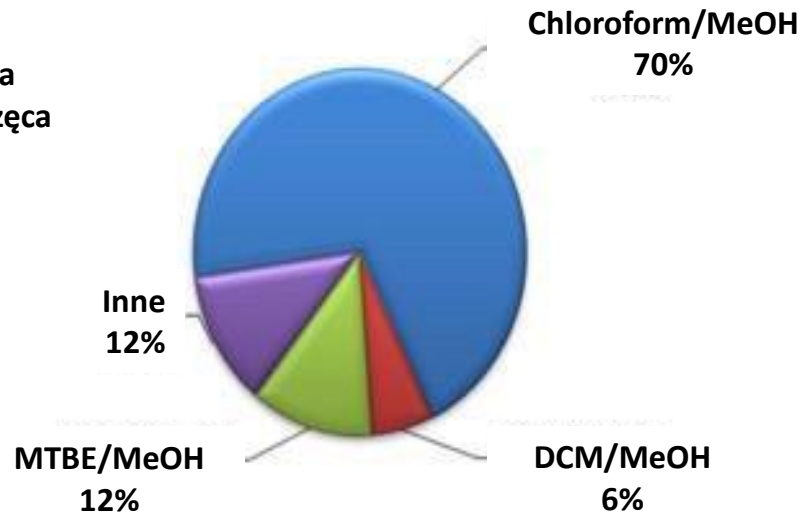


# Ekstrakcja lipidów

Analizowana matryca

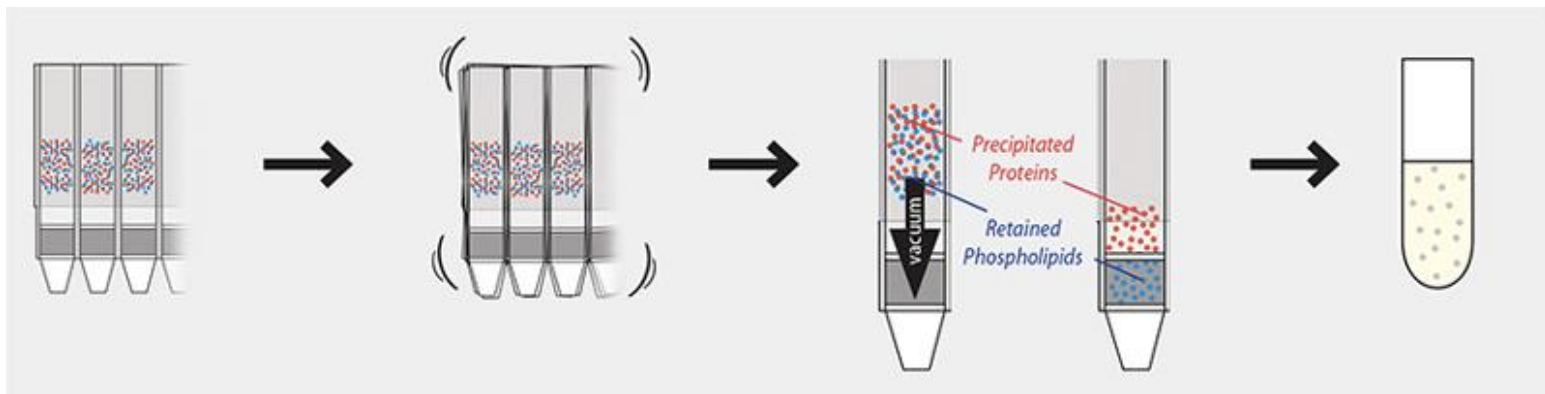
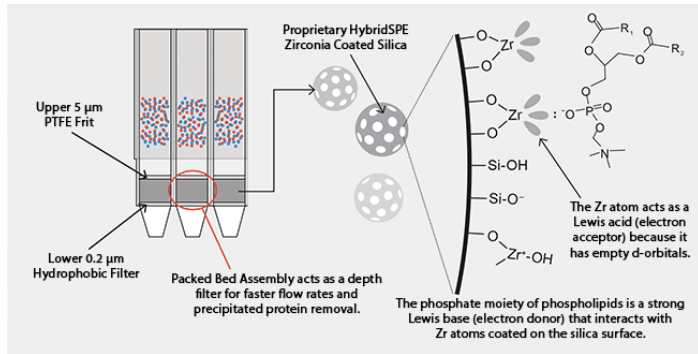


Metoda ekstrakcji



# Ekstrakcja lipidów

## Selektywna ekstrakcja lipidów z osocza przy użyciu hybrydowych kolumnienek SPE

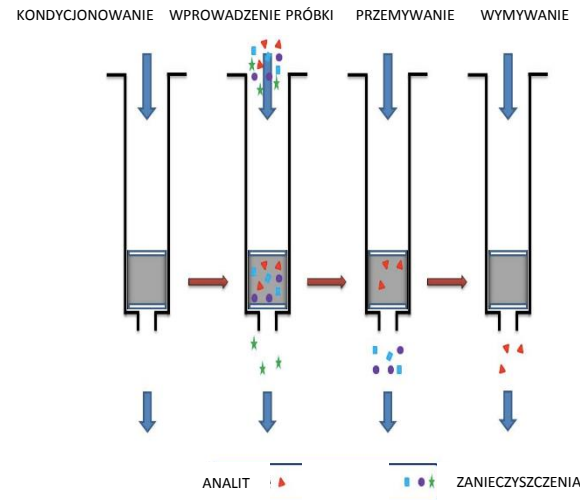


# Frakcjonowanie ekstraktu lipidów

## Ekstrakcja do fazy stałej

W celu oddzielenia obojętnych lipidów  
od polarnych

Obojętne lipidy (TG) od PL

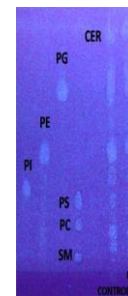


## Metody chromatograficzne

TLC (Chromatografia cienkowarstwowa)

HPLC (Wysokosprawna chromatografia cieczowa)

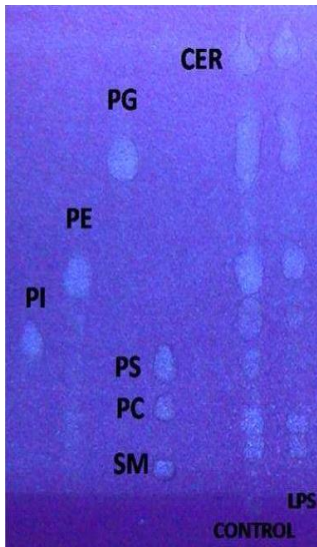
W celu rozdzielenia klas lipidów/związków lipidowych



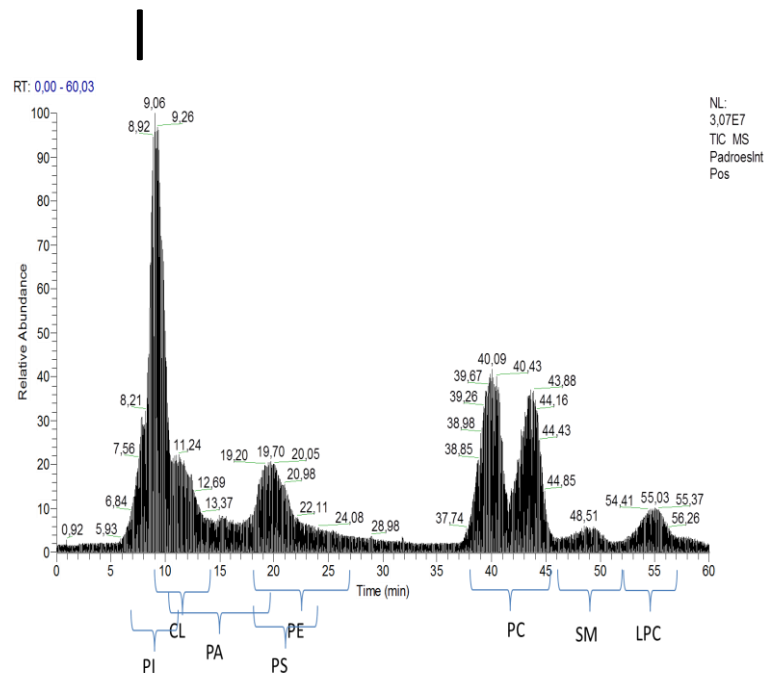
# Rozdzielanie klas fosfolipidów

Fosfolipidy można rozdzielić w zależności od ich polarności przy użyciu:

TLC

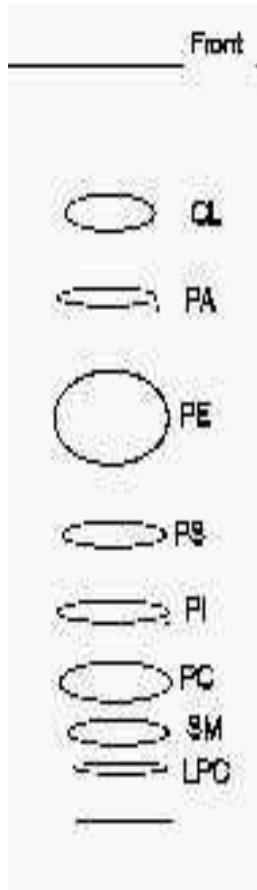


HPLC



# TLC – Chromatografia cienkowarstwowa

## Różne układy eluentów – różne profile TLC



CL-Cardiolipiny

PA-Kwas fosfatyglowy

PE-Fosfatydyloetanolaminy

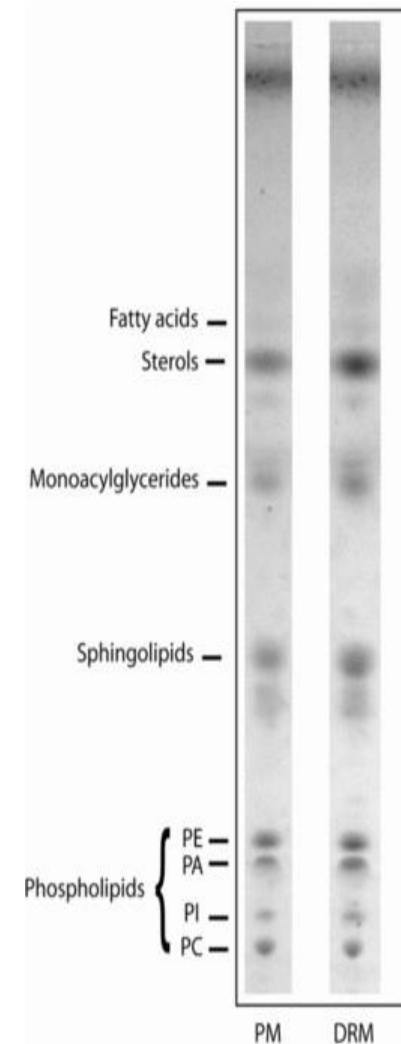
PS-Fosfatydyloseryny

PI-Fosfatydyloinozytol

PC-Fosfatydylocholiny

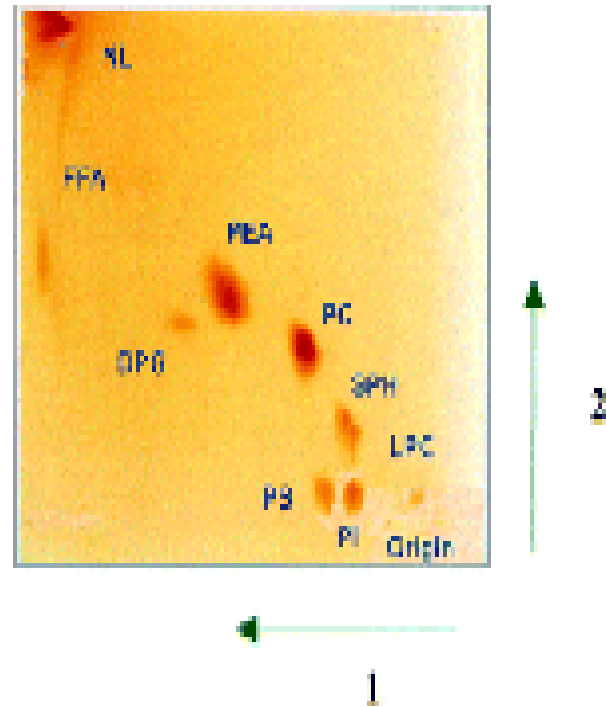
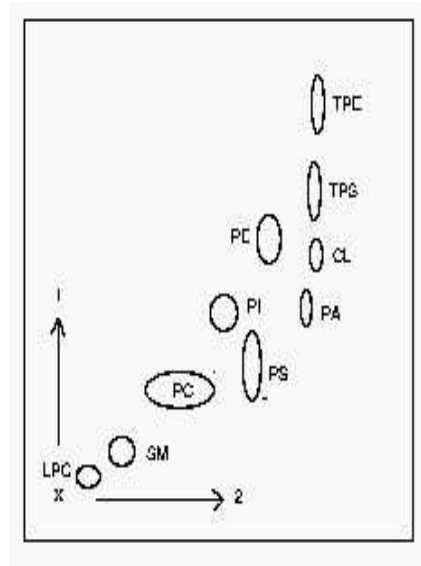
SM-Sfingomieliney

LPC-Lizofosfatydylocholiny





# 2D-TLC

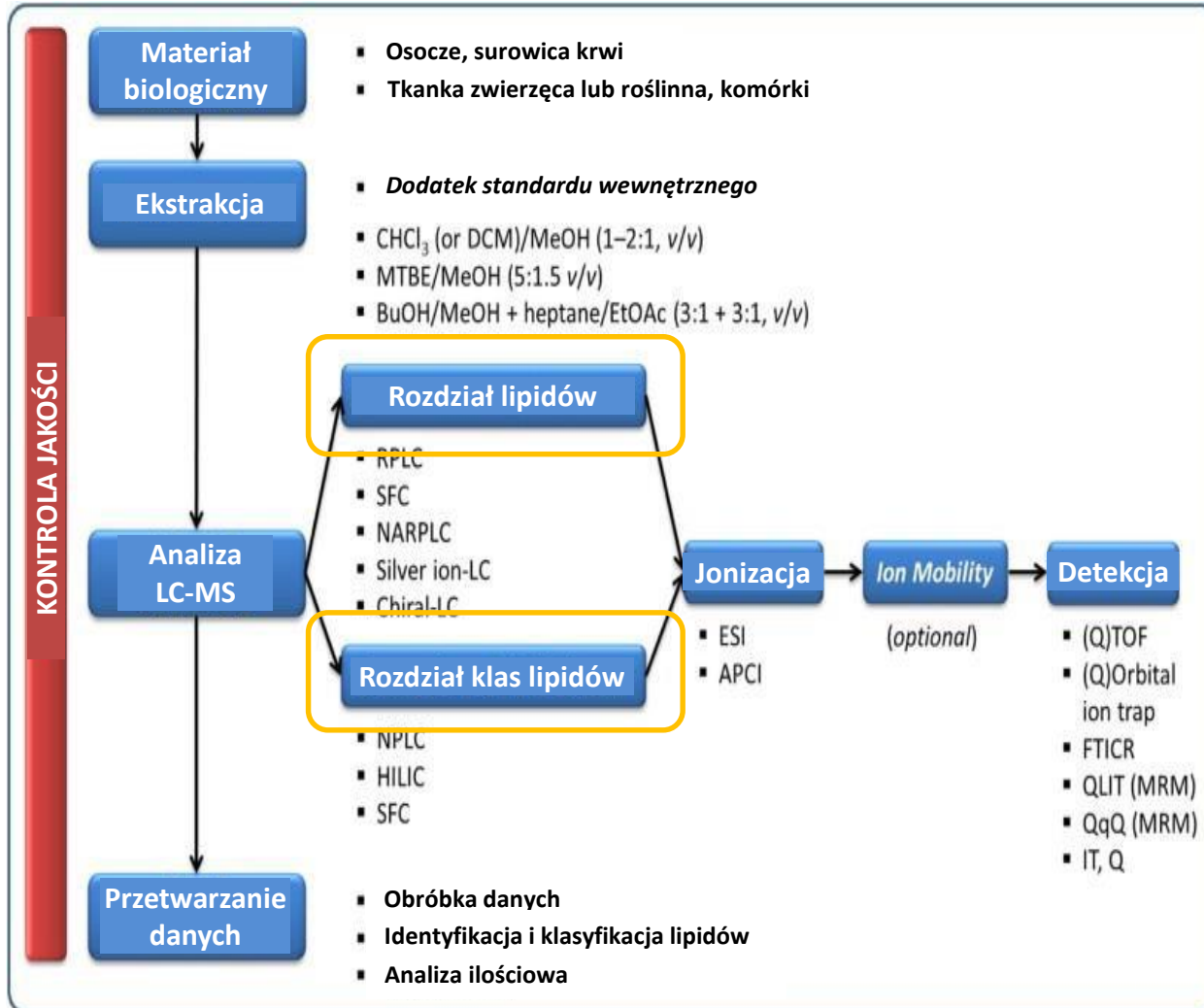
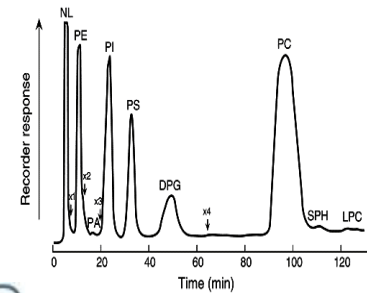


Dwa różne układy rozpuszczalników

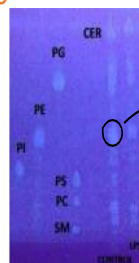
Prof Valerian Kagan lab



# HPLC-MS



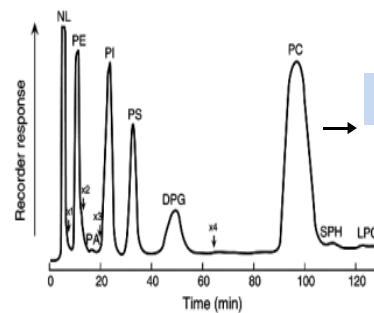
## off-line TLC-MS



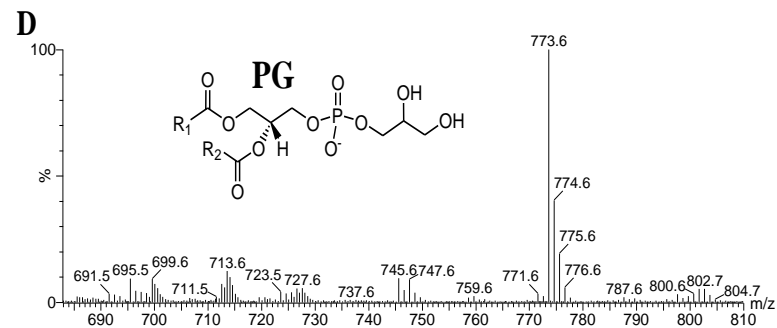
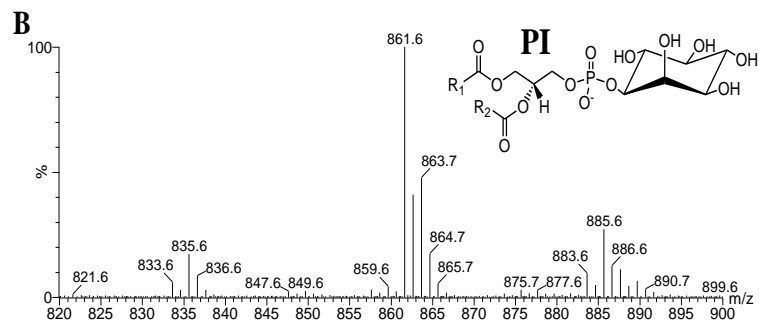
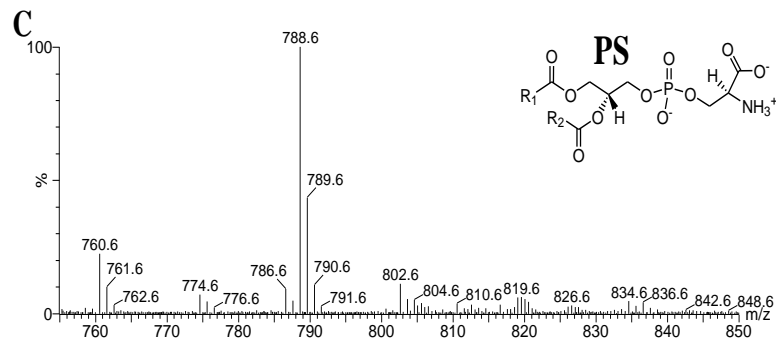
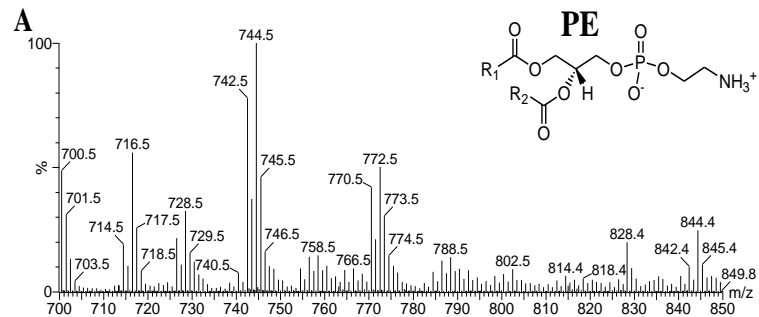
Ekstrakcja lipidów

Analiza MS

## on-line HPLC-MS



Analiza MS

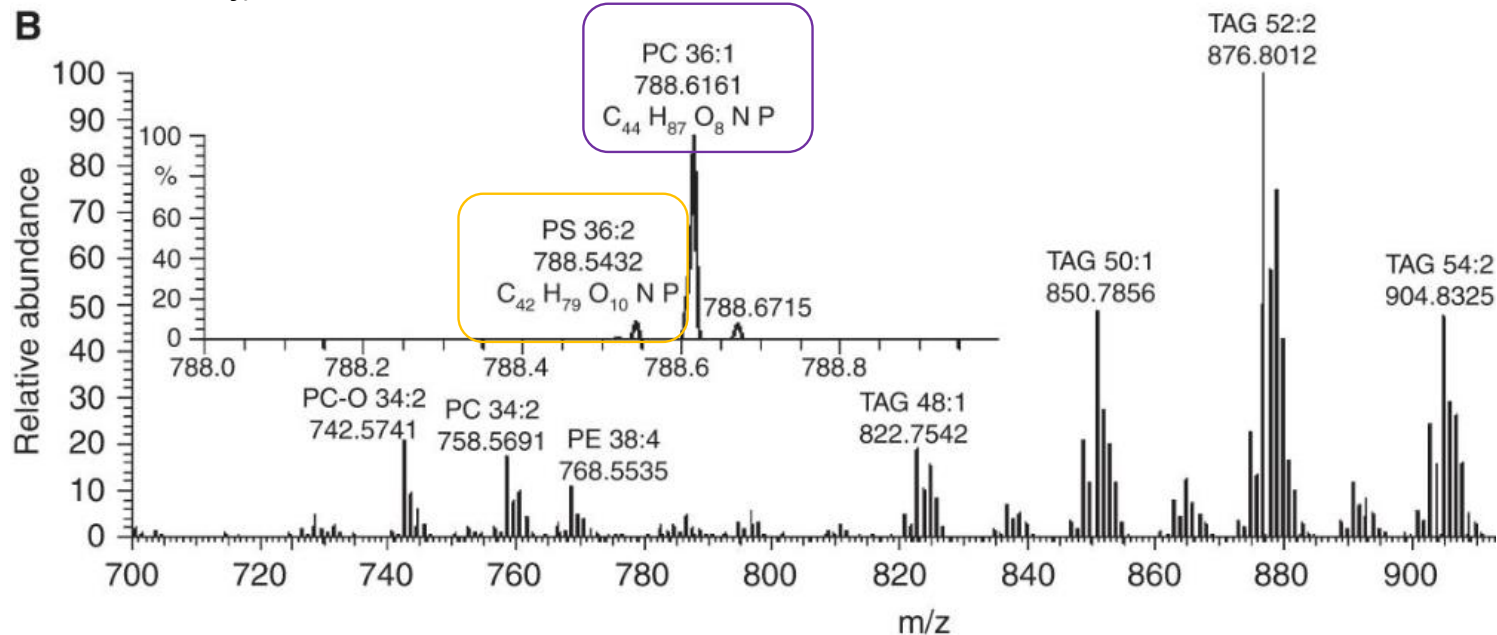


# ANALIZA DANYCH MS:

## Jony powstające w wyniku procesu jonizacji lipidów

Klasa lipidów	Tryb pozytywny	Tryb negatywny
LPC, PC	$[M + H]^+$ , $[M + Na]^+$	$[M - H]^-$ , $[M + HCOO]^-$ , $[M + CH_3COO]^-$
<b>LPE, PE</b>	<b><math>[M + H]^+</math>, <math>[M + Na]^+</math></b>	<b><math>[M - H]^-</math></b>
PG	$[M + H]^+$ , $[M + NH_4]^+$ , $[M + Na]^+$	$[M - H]^-$
PI	$[M + H]^+$ , $[M + NH_4]^+$ , $[M + Na]^+$	$[M - H]^-$
<b>PS</b>	<b><math>[M + H]^+</math></b>	<b><math>[M - H]^-</math></b>
PA		$[M - H]^-$
CE	$[M + NH_4]^+$ , $[M + Na]^+$	
SM	$[M + H]^+$	$[M + HCOO]^-$ , $[M + CH_3COO]^-$
Cholesterol	$[M - H_2O + H]^+$	
MG, DG, TG	$[M + NH_4]^+$ , $[M + Na]^+$	
MGDG, DGDG, SQDG	$[M + NH_4]^+$ , $[M + Na]^+$	$[M - H]^-$
<b>Kwasy tłuszczowe</b>		$[M - H]^-$
CL	$[M + H]^+$ , $[M + NH_4]^+$ , $[M + Na]^+$	$[M - H]^-$ , $[M - 2H]^{2-}$
Cer, GluCer, LacCer	$[M + H]^+$ , $[M + NH_4]^+$ , $[M + Na]^+$	$[M - H]^-$ , $[M + HCOO]^-$ , $[M + CH_3COO]^-$

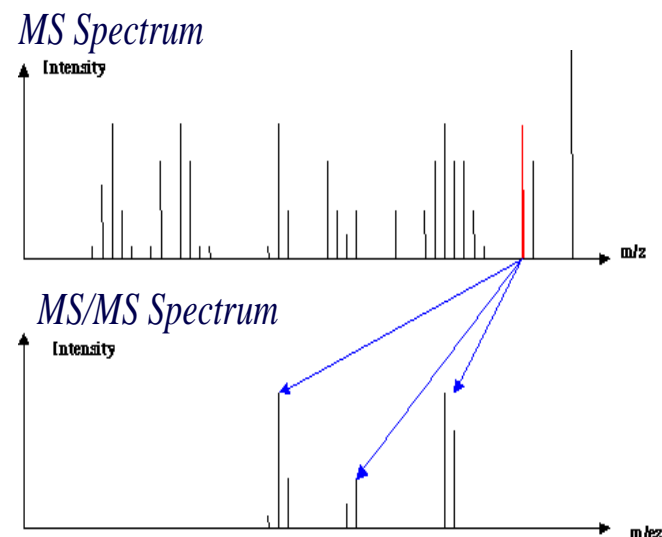
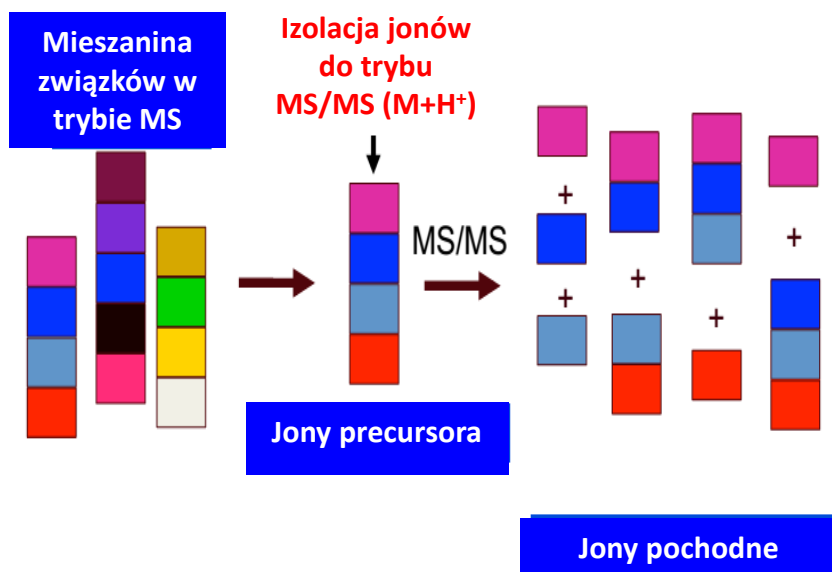
- Duża dokładność pomiaru masy wykorzystywana do:
  - ustalania masy cząsteczkowej
  - wyznaczania wzoru sumarycznego i składu elementarnego
  - ustalania struktury molekularnej
  - rozróżniania jonów molekularnych związków izobarycznych (o tej samej wartości stosunku  $m/z$  lecz różnym wzorze sumarycznym i strukturze molekularnej)



# Tandemowa spektrometria mas (MS/MS) analiza danych

## Fragmentacja:

- Izolacja wybranych jonów w trybie MS
- Wytworzenie jonów pochodnych w trybie MS/MS
- Informacja na temat budowy strukturalnej



Interpretacja widm MS/MS jest jak układanie puzzli



Pozwalają uzyskać informacje na temat budowy strukturalnej związków wyjściowych.

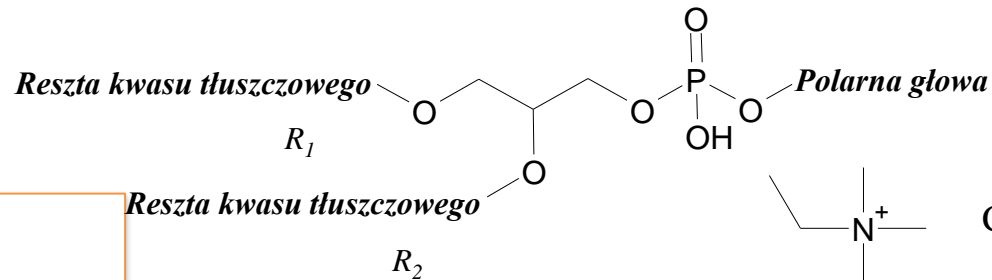
# Tandemowa spektrometria mas (MS/MS)

---

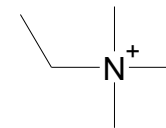
AAcLifeSci 



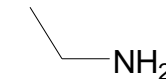
# Glicerofosfolipidy lub fosfolipidy (PL)



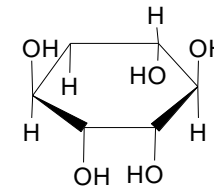
$R_1$  oraz  $R_2$  grupy alkilowe  
w pozycji *sn*-1 i *sn*-2



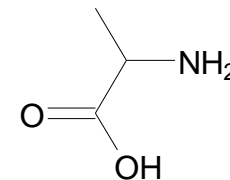
Cholina



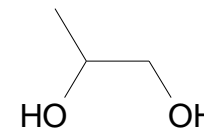
Etanolamina



Inozytol



Seryna



Glicerol

-H

Wodór

## Wymagane informacje dla

### potwierdzenia struktury:

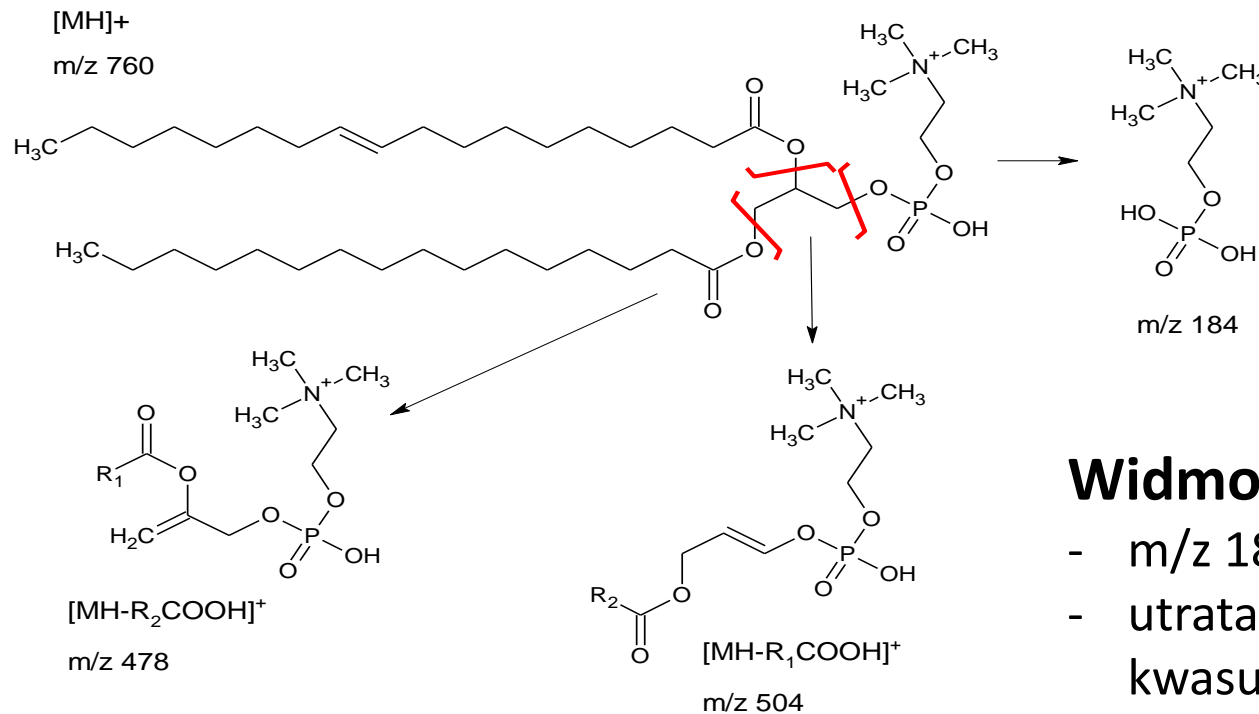
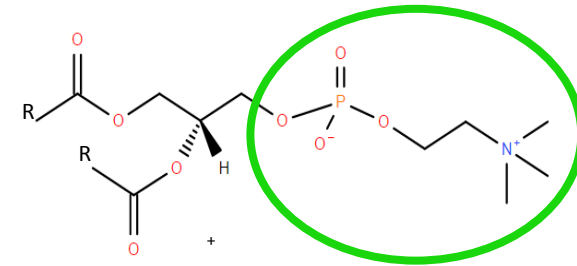
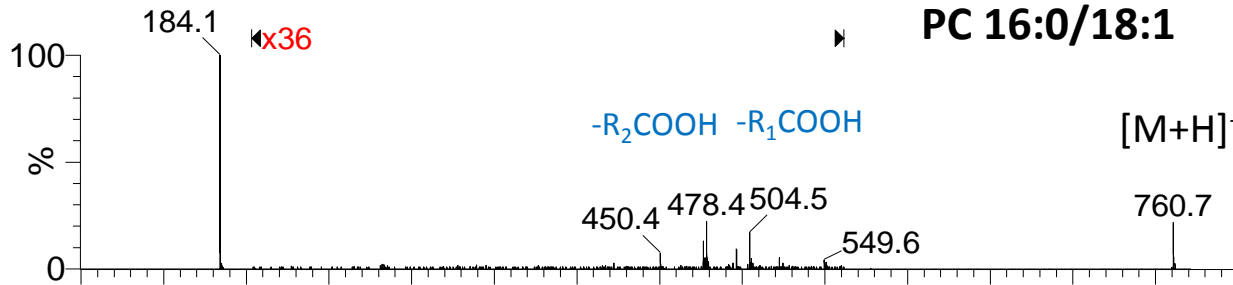
- identyfikacja polarnej głowy
- identyfikacja kwasów tłuszczowych

### Fragmentacja zależy od:

- rodzaju jonu prekursora
- energii kolizji
- innych czynników



# Fosfatydylocholiny – MS/MS $[M+H]^+$

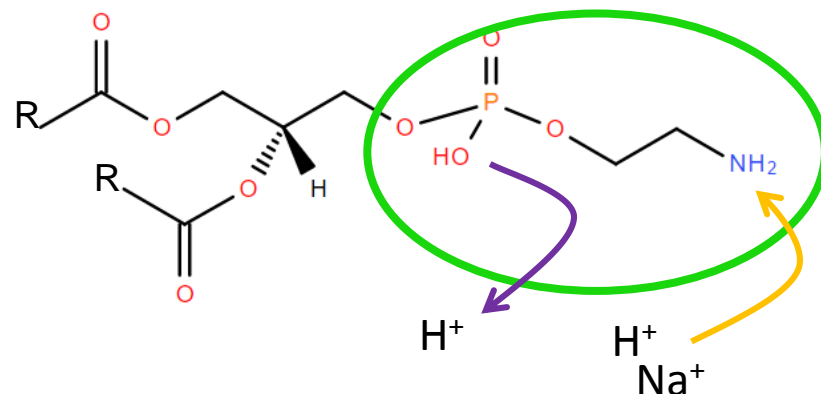


## Widmo MS/MS

- m/z 184
- utrata cząsteczki kwasu tłuszczowego

# Fosfatydyletanolaminy – PE

- Tryb jonizacji pozytywnej  $[M+H]^+$
- Tryb jonizacji negatywnej  $[M-H]^-$

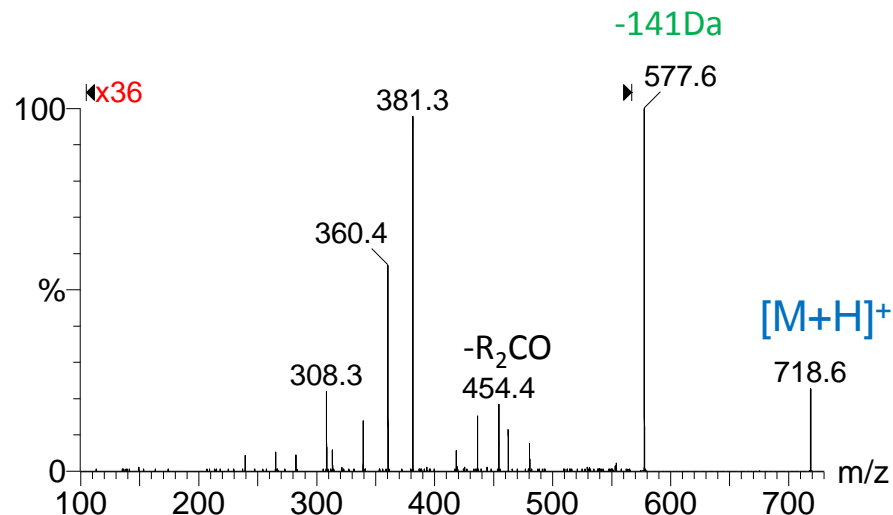


## ESI-MS/MS $[M+H]^+$

Utrata charakterystycznego fragmentu o masie **141 Da**

utrata  $RCOOH$  i  $RC=O$   
( $R_1=CO^+$  i  $R_2=CO^+$ ).

PE 16:0/18:1

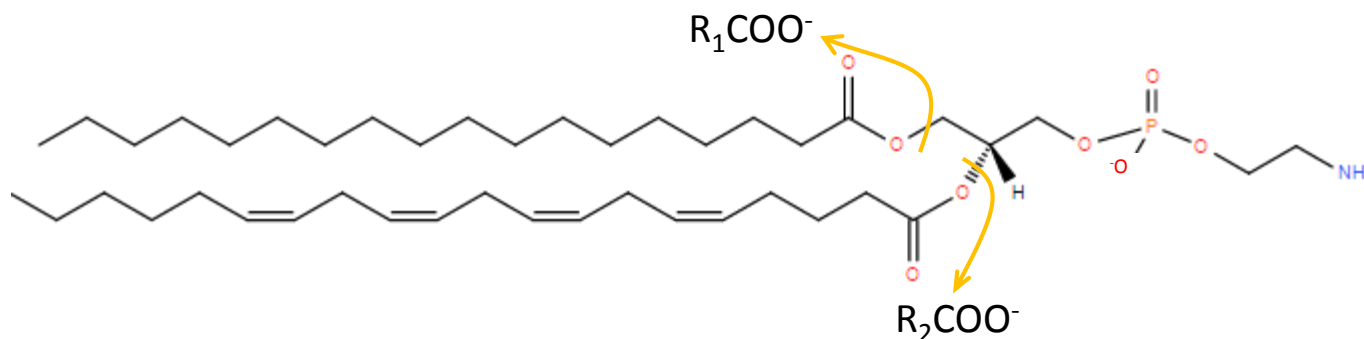
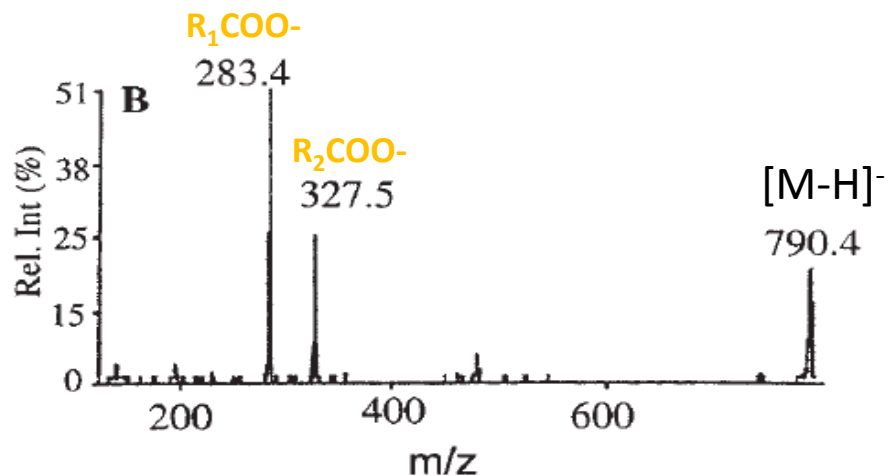


# Fosfatydyletanolaminy – MS/MS

## ESI-MS/MS $[M-H]^-$

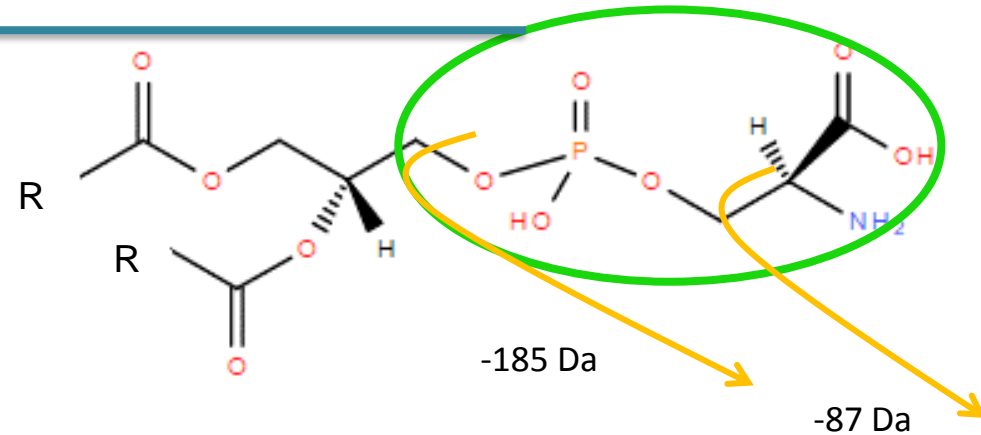
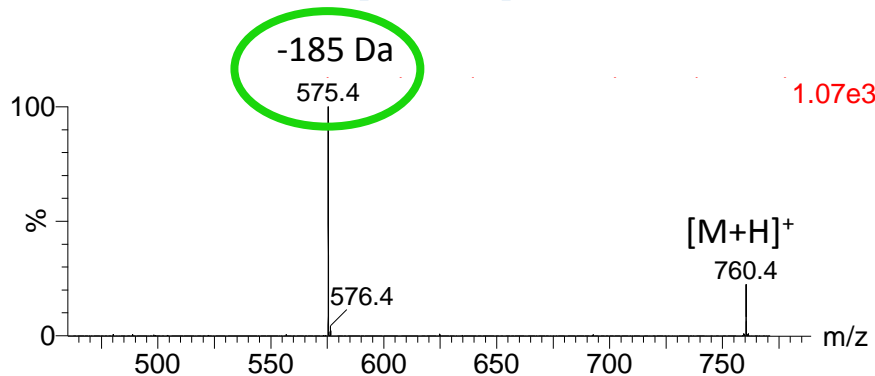
Widmo MS/MS w trybie negatywnej jonizacji

- jon  $R_1COO^-$
- jon  $R_2COO^-$



# Fosfatidylseryny – PS

## ESI-MS/MS $[M+H]^+$ POPS

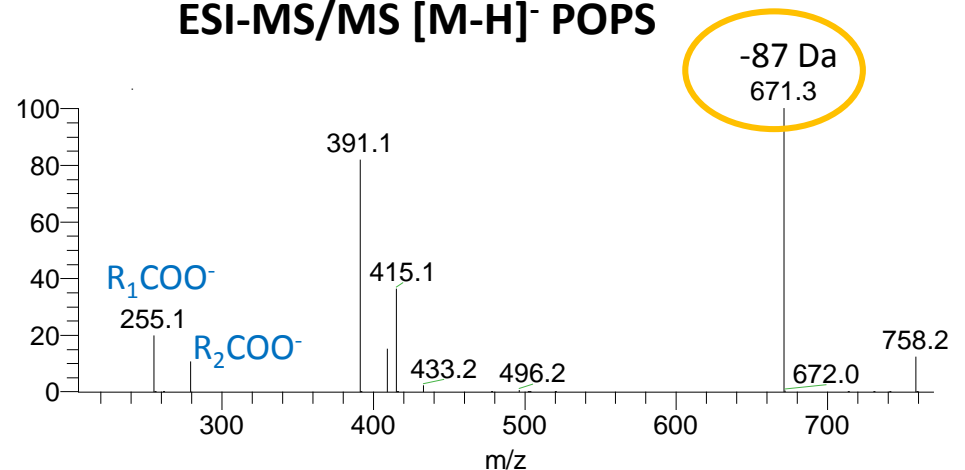


MS/MS  $[M+H]^+$  utrata charakterystycznego fragmentu o masie 185 Da

## ESI-MS/MS jonu $[M-H]^-$

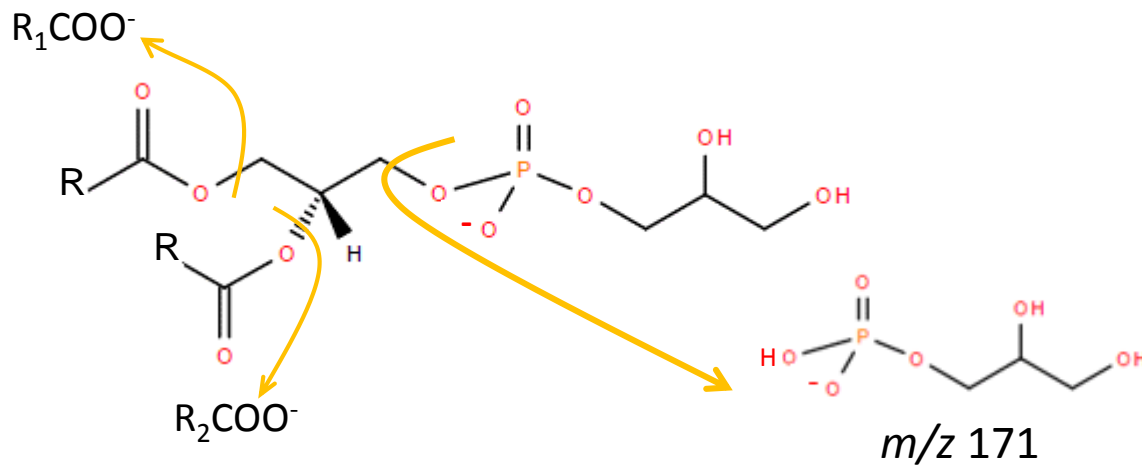
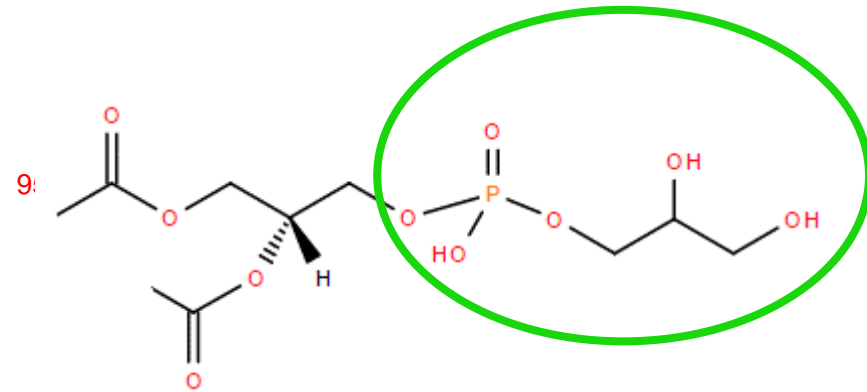
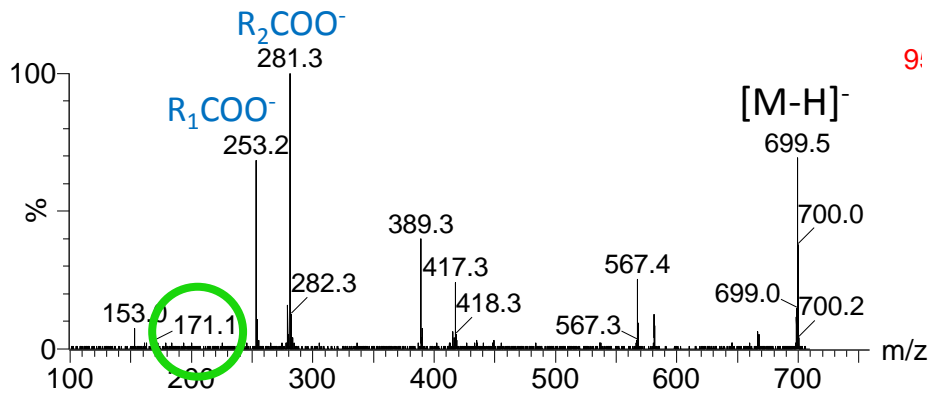
- utrata fragmentu o masie 87 Da
- jon  $R_1COO^-$  ion
- jon  $R_2COO^-$  ion

## ESI-MS/MS $[M-H]^-$ POPS



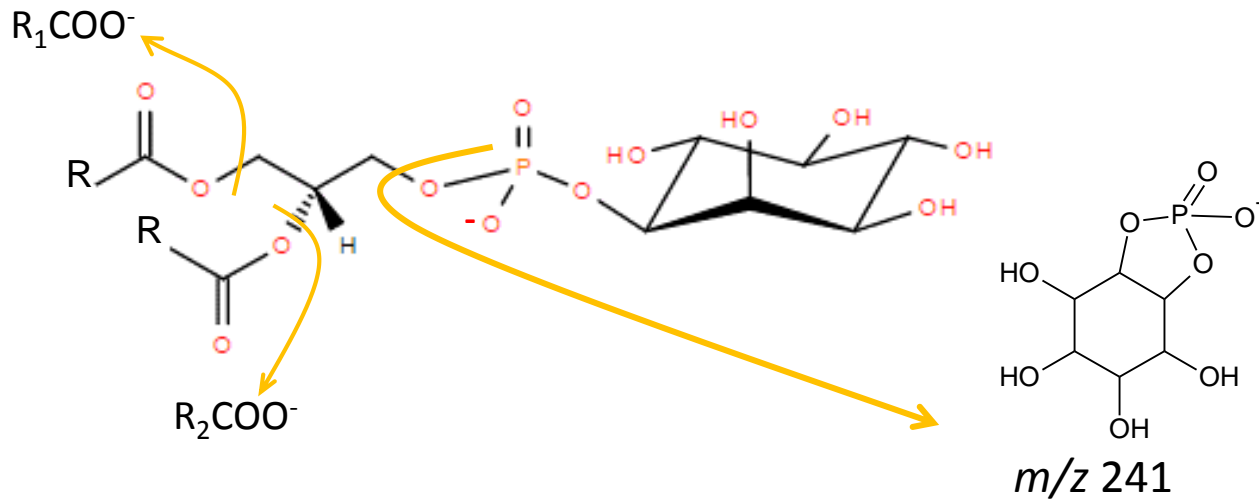
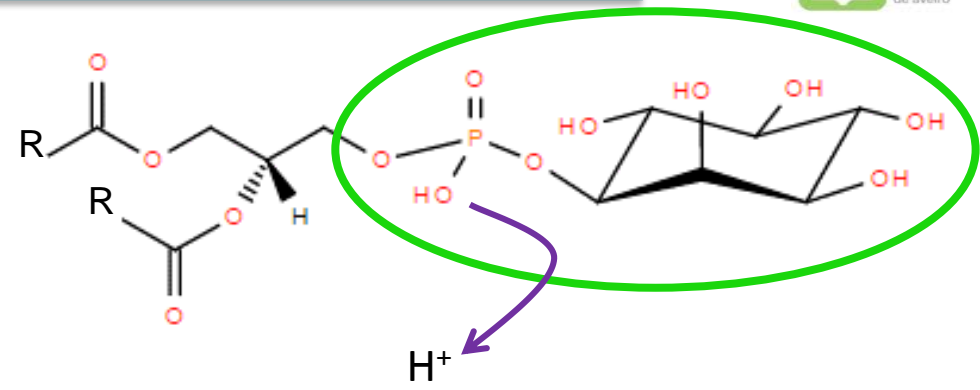
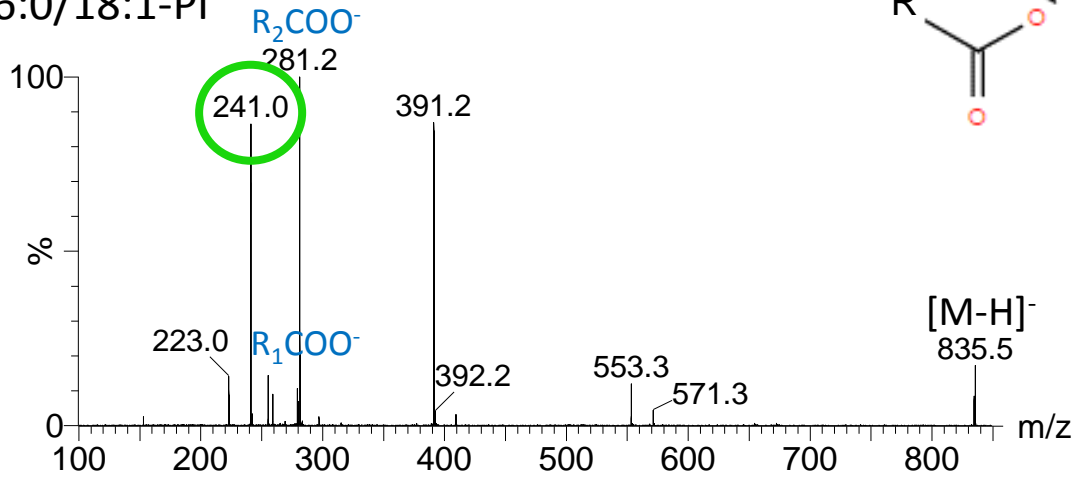
# Fosfatydyloglicerol- PG

Widmo ESI-MS/MS w trybie jonizacji negatywnej



# Fosfatidyloinozitol -PI

ESI-MS/MS of  $[M-H]^-$   
 16:0/18:1-PI



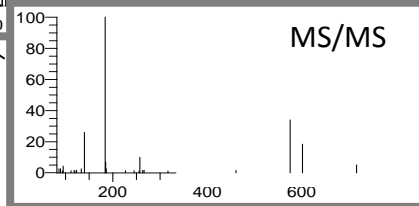
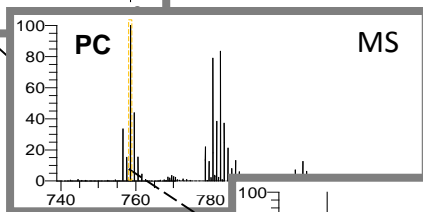
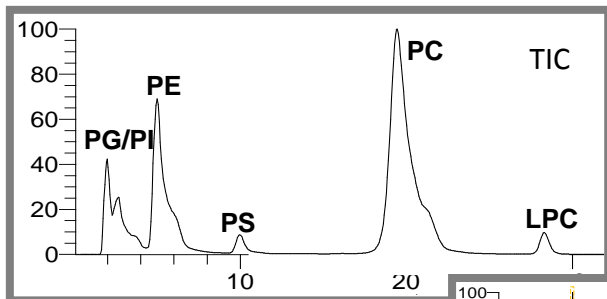
# Klasyfikacja fosfolipidów (MS/MS)

Klasa fosfolipidów	Wykrywane jony	Tryb pozytywny	Tryb negatywny
Fosfatydylocholiny (PC)	$[M+H]^+$	PIS m/z 184	-
Fosfatydyloetanolaminy (PE)	$[M+H]^+$ $[M-H]^-$	NLS 141 Da	-
Fosfatydyloseryny (PS)	$[M+H]^+$ $[M-H]^-$	NLS 185 Da	NLS 87 Da
Fosfatydyloglicerole (PG)	$[M-H]^-$	-	NLS 74 Da
Fosfatydyloinosytle (PI)	$[M-H]^-$	-	PIS m/z 241
Kardiolipiny (CL)	$[M-H]^-$ $[M_2H]^{2-}$		
Sfingomieliny (SM)	$[M+H]^+$	PIS m/z 184	-

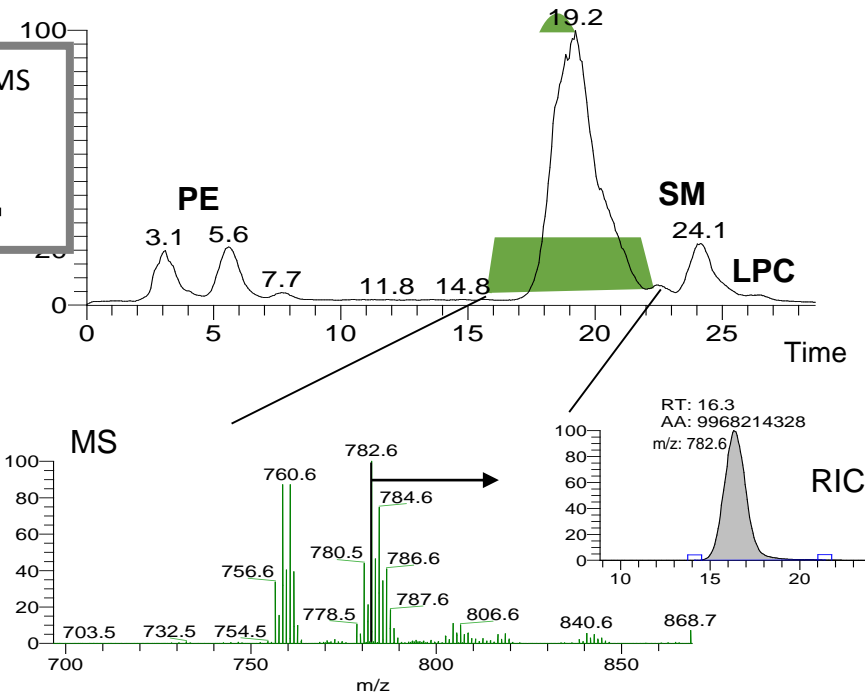
## Informacje dotyczące reszt kwasów tłuszczowych

- Tryb jonizacji pozytywnej
  - utrata  $RCOOH$  i  $R=C=O$
- Tryb jonizacji negatywnej
  - wytworzenie anionu karboksylowego  $RCOO^-$

# Nietargetowa analiza lipidomiczna



TIC  
Tryb jonizacji pozytywnej



Bioinformatyczne narzędzia:

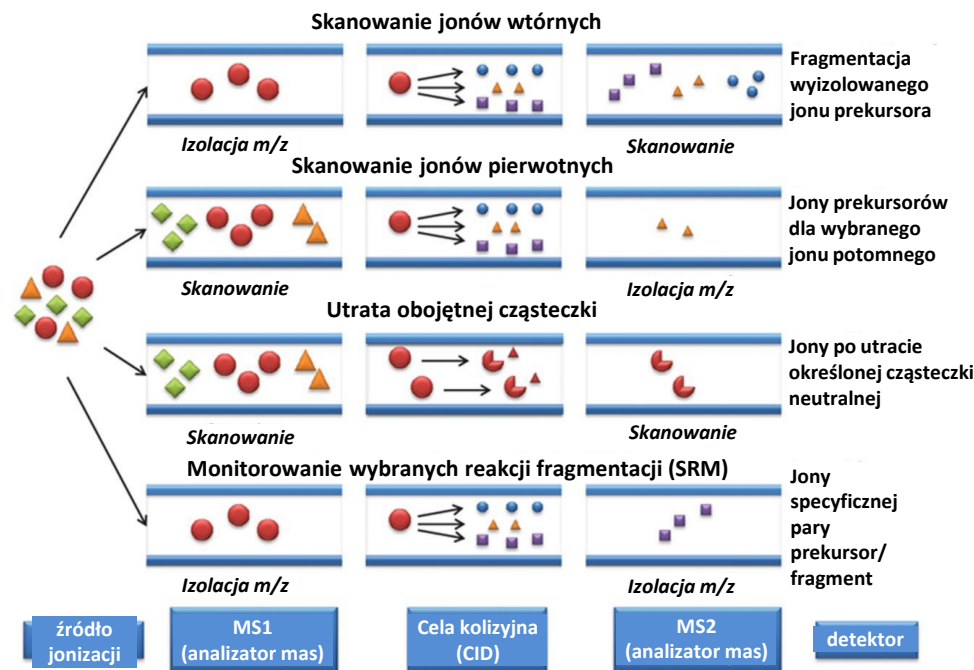
- Mzmine
- Compound Discover
- Lipidizer
- LipidBlast
- inne



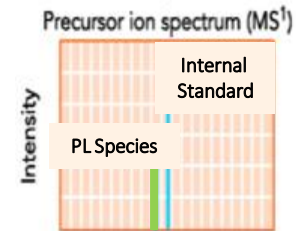
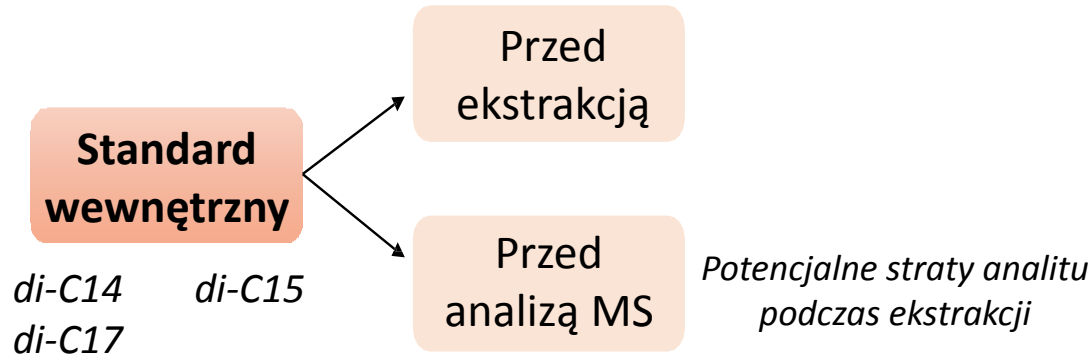
# Targetowa analiza lipidomiczna

## Lipidomika Shotgun

Klasa fosfolipidów	Tryb pozytywny	Tryb negatywny
Fosfatydylocholiny (PC)	PIS m/z 184	-
Fosfatydyloetanolaminy (PE)	NLS 141 Da	-
Fosfatydyloseryny (PS)	NLS 185 Da	NLS 87 Da
Fosfatydyloglicerole (PG)	-	NLS 74 Da
Fosfatydyloinosytole (PI)	-	PIS m/z 241
Kardiolipiny (CL)		
Sfingomieliny (SM)	PIS m/z 184	-



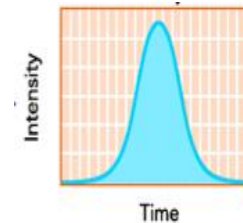
# Analiza ilościowa przy użyciu LC-MS



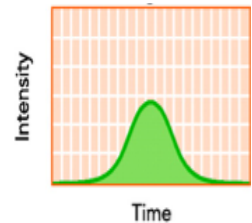
Chromatogram wybranych jonów

- Odpowiedź ilościowa jest inna zależnie od rodzaju związku lipidowego
- Inny standard wewnętrzny dla każdej klasy lipidów

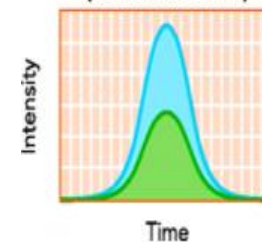
Standard wewnętrzny



Analit



Relative quantification  
(area under curves)



Normalizacja sygnału poszczególnych związków lipidowych do sygnału standardu wewnętrznego odpowiedniej klasy lipidów

---

Ten projekt został sfinansowany przy wsparciu Unii Europejskiej.

Niniejsza prezentacja odzwierciedla jedynie poglądy autorów, a Komisja nie ponosi odpowiedzialności za jakiegokolwiek wykorzystanie informacji zawartych w tej prezentacji



# Proteomika z wykorzystaniem spektrometrii mas

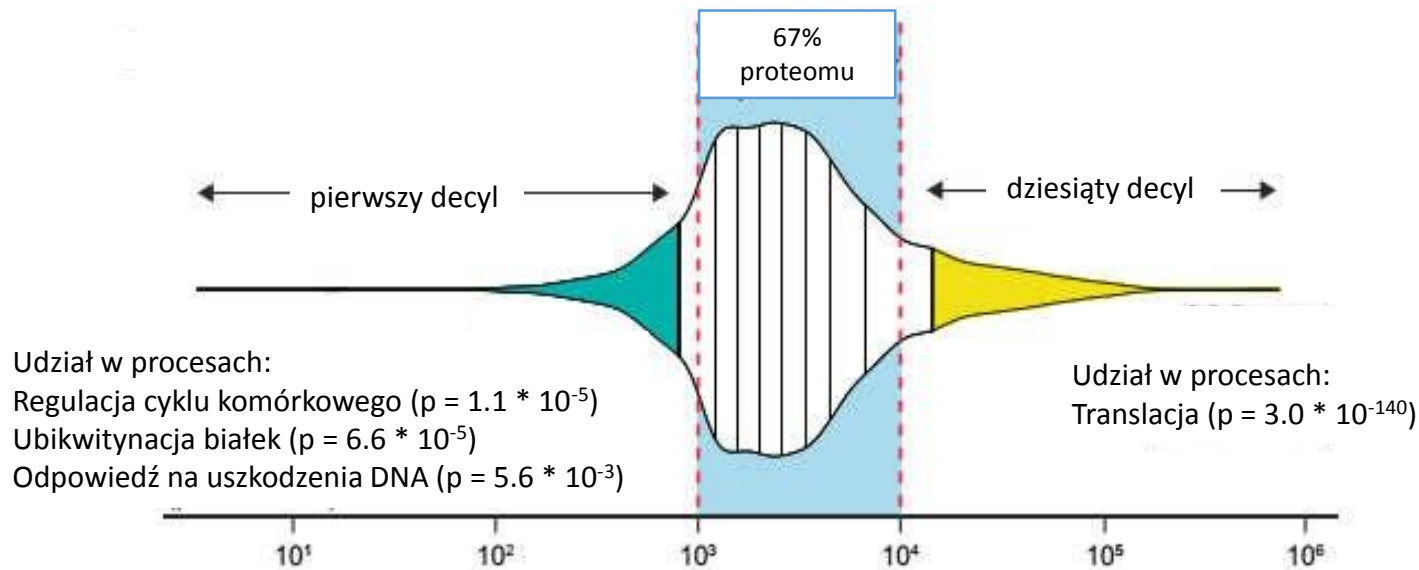
Pedro Domingues  
Rosário Domingues  
Rita Ferreira  
Tânia Melo  
Eliana Alves



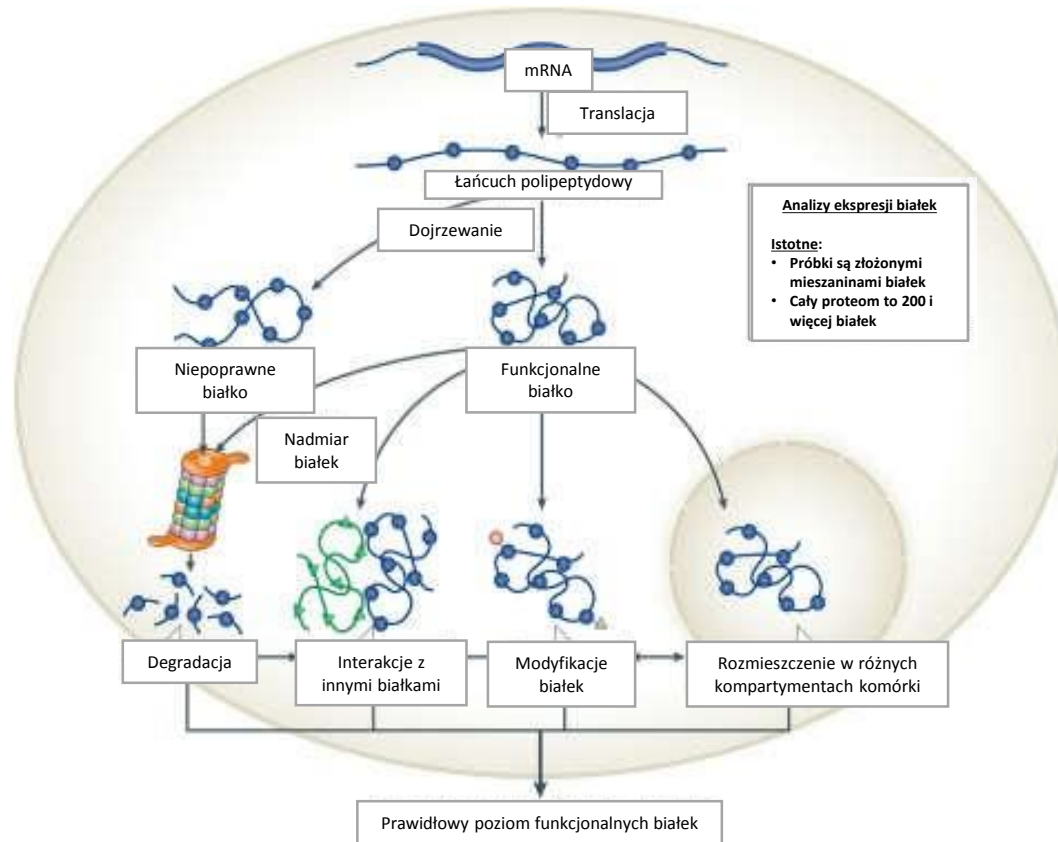
*S. Serevisiae*

Ilość białek (proteom): 5858

Całkowita ilość białek/komórkę:  $42 \times 10^6$



Mediana ilości białek (ilość cząsteczek/komórkę)



**Analizy ekspresji białek**

**Istotne:**

- Próbkki są złożonymi mieszaninami białek
- Cały proteom to 200 i więcej białek

Nature Reviews | Molecular Cell Biology

**Proteomika funkcjonalna**

**Istotne:**

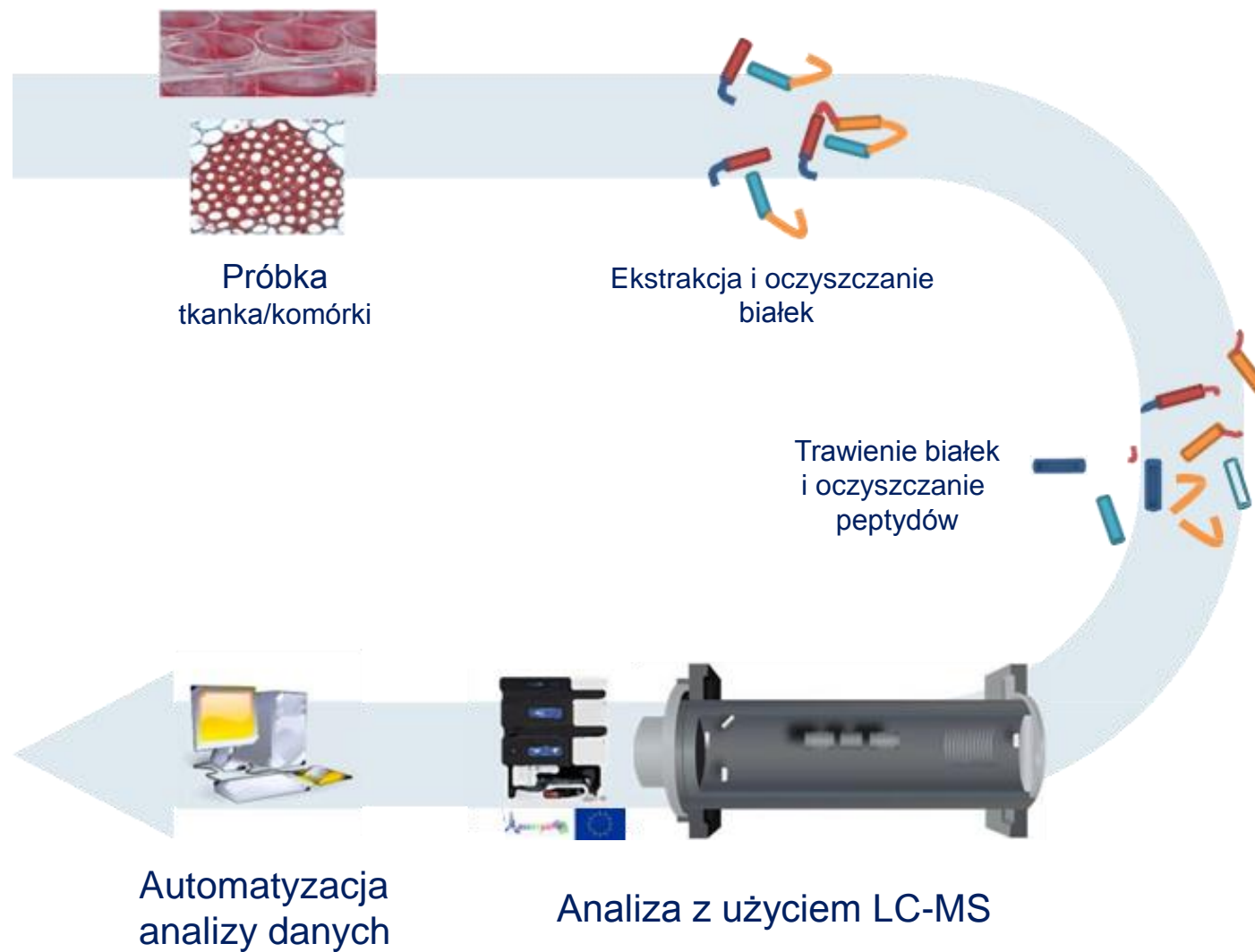
- Próbkki są złożonymi mieszaninami białek
- Frakcje białkowe
- Organella
- 1-200 białek

**Analizy PTM**

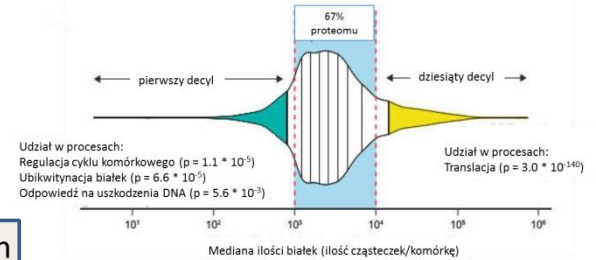
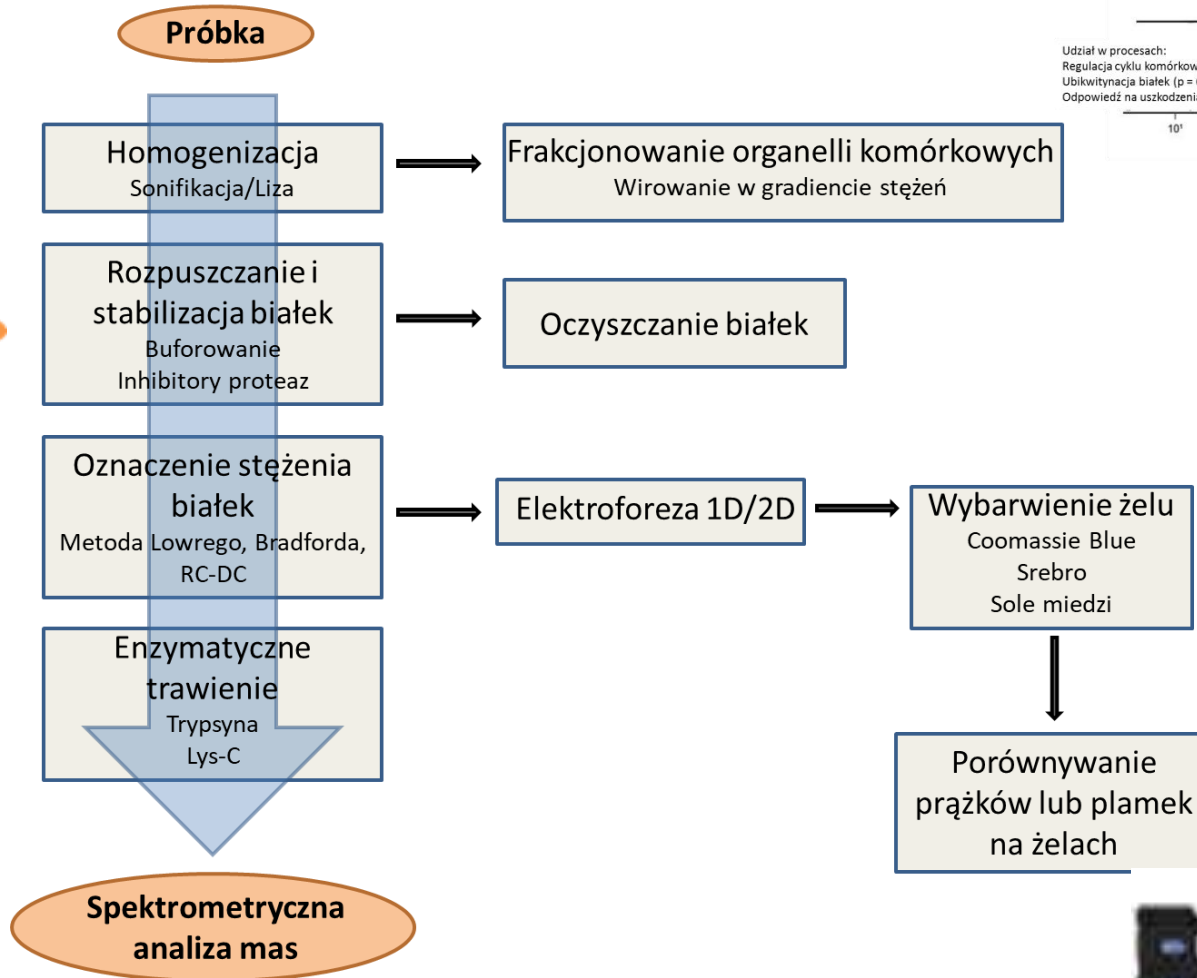
**Istotne:**

- Próbkki nie są złożone
- 1-20 białek

# Proteomika



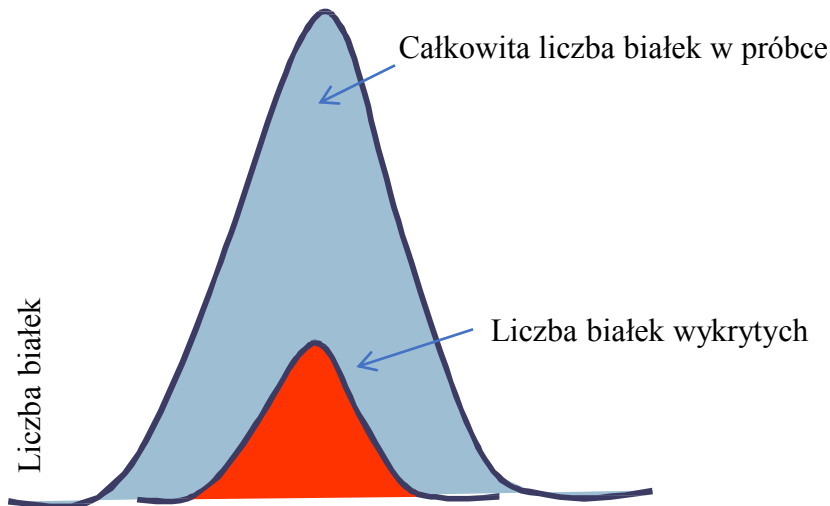
# Proteomika oparta na spektrometrii mas





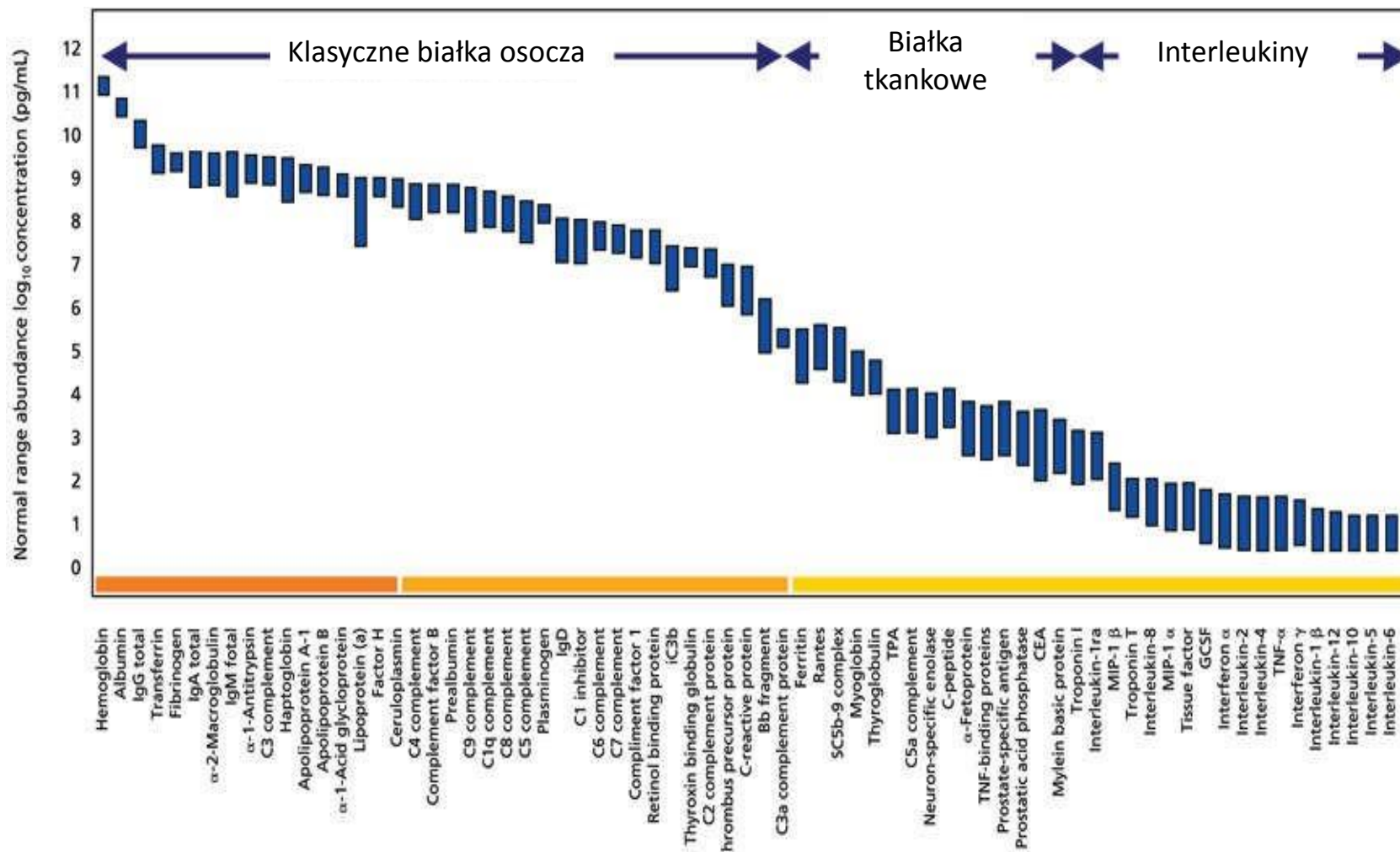
- **Czułość**
  - odpowiedź przyrządu/systemu pomiarowego
- **Limit detekcji**
  - Stosunek sygnału do szumu („signal to noise”, S/N)
  - Zależny od poziomu złożoności matrycy

W badaniach omicznych:

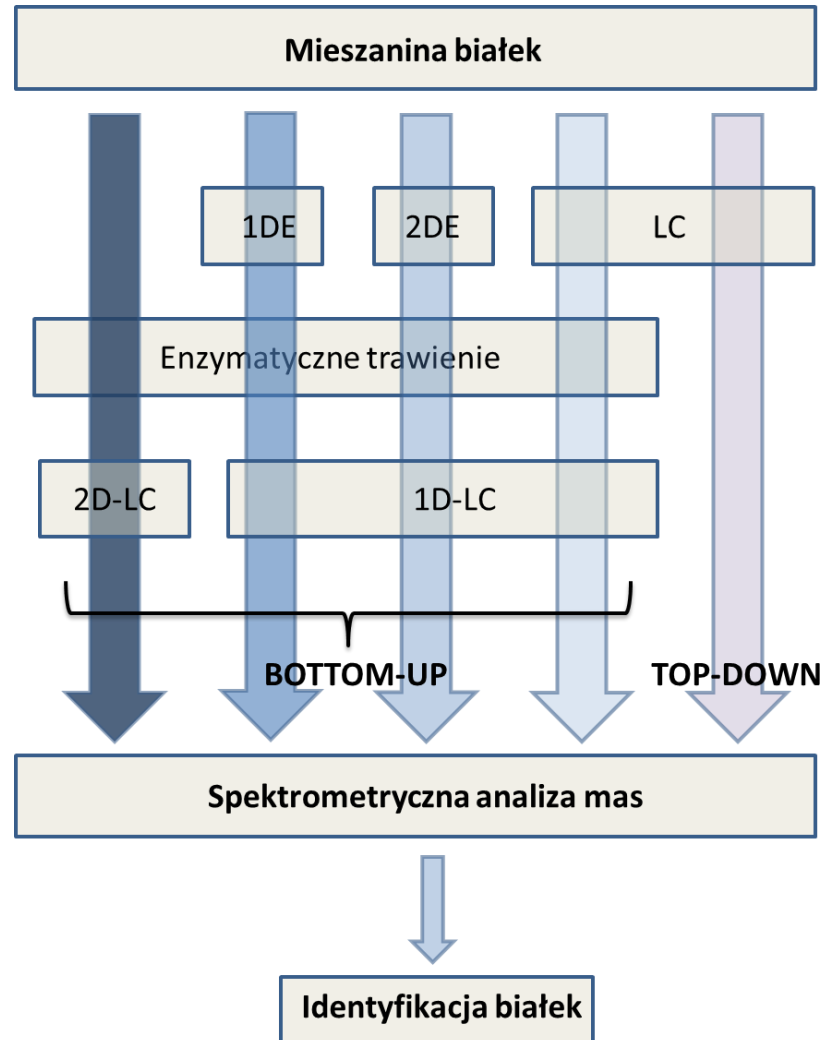


$$\text{Czułość} = \frac{\text{Liczba białek wykrytych}}{\text{Całkowita liczba białek w próbce}}$$

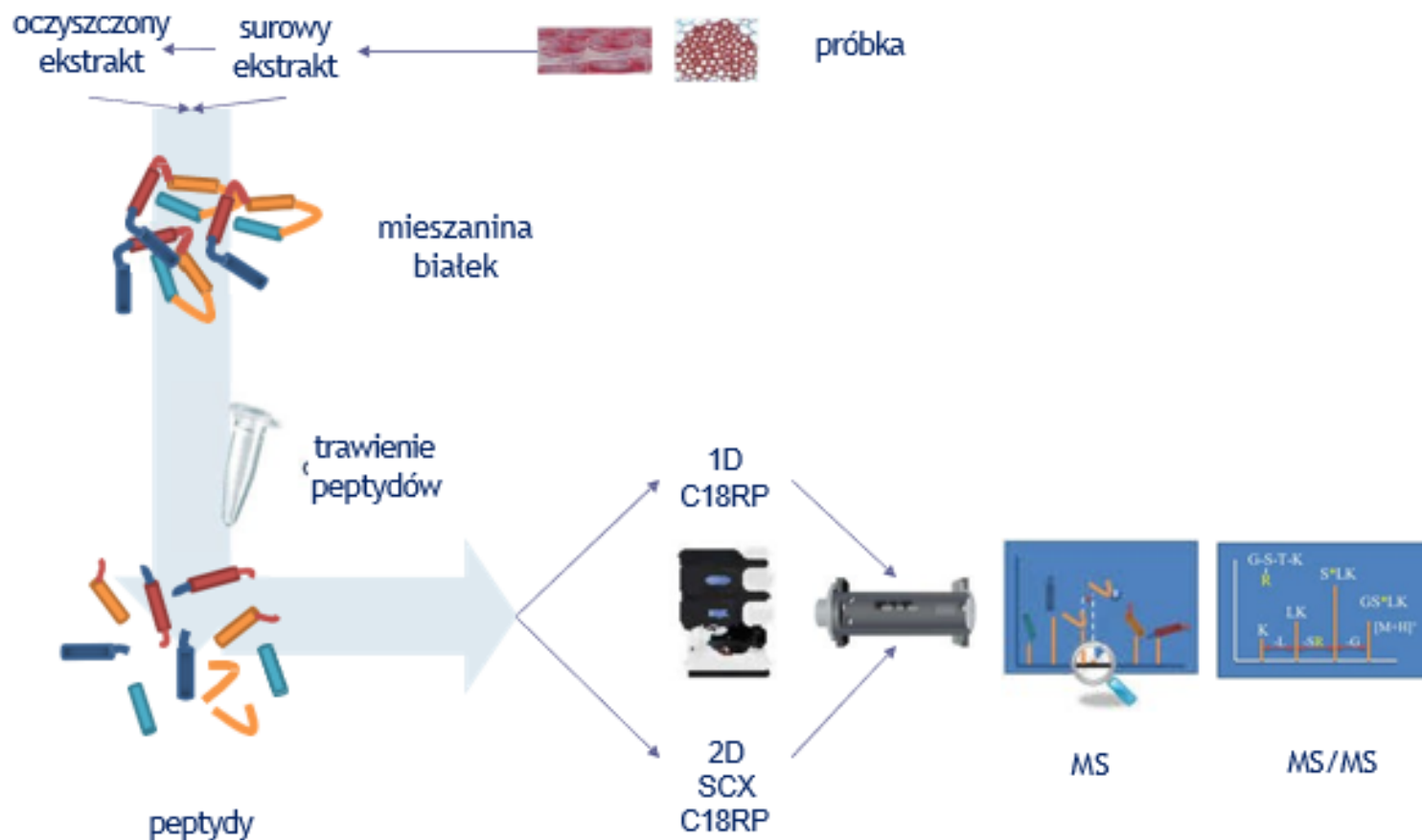
# Białka osocza



# Metody oparte na analizie sekwencji peptydów lub białek

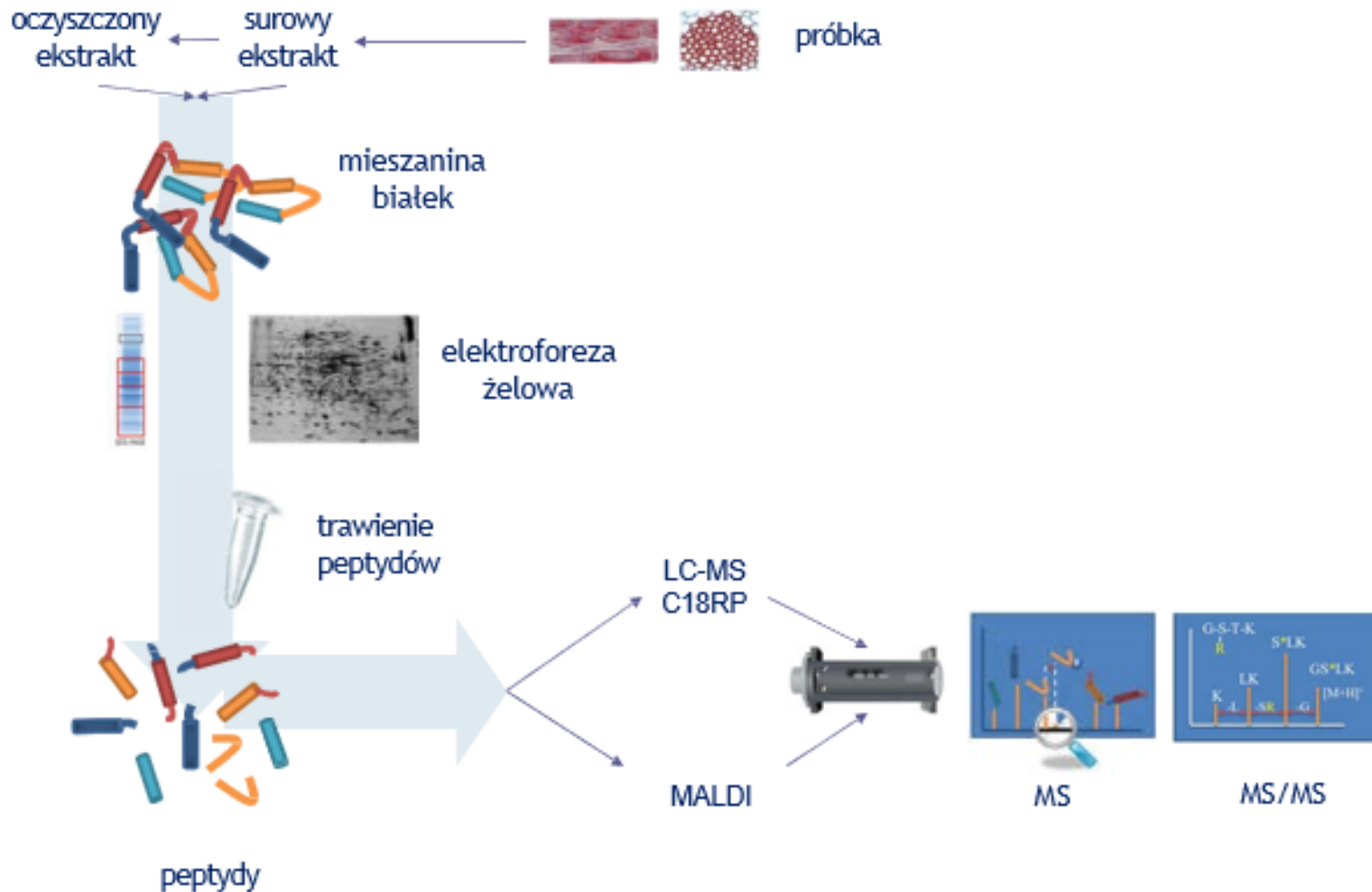


# Analiza peptydów (shotgun)



Wielowymiarowa technologia identyfikacji białek „MuDPIT”  
(Multidimensional Protein Identification Technology)

# Analiza peptydów połączona z rozdzielaniem elektroforetycznym

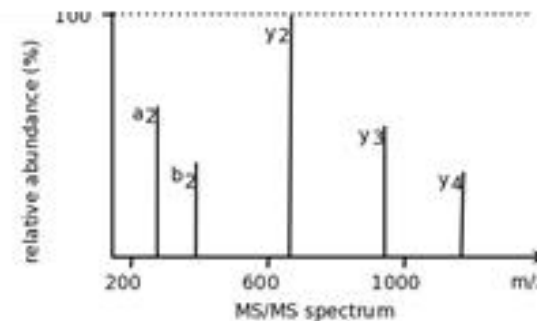
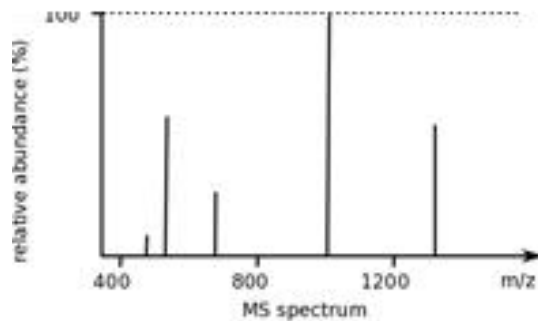


# Oznaczenia białek oparte na MS: koncepcja



Q1

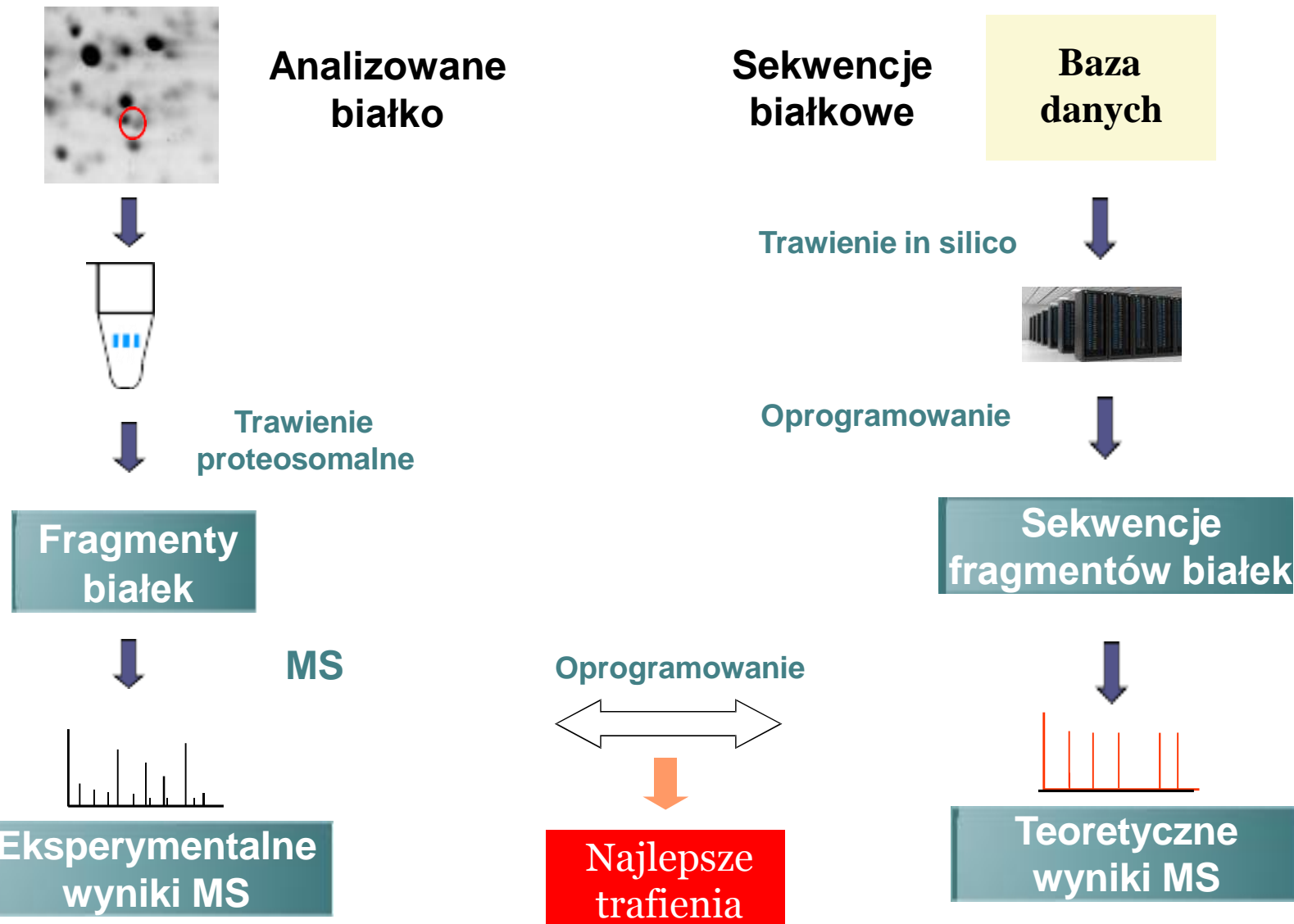
Q3



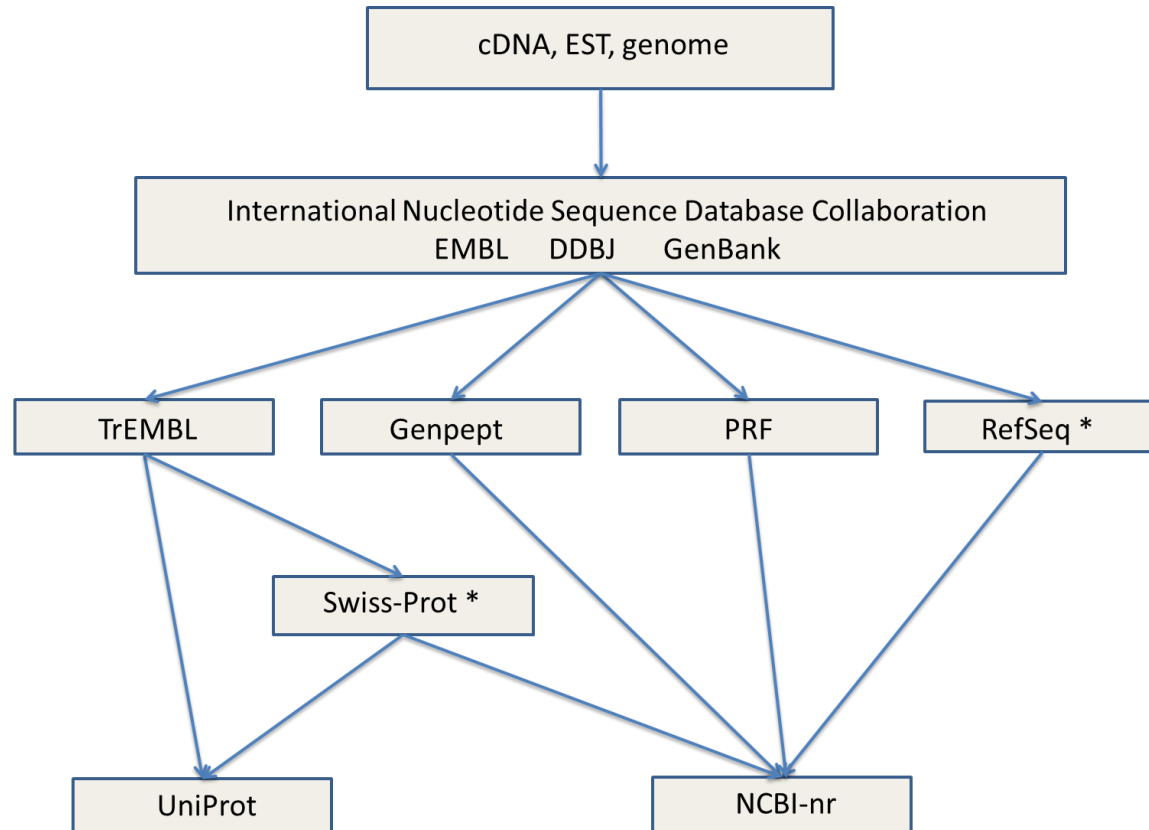
(PMF - Peptide Mass Fingerprinting)

(PFF - Peptide Fragment Fingerprinting)

# Oznaczenia białek oparte na MS: koncepcja



# Bazy danych sekwencji białkowych



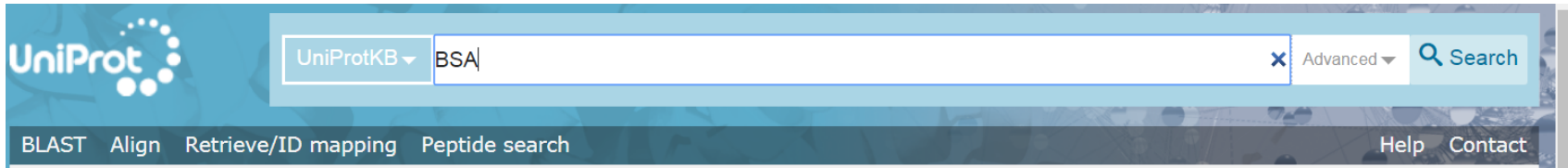
**UniProt:** Swiss-Prot + TrEMBL + (PIR)

**NCBI-nr:** Swiss-Prot + GenPept + (PIR) + RefSeq + PDB + PRF

\* proteomiczne bazy danych zawierające niezrecenzowane sekwencje białkowe



# Bazy danych sekwencji białkowych - UniProt



The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

**UniProtKB**  
UniProt Knowledgebase

Swiss-Prot (557,992)  
Manually annotated and reviewed.

TrEMBL (120,243,849)  
Automatically annotated and not reviewed.


**UniRef**  
Sequence clusters








**UniParc**  
Sequence archive




**Proteomes**



**Supporting data**

Literature citations 	Taxonomy 	Subcellular locations 
Cross-ref. databases 	Diseases XXX	Keywords 

**News** 

[Forthcoming changes](#)  
Planned changes for UniProt

---

[UniProt release 2018\\_07](#)  
Ubiquitin ligation: new insight into mechanistic diversity | Cross-references to UniLectin | Change of UniProt license

---

[UniProt release 2018\\_06](#)  
Neuronal express mRNA delivery service | International protein nomenclature

[News archive](#)

# Teoretyczne trawienie białek (in silico)

Display

None

Sequence status<sup>1</sup>: Complete.

Sequence processing<sup>1</sup>: The displayed sequence is further processed into a mature form.

P02769-1 [UniParc] [FASTA](#) [Add to basket](#)

< Hide

Length: 607

Mass (Da): 69,293

Last modified: February 1, 1996 - v4

Checksum: 39167DFE768585D4

PeptideMass  [GO](#)

```

10      20      30      40      50
MKWTFISLL LFSSAYSRG VFRDTHKSE IAHRFKDLGE EHFKGLVLIA
60      70      80      90     100
FSQYLQQCPF DEHKLVNLEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
110     120     130     140     150
VASLRETYGD MADCCKEQEP ERNECFLSHK DDSPDLPLKLP PDPNLTLCDEF
160     170     180     190     200
KADEKFKWGK YLVEIARRHP YFYAPPELLYY ANKYNQVQFE CCQAEDKGAC
210     220     230     240     250
LLPKIETHRE KVLASSARQR LRCASIQKFG ERALKANSVA RLSQKFPKAE
260     270     280     290     300
FVEVTKLVDL LTKVHKECH  GOLLECCADD ADLAKYICDN QDTISSKLEK
310     320     330     340     350
CCDKPLLEKS HCIAEVEKDA IPENLPLPTA DFAEDKDVCK NYQEAQDAFL
360     370     380     390     400
GSFLVEYSRR HPEYAVSVLL RLAKEYEATL EECCAQDDPH ACYSTVFDKL
410     420     430     440     450
KHLVDEPNQL TKQNCQDFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
460     470     480     490     500
RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
510     520     530     540     550
TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
560     570     580     590     600
ALVELLKHKP KATEEQKLTV MENFVAFVCK CCAADDKEAC FAVEGPKLVV
    
```

STQTALA



PeptideMass

## PeptideMass

The entered protein is: P02769

The selected enzyme is: Trypsin

Maximum number of missed cleavages (MC): 2

All cysteines in reduced form.

Methionines have not been oxidized.

Displaying peptides with a mass bigger than 500 Dalton.

Using monoisotopic masses of the occurring amino acid residues and giving peptide masses as [M+H]<sup>+</sup>.

You have selected ALBU\_BOVIN (P02769) from UniProtKB/Swiss-Prot:

Serum albumin precursor (BSA) (Allergen Bos d 6)  
Signal and propep in positions 1-24 have been removed.

- Chain Serum albumin at positions 25 - 607 [Theoretical pI: 5.60 / Mw (average mass): 66432.96 / Mw (monoisotopic mass): 66389.86]

mass	position	#MC	modifications	peptide sequence
4910.3837	45-88	2	PHOS: 82	4990.3500 GLVLIAFSQYLQQCPFDEHV KLVNELTEFAKTCVADESHA GCEK
4535.2954	37-75	2		DLGEEHFKGLVLIAFSQYLQ QCPFDEHVKLVNELTEFAK
4246.0511	508-544	2	PHOS: 512	4326.0174 RPCFSALTPDETYVPKAFDE KLFTFHADICTLPDTEK
4185.9547	169-204	2		HPYFYAPPELLYYANKYNGV FQECCQAEDKGACLLPK
4110.9707	300-336	2		ECCDKPLLEKSHCIAEVEKD AIPENLPLTADFAEDK
3837.8131	66-100	2	PHOS: 82, 89	3997.7458 LVNELTEFAKTCVADESHAG CEKSLHTLFGDELCK
3758.8958	402-433	2		HLVDEPNQLIKQNCQDFEKL GEYGFQNALIVR
3665.8460	35-65	2		FKDLGEEHFKGLVLIAFSQY LQQCPFDEHVK
3659.6722	168-197	2		RHPYFYAPPELLYYANKYNGV FQECCQAEDK
3579.8555	45-75	1		GLVLIAFSQYLQQCPFDEHV KLVNELTEFAK
3523.6874	460-489	2		CCTKPESERMPCTEDYLSL I NLRNLCVLEHEK
3503.5711	169-197	1		HPYFYAPPELLYYANKYNGV FQECCQAEDK
3444.4810	267-297	2	PHOS: 296	3524.4473 ECCHGDILLECADDRADLAKY ICNDQDTISSK
3419.6068	499-528	2	PHOS: 512	3499.5731 CCTESLVNRRPCFSALTPDE TYVPKAFDEK
3397.6289	310-340	2		SHCIAEVEKDAIPENLPLPT ADFAEKDVCK
3390.6826	37-65	1		DLGEEHFKGLVLIAFSQYLQ QCPFDEHVK

Sekwencje białkowe

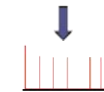
Baza danych

Trawienie in silico



Oprogramowanie

Sekwencje fragmentów białek



Teoretyczne wyniki MS

# Teoretyczne trawienie białek (in silico)

**N**~~R~~PCHSHT**K**ECESAW**K**~~N~~R**P**CHSHT**K**K**P**CHSHT**K**~~K~~~~N~~R**K**VW**K**IPPF**W**

trawienie trypsyną

~~N~~      ECESAW**K**      **K**PCHSHT**K** ~~N~~      IPPFF**W**  
**R**PCHSHT**K**      **N**R**P**CHSHT**K**      ~~K~~ ~~N~~ ~~K~~

Trypsyna w przeważającej mierze degradowuje białka po stronie karboksylowej (lub "C-końcowej") aminokwasów:

- lizyny (K)
- argininy (R)
- z wyjątkiem sytuacji, gdy jedna z powyższych jest związana z proliną (P).

# PMF (peptide mass fingerprinting)

Lista pików (m/z):

COM=10 pmol digest of Sample BSA

MASS=Monoisotopic

USERNAME=Pedro Domingues

USEREMAIL=p.domingues@ua.pt

TITLE= Cmpd 7, +MS, 16.8 min

847.50413

868.97220

922.46673

923.48150

927.49393

1022.45510

1050.45330

1163.63123

1164.65310

1193.60273

1249.62173

1250.71030

1296.75560

1297.74990

1305.71663

1416.79290

1439.81233

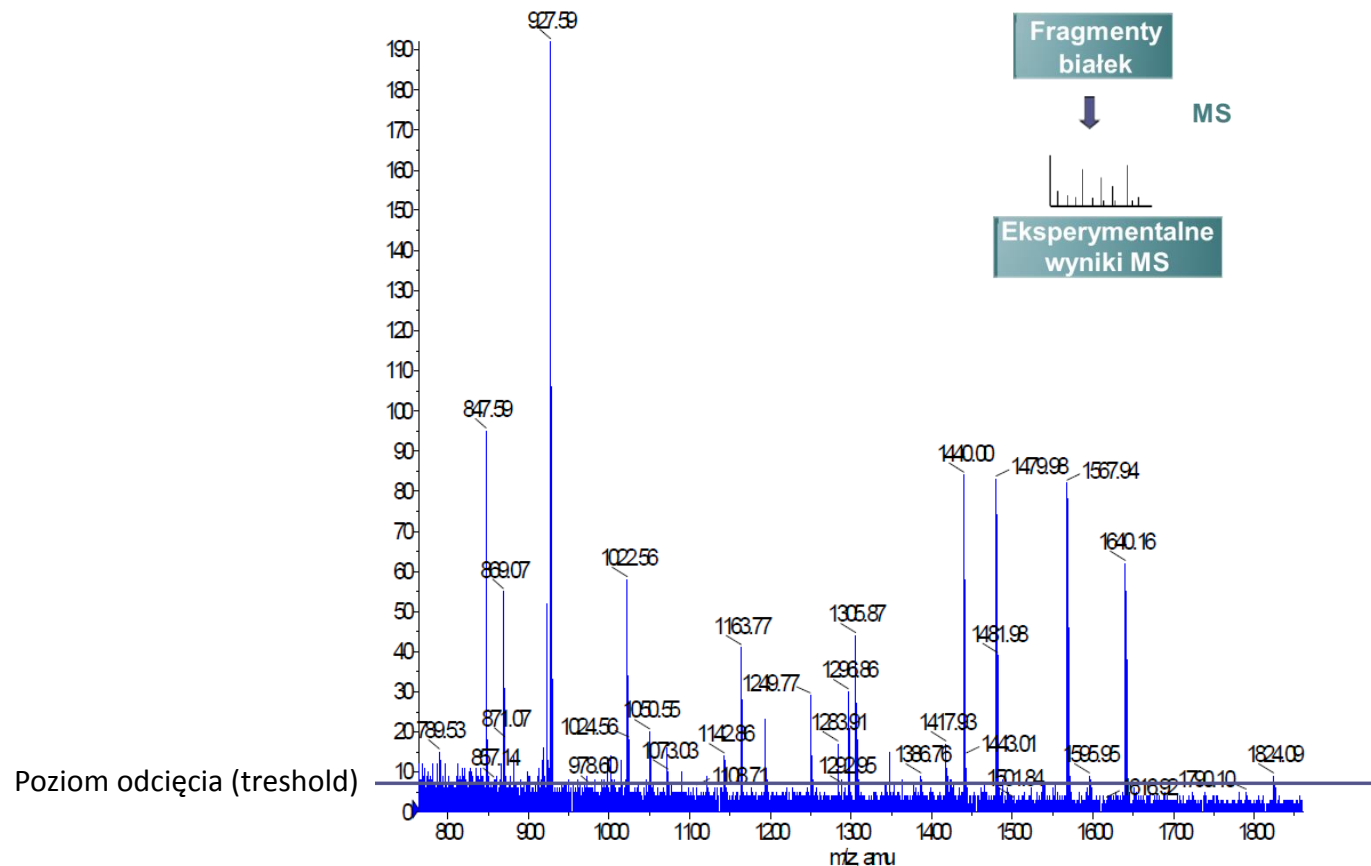
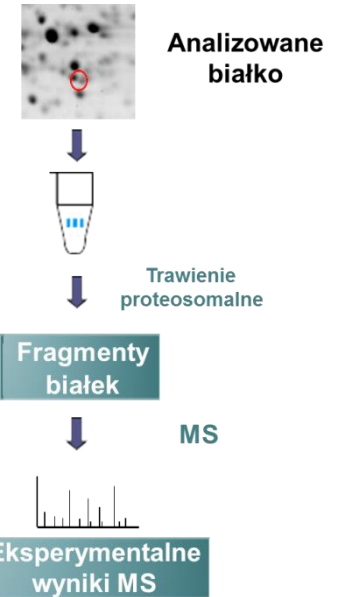
1479.79593

1482.75830

1567.74323

1639.93833

1823.86000



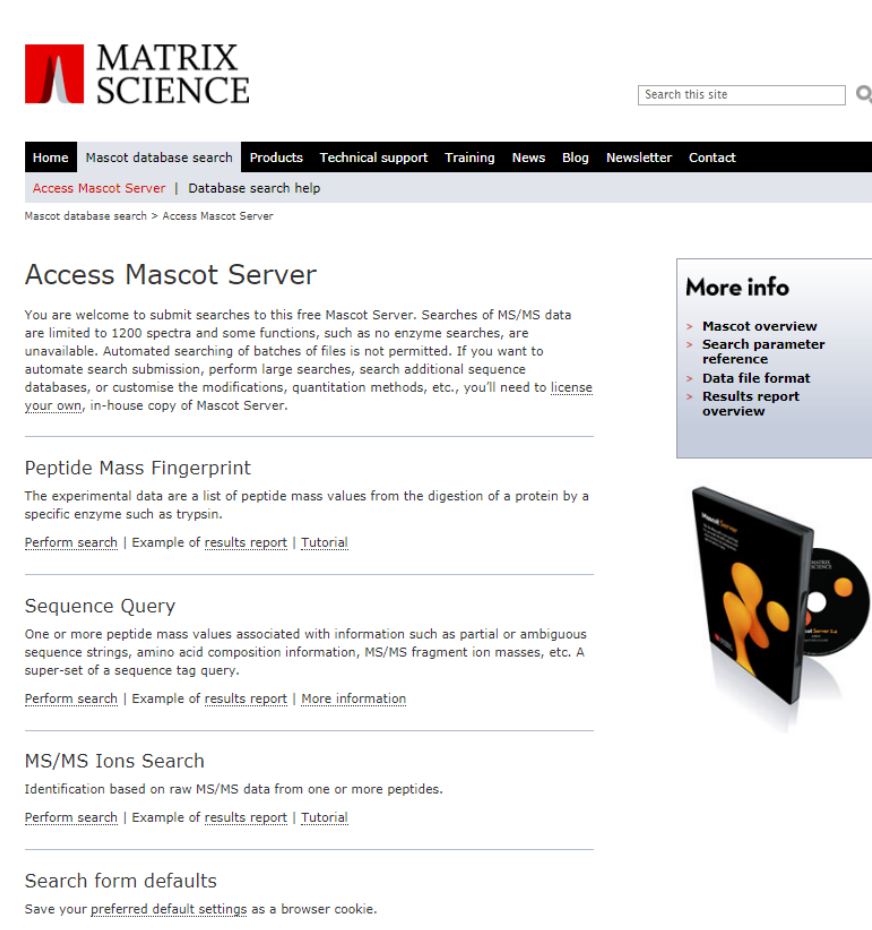
# Wyszukiwarki identyfikujące białka

---

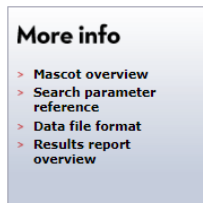
- Mascot
- X!Tandem,
- MS-GF+,
- MS Amanda,
- MyriMatch,
- Comet,
- Tide,
- Andromeda
- OMSSA
- ProteinProspector
- Sequest
- PEAKS
- PRIDE
- ....

# PMF (peptide mass fingerprinting)

## Wyszukiwanie sekwencji białek



The screenshot shows the Matrix Science Mascot Server website. At the top left is the Matrix Science logo. To its right is a search bar with the text "Search this site" and a magnifying glass icon. Below the logo and search bar is a navigation menu with links: Home, Mascot database search, Products, Technical support, Training, News, Blog, Newsletter, and Contact. Underneath the menu is a sub-menu with "Access Mascot Server" and "Database search help". The main content area is titled "Access Mascot Server" and contains a welcome message: "You are welcome to submit searches to this free Mascot Server. Searches of MS/MS data are limited to 1200 spectra and some functions, such as no enzyme searches, are unavailable. Automated searching of batches of files is not permitted. If you want to automate search submission, perform large searches, search additional sequence databases, or customise the modifications, quantitation methods, etc., you'll need to [license your own](#), in-house copy of Mascot Server." Below this is a section for "Peptide Mass Fingerprint" with the text: "The experimental data are a list of peptide mass values from the digestion of a protein by a specific enzyme such as trypsin." and links for "Perform search", "Example of results report", and "Tutorial". The next section is "Sequence Query" with the text: "One or more peptide mass values associated with information such as partial or ambiguous sequence strings, amino acid composition information, MS/MS fragment ion masses, etc. A super-set of a sequence tag query." and links for "Perform search", "Example of results report", and "More information". The following section is "MS/MS Ions Search" with the text: "Identification based on raw MS/MS data from one or more peptides." and links for "Perform search", "Example of results report", and "Tutorial". The final section is "Search form defaults" with the text: "Save your preferred default settings as a browser cookie."



**More info**

- > Mascot overview
- > Search parameter reference
- > Data file format
- > Results report overview



# PMF (peptide mass fingerprinting)

Lista pików (m/z):

847.50413  
868.97220  
922.46673  
923.48150  
927.49393  
1022.45510  
1050.45330  
1163.63123  
1164.65310  
1193.60273  
1249.62173  
1250.71030  
1296.75560  
1297.74990  
1305.71663  
1416.79290  
1439.81233  
1479.79593  
1482.75830  
1567.74323  
1639.93833  
1823.86000

## MASCOT Peptide Mass Fingerprint

Your name  Email

Search title

Database(s)   
NCBInr  
contaminants  
cRAP

Enzyme

Allow up to  missed cleavages

Taxonomy

Fixed modifications

Variable modifications

Display all modifications

Acetyl (K)  
Acetyl (N-term)  
Acetyl (Protein N-term)  
Amidated (C-term)  
Amidated (Protein C-term)  
Ammonia-loss (N-term C)  
Biotin (K)  
Biotin (N-term)  
Carbamidomethyl (C)  
Carbamyl (K)  
Carbamyl (N-term)

Protein mass  kDa

Peptide tol.  ppm

Mass values  MH<sup>+</sup>  M<sub>r</sub>  M-H<sup>-</sup>

Monoisotopic  Average

Data file  No file chosen

Query

Decoy

Report top  hits

# PMF (Wyniki w programie Mascot)

## Decoy („wabik”)

1. Podczas wyszukiwania, za każdym razem, gdy testowana jest sekwencja białek z docelowej bazy danych, sekwencja decoy o tej samej długości jest automatycznie generowana i testowana.
2. Dopasowania i wyniki dla sekwencji wabików są rejestrowane oddzielnie w pliku wynikowym.
3. Po zakończeniu wyszukiwania w nagłówku wyników zostaną podane liczby dopasowań i współczynnik fałszywych odkryć.

## Zanieczyszczenia:

Sekwencje dla typowych zanieczyszczeń, takich jak keratyny, BSA i trypsina.

Liczba „**Expect**” pokazuje ile razy spodziewamy się uzyskać równy lub wyższy wynik, wyłącznie przypadkowo (im mniej, tym lepiej).

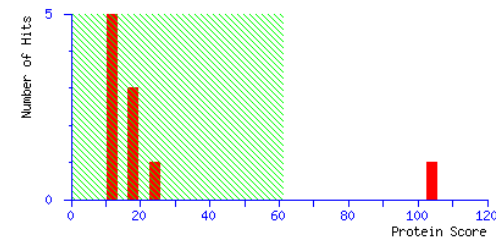
## MATRIX SCIENCE Mascot Search Results

User : pedro  
Email : p.domingues@ua.pt  
Search title :  
Database : SwissProt 2015\_10 (549646 sequences; 195983064 residues)  
Taxonomy : Mammalia (mammals) (66401 sequences)  
Timestamp : 20 Oct 2015 at 11:43:21 GMT  
Top Score : 104 for **ALBU\_BOVIN**, Serum albumin OS=Bos taurus GN=ALB PE=1 SV=4

	SwissProt	Decoy
Protein hits above identity threshold	1	0
Highest scoring protein hit	104	20

## Mascot Score Histogram

Protein score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event. Protein scores greater than 61 are significant ( $p < 0.05$ ).



## Concise Protein Summary Report

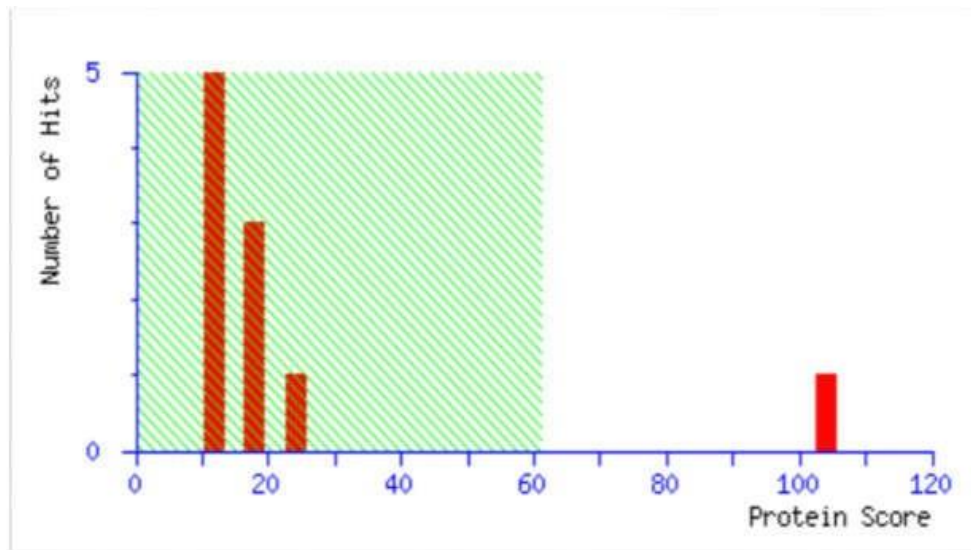
Format As  [Help](#)  
Significance threshold  $p < 0.05$  Max. number of hits   
Preferred taxonomy

1. [ALBU\\_BOVIN](#) Mass: 69248 Score: **104** Expect: 2.6e-06 Matches: 11  
Serum albumin OS=Bos taurus GN=ALB PE=1 SV=4  
[ALBU\\_FELCA](#) Mass: 68615 Score: 24 Expect: 2.6e+02 Matches: 4  
Serum albumin OS=Felis catus GN=ALB PE=1 SV=1  
[ALBU\\_CAPHI](#) Mass: 10048 Score: 24 Expect: 2.8e+02 Matches: 2  
Serum albumin (Fragments) OS=Capra hircus GN=ALB PE=1 SV=2  
[ALBU\\_SHEEP](#) Mass: 69143 Score: 22 Expect: 4e+02 Matches: 4



# PMF (Wyniki w programie Mascot)

- Wynik w Mascot Score jest podawany jako  $S = -10 \cdot \log(P)$ , gdzie:
  - P to prawdopodobieństwo, że obserwowany mecz jest przypadkowym zdarzeniem
  - $P = E \cdot N^{-1}$ 
    - E = wartość oczekiwana
    - N = liczba białek w bazie danych
- Znaczenie tego wyniku zależy od wielkości przeszukiwanej bazy danych. Mascot zaznacza na zielono trafienia przy wartości odcięcia E = 0,05



# PMF (Wyniki w programie Mascot)

## MATRIX SCIENCE MASCOT Search Results

### Protein View: ALBU\_BOVIN

Serum albumin OS=Bos taurus GN=ALB PE=1 SV=4

Database: SwissProt  
Score: 104  
Expect: 2.6e-06  
Nominal mass (M<sub>r</sub>): 69248  
Calculated pI: 5.82  
Taxonomy: [Bos taurus](#)

Sequence similarity is available as [an NCBI BLAST search of ALBU\\_BOVIN against nr](#).

### Search parameters

Enzyme: Trypsin: cuts C-term side of KR unless next residue is P,  
Variable modifications: [Oxidation \(M\)](#)  
Mass values searched: 22  
Mass values matched: 11

### Protein sequence coverage: 18%

Matched peptides shown in **bold red**.

```
1 MKWVTFISLL LFFSSAYSRG VFRRDDTHKSE IAHRFKDLGE EHPKGLVLIA
51 FSQYLQQCPF DEHVKLVEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
101 VASLRRETYGD MADOCCEKQEP ERNECFLSHK DDSFDLPKPKL PDPNTLCDEF
151 KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNQVVFQE CQQAEDKQAC
201 LLPKIETMRE KVLISSARQR LRCASIQKFG ERALKAWNSVA RLSQKFPKAE
251 FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDKPLLEKS HCIAEVEKDA IPENLPLTA DPAEDKDVCK NYQEAKDAFL
351 GSPLYEYSRR HPEYAVSVLL RLAKYEATL EECCKADDPH ACYSTVFDKL
401 KHLVDEPQNL IKQNCDFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
451 RSLGKVGTRC CTKPESERMP CTEDVLSLIL NRLCVLHEKT EVSEKVTKCC
501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADKKEAC FAVEGPKLVV
601 STQTALA
```

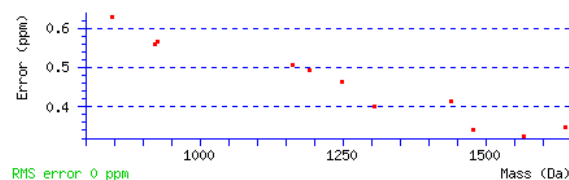
Unformatted sequence string: [607 residues](#) (for pasting into other applications).

Sort peptides by  Residue Number  Increasing Mass  Decreasing Mass

Show predicted peptides also

Start - End	Observed	Mr (expt)	Mr (calc)	ppm	M	Peptide
25 - 34	1193.6027	1192.5955	1192.5949	0.49	1	R.DTHKSEIAHR.F
35 - 44	1249.6217	1248.6145	1248.6139	0.46	1	R.FRDLGEEHFK.G
66 - 75	1163.6312	1162.6240	1162.6234	0.50	0	K.LVNELTEFAK.T
161 - 167	927.4939	926.4867	926.4861	0.57	0	K.YLYEIA.R
205 - 211	922.4667	921.4595	921.4589	0.56	1	K.IETMREK.V + Oxidation (M)
242 - 248	847.5041	846.4969	846.4963	0.63	1	R.LSQKFPK.A
347 - 359	1567.7432	1566.7360	1566.7354	0.33	0	K.DAFLGSPLYEYSR.R
360 - 371	1439.8123	1438.8051	1438.8045	0.41	1	R.RHPEYAVSVLLR.L
402 - 412	1305.7166	1304.7094	1304.7088	0.40	0	K.HLVDEPQNLIK.Q
421 - 433	1479.7959	1478.7887	1478.7881	0.34	0	K.LGEYGFQNALIVR.Y
437 - 451	1639.9383	1638.9311	1638.9305	0.35	1	R.KVPQVSTPTLVEVSR.S

No match to: 868.9722, 923.4815, 1022.4551, 1050.4533, 1164.6531, 1250.7103, 1296.7556, 1297



# PMF (peptide mass fingerprinting)

Lista pików (m/z):

COM=10 pmol digest of Sample BSA

MASS=Monoisotopic

USERNAME=Pedro Domingues

USEREMAIL=p.domingues@ua.pt

TITLE= Cmpd 7, +MS, 16.8 min

847.50413

868.97220

922.46673

923.48150

927.49393

1022.45510

1050.45330

1163.63123

1164.65310

1193.60273

1249.62173

1250.71030

1296.75560

1297.74990

1305.71663

1416.79290

1439.81233

1479.79593

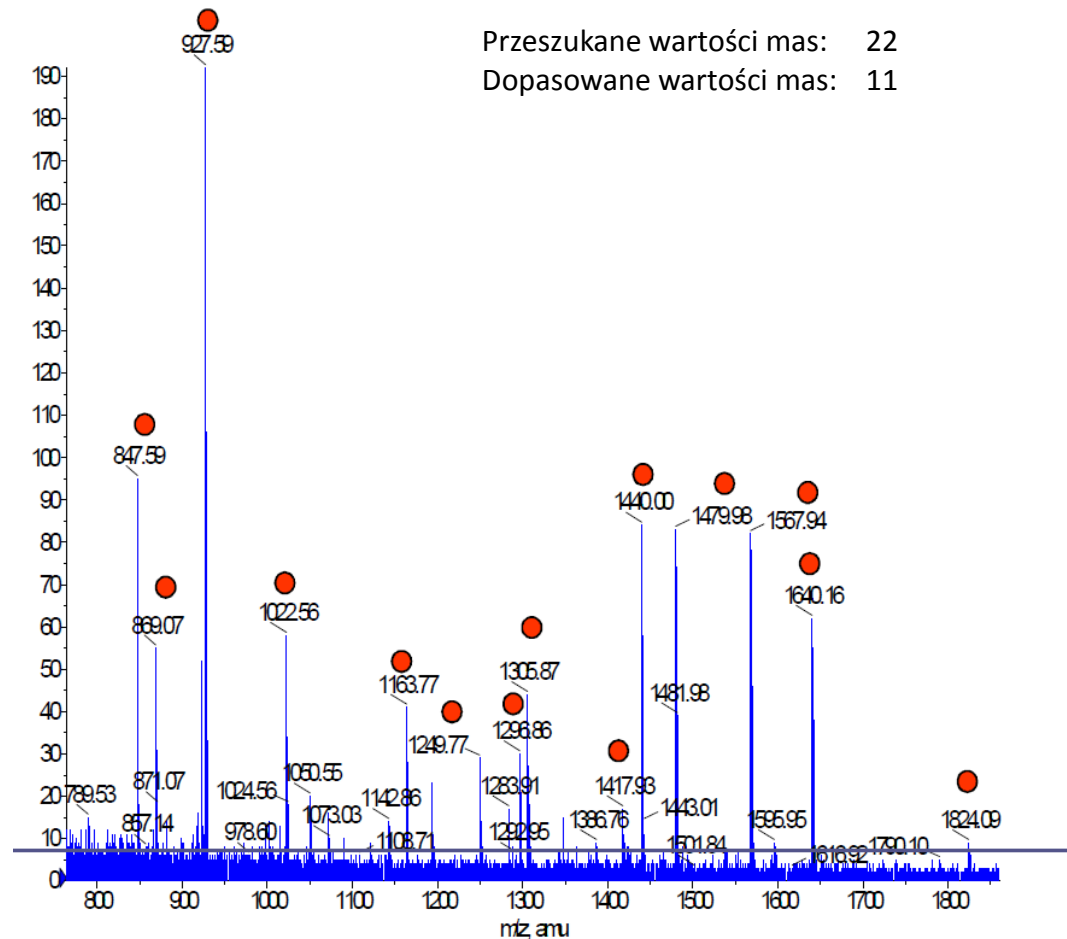
1482.75830

1567.74323

1639.93833

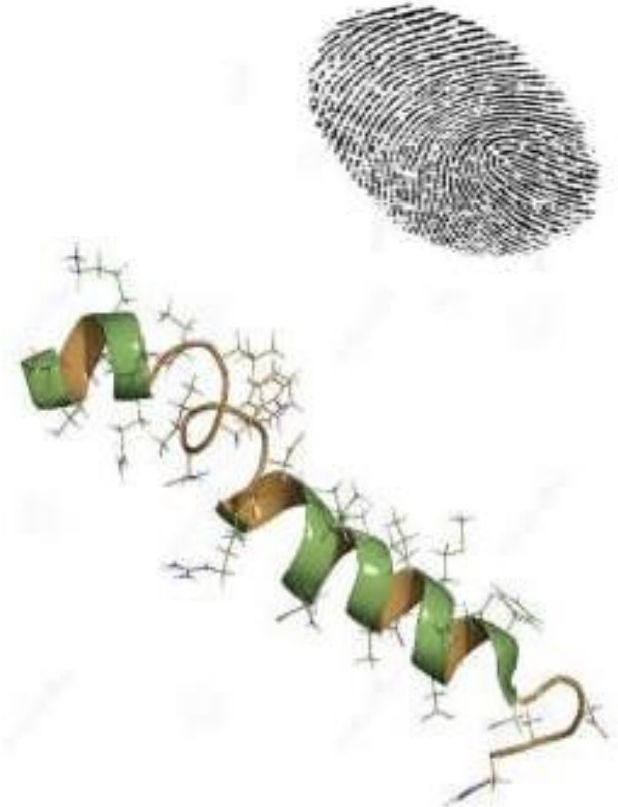
1823.86000

Poziom odcięcia (treshold)

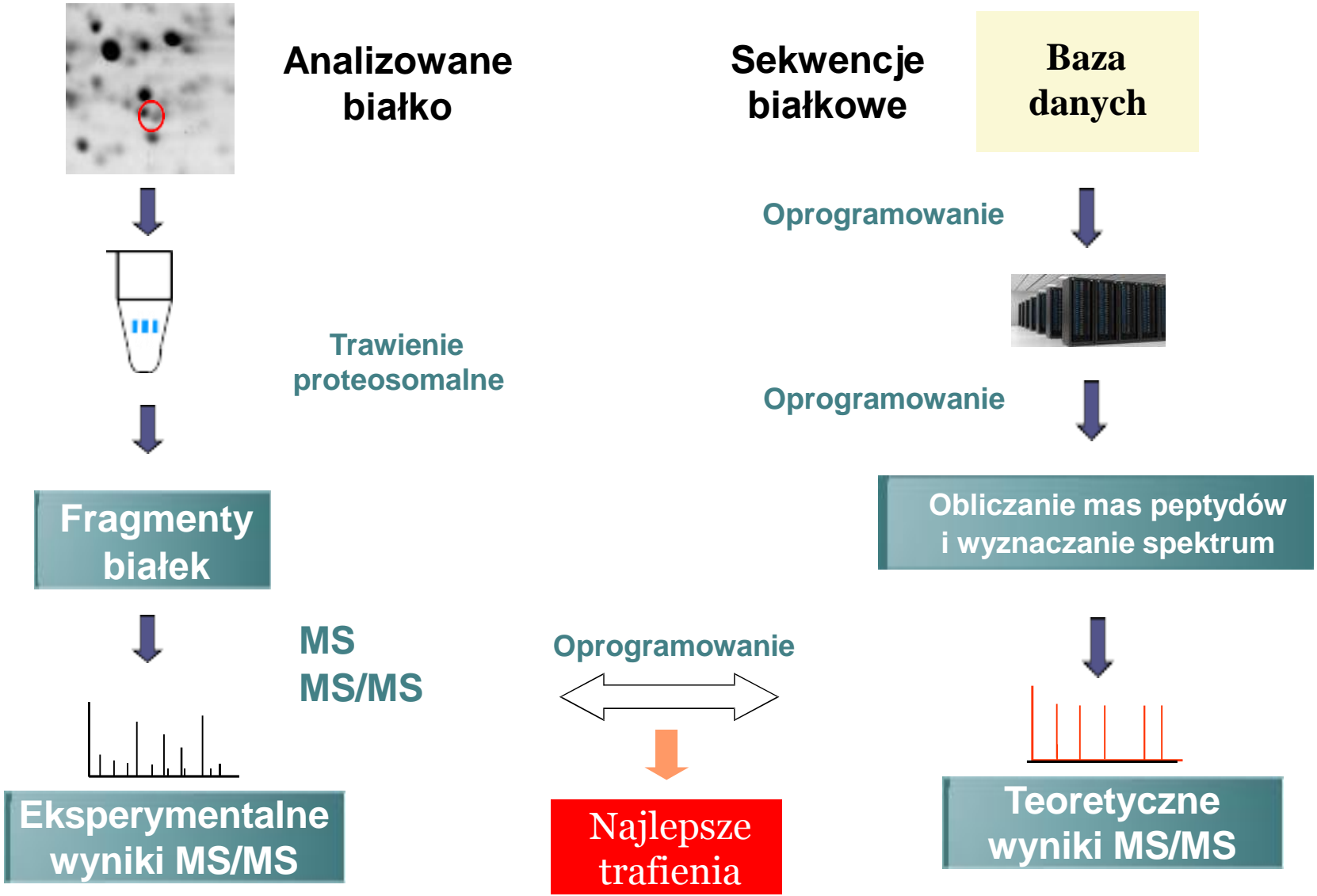


# PMF (peptide mass fingerprinting)

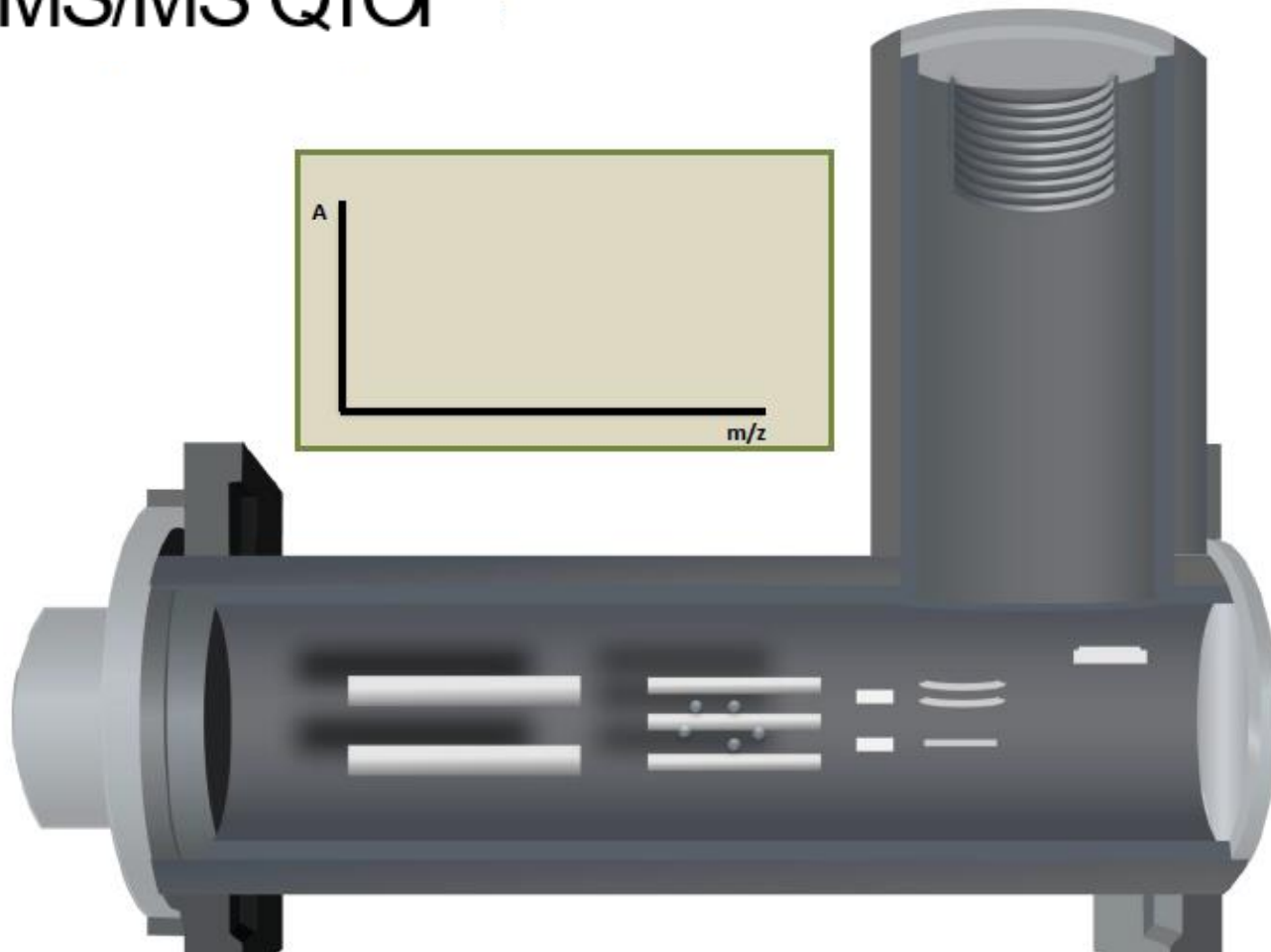
- Szybka, prosta analiza
- Wysoka czułość
- Nie nadaje się do analiz mieszanin (szczególnie w przypadku małych cząsteczek)
- Niezbędna jest wysoka dokładność analizatora mas.



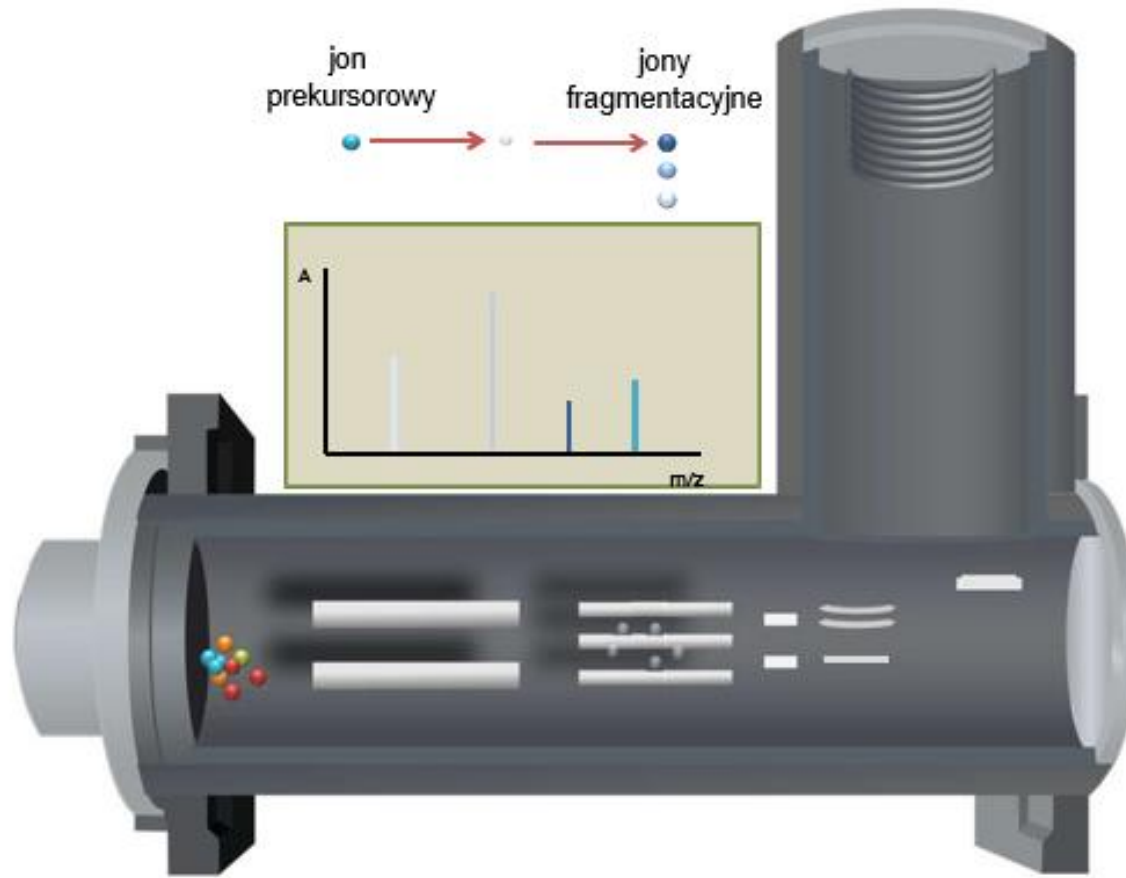
# Oznaczenia białek oparte na MS: koncepcja PFF



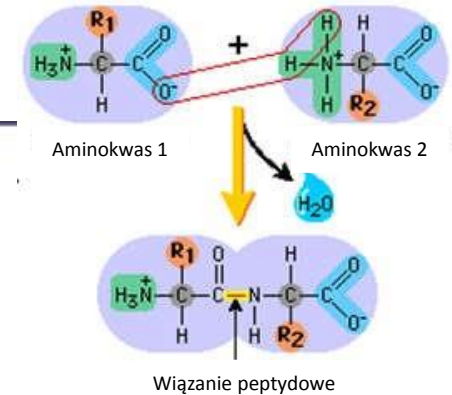
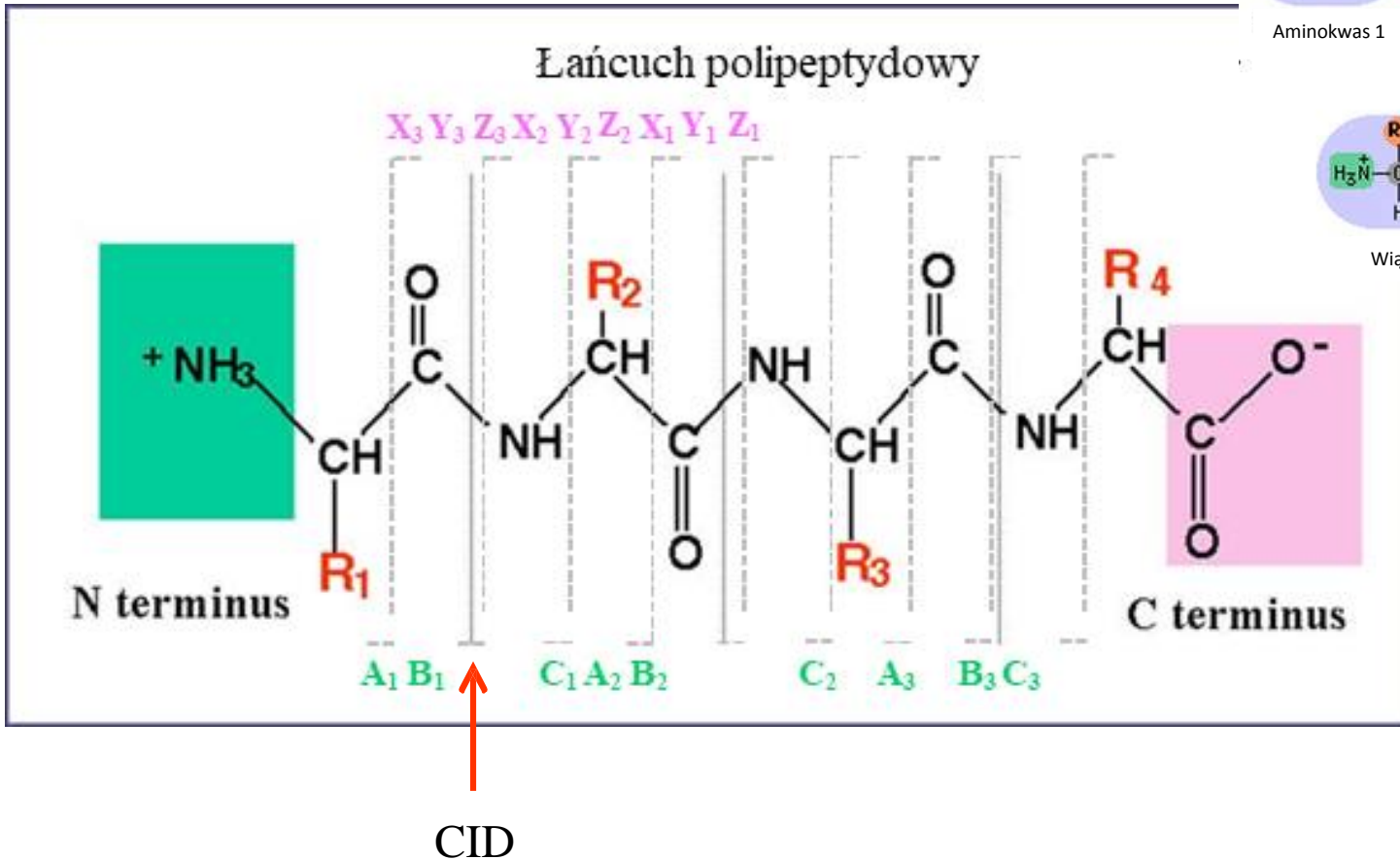
# MS/MS QTOF



# MS/MS QTOF

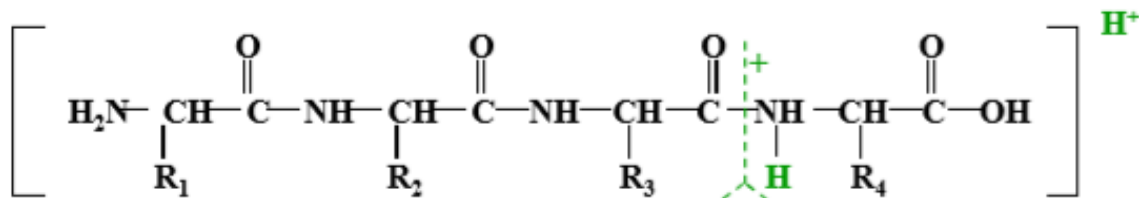
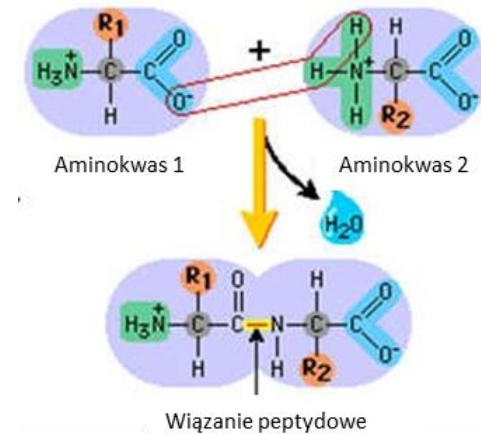


# Fragmentacja peptydów





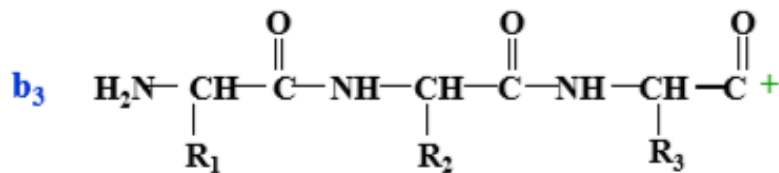
# Fragmentacja zależna od ładunku



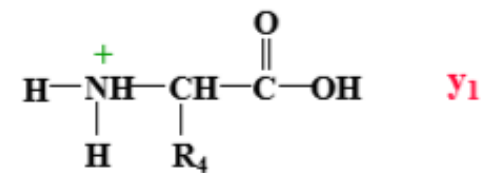
Formowanie jonów typu b

oraz/lub

Formowanie jonów typu y

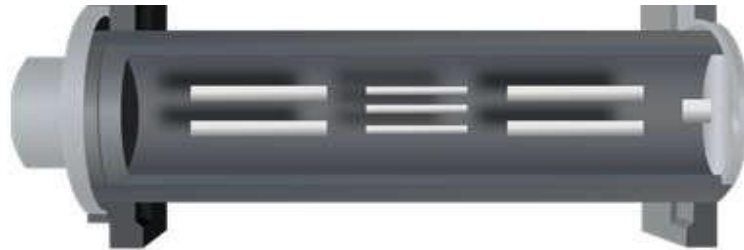
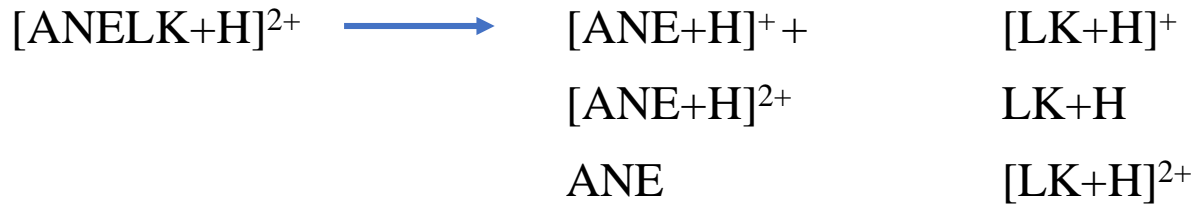


+  
cząsteczka neutralna odebrana przez system próżniowy

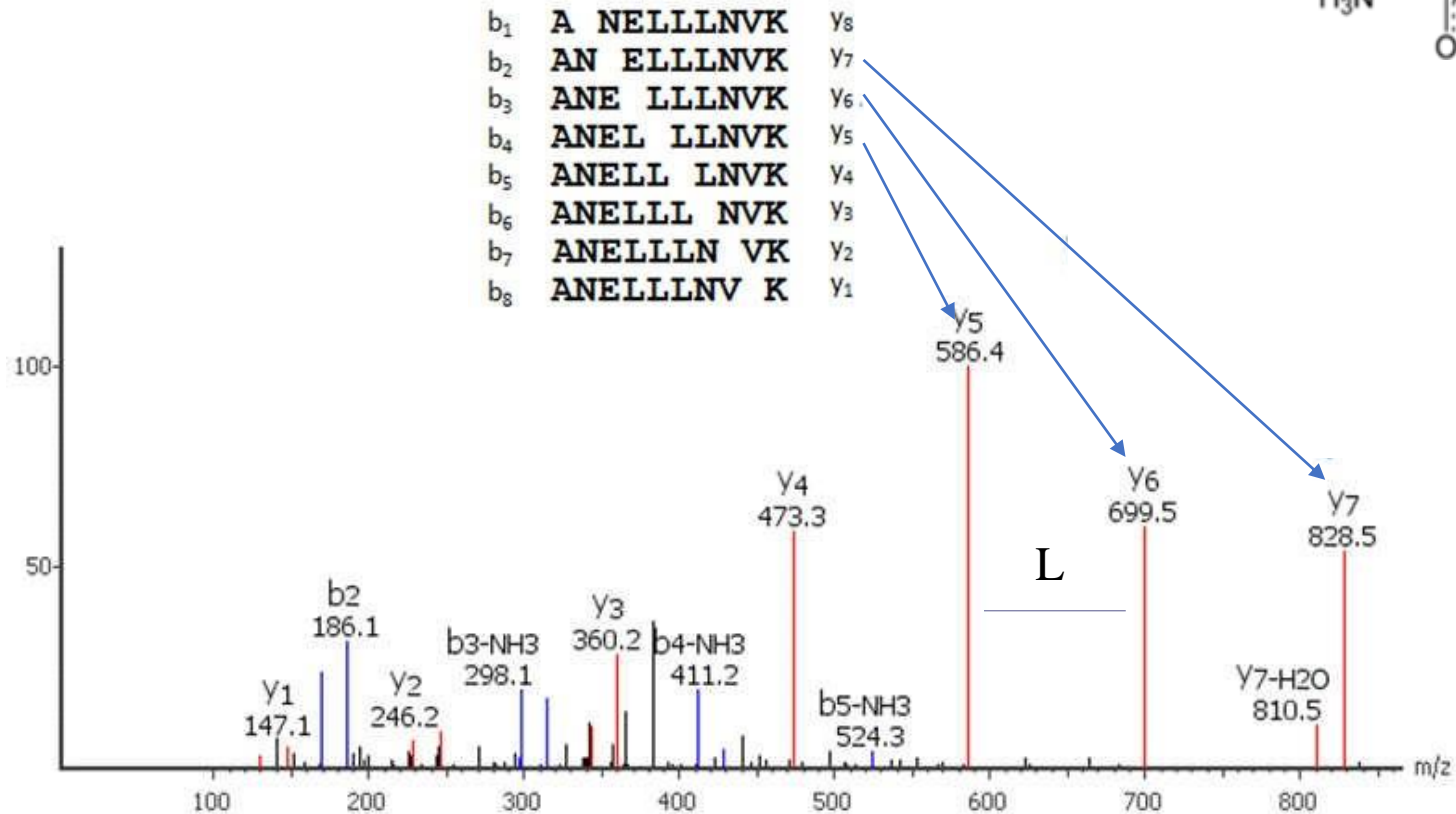
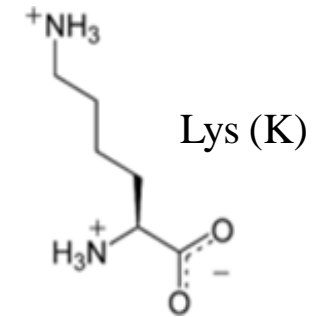
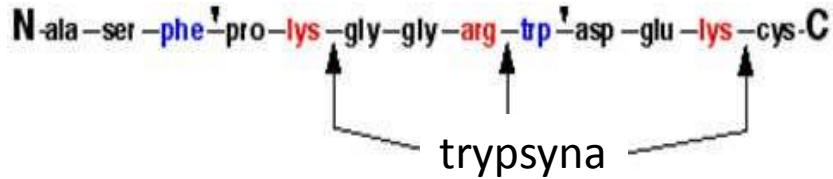


+  
cząsteczka neutralna odebrana przez system próżniowy

# Dopełniające się pary jonów b/y (wielokrotnie naładowane jony)



# Dopełniające się pary jonów b/y



# Modele fragmentacji peptydów in silico

## Fragment Ion Calculator Results

Sequence: AEFVEVTK, pI: 4.53158

### Fragment Ion Table, monoisotopic masses

Seq	#	B	Y	# (+1)
A	1	72.04444	922.48807	8
E	2	201.08703	851.45095	7
F	3	348.15544	722.40836	6
V	4	447.22386	575.33995	5
E	5	576.26645	476.27153	4
V	6	675.33486	347.22894	3
T	7	776.38254	248.16053	2
K	8	904.47750	147.11285	1

### Mass/Charge Table

	Mass	
	Mono	Avg
(M)	921.48079	922.04593
(M+H) <sup>+</sup>	922.48807	923.05320
(M+2H) <sup>2+</sup>	461.74769	462.03026
(M+3H) <sup>3+</sup>	308.16757	308.35595
(M+4H) <sup>4+</sup>	231.37751	231.51879

Sekwencje białkowe

Baza danych

Oprogramowanie



Oprogramowanie

Obliczanie mas peptydów i wyznaczenie spektrum



Teoretyczne wyniki MS/MS

<http://db.systemsbioology.net:8080/proteomicsToolkit/FragIonServlet.html>

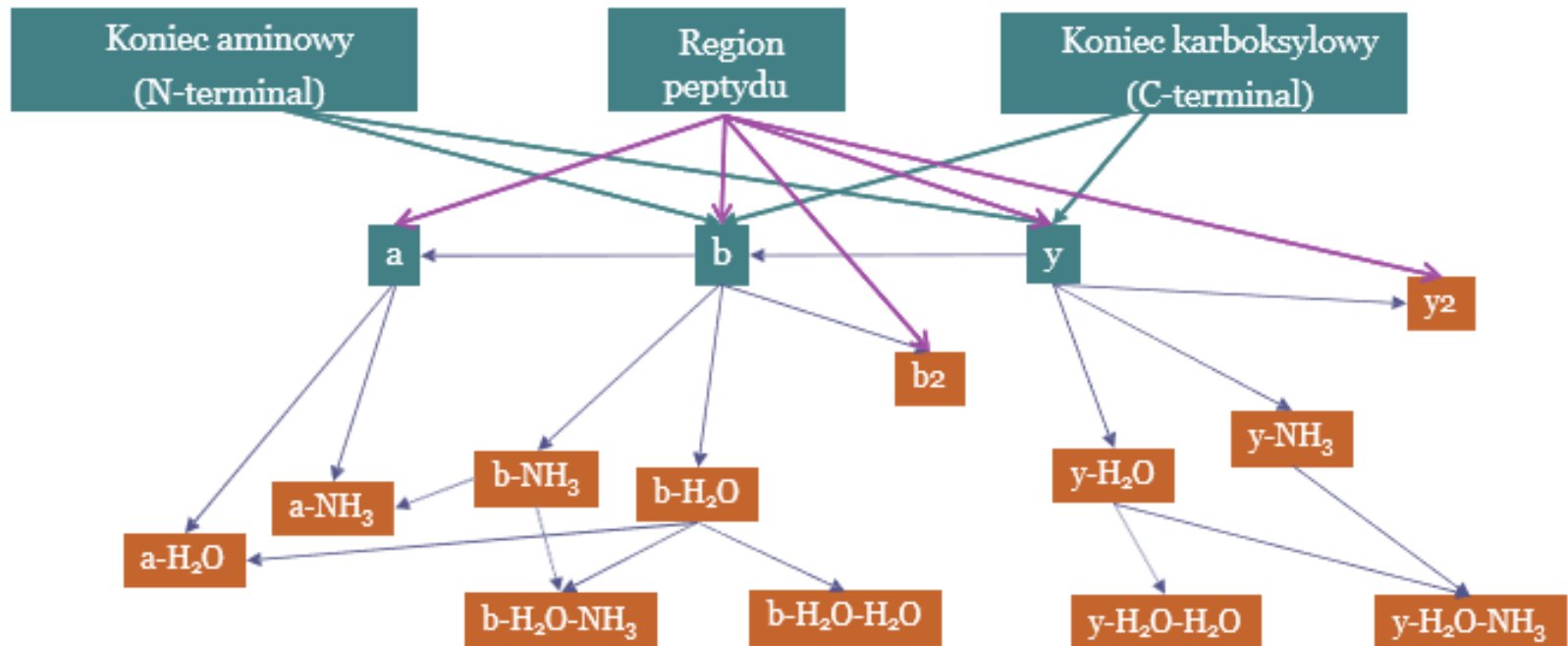
**AEFVEVTK**

# Modele fragmentacji peptydów

Sposób jonizacji

Metoda fragmentowania

Energia fragmentacji



- ❖ korelacje między jonami
- ❖ zależności wynikające z pozycji miejsca cięcia w peptydzie
- ❖ wpływ flankujących aminokwasów na miejsce cięcia

# Wyszukiwarki identyfikujące peptydy

---

- Mascot
- X!Tandem,
- MS-GF+,
- MS Amanda,
- MyriMatch,
- Comet,
- Tide,
- Andromeda
- OMSSA
- ProteinProspector
- Sequest
- PEAKS
- PRIDE
- ....

[Home](#) [Mascot database search](#) [Products](#) [Technical support](#) [Training](#) [News](#) [Blog](#) [Newsletter](#) [Contact](#)

[Access Mascot Server](#) | [Database search help](#)

[Mascot database search](#) > [Access Mascot Server](#)

## Access Mascot Server

You are welcome to submit searches to this free Mascot Server. Searches of MS/MS data are limited to 1200 spectra and some functions, such as no enzyme searches, are unavailable. Automated searching of batches of files is not permitted. If you want to automate search submission, perform large searches, search additional sequence databases, or customise the modifications, quantitation methods, etc., you'll need to [license your own](#), in-house copy of Mascot Server.

### Peptide Mass Fingerprint

The experimental data are a list of peptide mass values from the digestion of a protein by a specific enzyme such as trypsin.

[Perform search](#) | [Example of results report](#) | [Tutorial](#)

### Sequence Query

One or more peptide mass values associated with information such as partial or ambiguous sequence strings, amino acid composition information, MS/MS fragment ion masses, etc. A super-set of a sequence tag query.

[Perform search](#) | [Example of results report](#) | [More information](#)

### MS/MS Ions Search

Identification based on raw MS/MS data from one or more peptides.

[Perform search](#) | [Example of results report](#) | [Tutorial](#)

### More info

- > [Mascot overview](#)
- > [Search parameter reference](#)
- > [Data file format](#)
- > [Results report overview](#)



# Specyficzna fragmentacja peptydów (PFF, Peptide Fragment Fingerprinting)

COM=10 pmol digest of Sample BSA

MASS=Monoisotopic

USERNAME=Pedro Domingues

[USEREMAIL=p.domingues@ua.pt](mailto:p.domingues@ua.pt)

BEGIN IONS

TITLE= Cmpd 7, +MSn(461.8), 16.8 min

PEPMASS=461.7491 18565 CHARGE=2+

147.05 229

154.98 787 1+

172.95 1734 1+

183.87 410 1+

200.91 2479 1+

218.92 144

248.16 150

347.22 250

347.96 200

476.17 670 1+

548.13 180

575.33 672

576.37 393 1+

675.33 231

722.45 13125 1+

723.14 357 1+

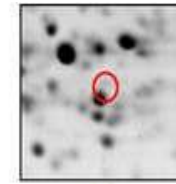
759.39 159

833.47 141

851.45 149

904.13 432 1+

END IONS



próbka



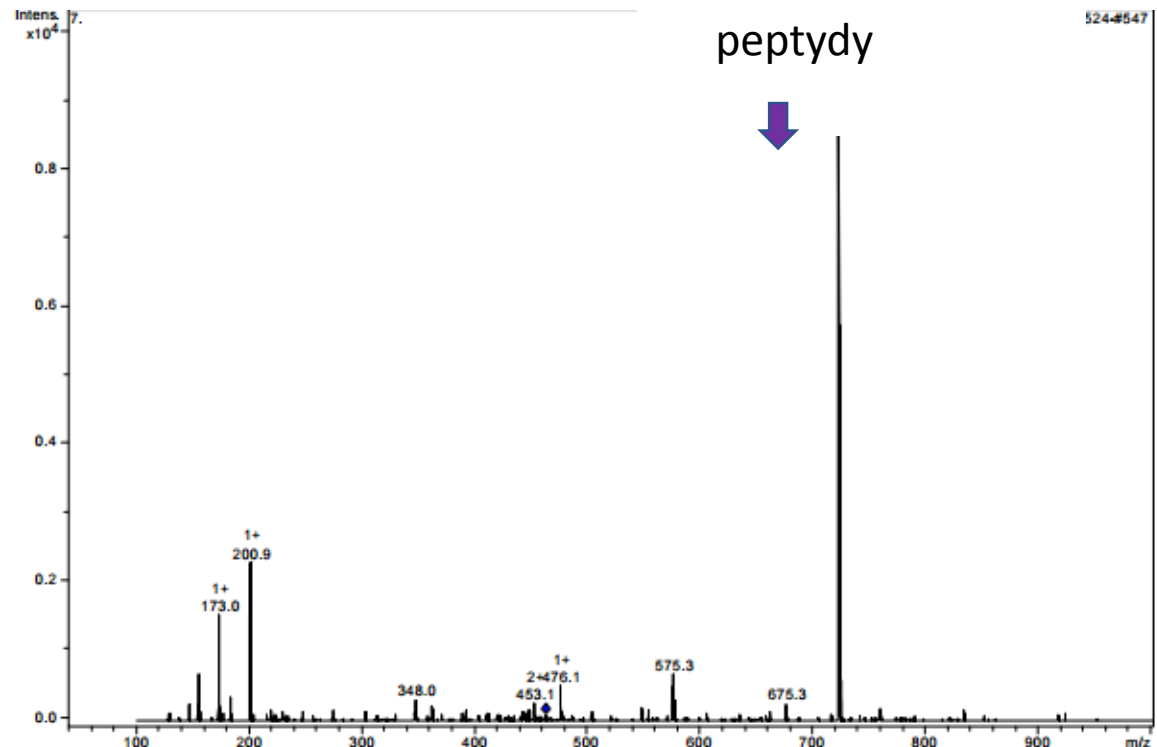
trawienie  
trypsyną



peptydy

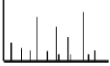


MS/MS 461.7





# PFF (Mascot)

  
Eksperymentalne  
wyniki MS/MS

Oprogramowanie



Najlepsze  
trafienia

  
Teoretyczne  
wyniki MS/MS



Search this site 

Home Mascot database search Products Technical support Training News Blog Newsletter Contact

Access Mascot Server | Database search help

Mascot database search > Access Mascot Server > MS/MS Ions Search

## MASCOT MS/MS Ions Search

Your name  Email

Search title

Database(s)

*Amino acid (AA)*  
contaminants  
cRAP  
NCBIprot  
*Nucleic acid (NA)*  
Environmental\_EST  
Fungi\_EST  
Human\_EST  
Invertebrates\_EST  
Mammals\_EST

Taxonomy

Enzyme   Allow up to  missed cleavages

Quantitation

Fixed modifications

Display all modifications

Variable modifications

Acetyl (K)  
Acetyl (N-term)  
Acetyl (Protein N-term)  
Amidated (C-term)  
Amidated (Protein C-term)  
Ammonia-loss (N-term C)  
Biotin (K)  
Biotin (N-term)  
Carbamidomethyl (C)  
Carbamyl (K)  
Carbamyl (N-term)

Peptide tol.  $\pm$   ppm  # <sup>13</sup>C

MS/MS tol.  $\pm$   ppm

Peptide charge

Monoisotopic  Average

Data file  Mascot MSMS2.txt

Data format

Precursor  m/z

Instrument

Error tolerant

Decoy

Report top   hits

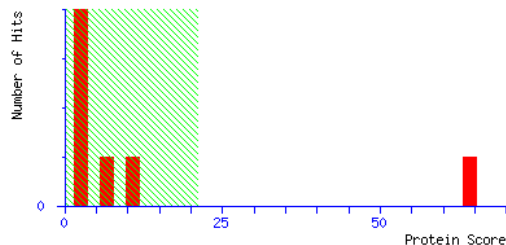
## MATRIX SCIENCE Mascot Search Results

User : Pedro Domingues  
 Email : p.domingues@ua.pt  
 Search title : 10 pmol digest of Sample BSA  
 MS data file : new 1.txt  
 Database : SwissProt 2015\_10 (549646 sequences; 195983064 residues)  
 Taxonomy : Mammalia (mammals) (66401 sequences)  
 Timestamp : 21 Oct 2015 at 13:31:53 GMT  
 Protein hits : [ALBU\\_BOVIN](#) Serum albumin OS=Bos taurus GN=ALB PE=1 SV=4

	SwissProt	Decoy	False discovery rate
Peptide matches above identity threshold	1	0	0.00 %
Peptide matches above homology or identity threshold	1	0	0.00 %

### Mascot Score Histogram

Ions score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event. Individual ions scores  $> 21$  indicate identity or extensive homology ( $p < 0.05$ ). Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



### Peptide Summary Report

Format As	Peptide Summary	<a href="#">Help</a>
Significance threshold p<	0.05	Max. number of hits
Standard scoring	<input checked="" type="radio"/> MudPIT scoring <input type="radio"/> Ions score or expect cut-off	0
Show pop-ups	<input checked="" type="radio"/> Suppress pop-ups <input type="radio"/> Sort unassigned	Decreasing Score
Preferred taxonomy	All entries	Require bold red <input type="checkbox"/>

Select All Select None Search Selected  Error tolerant

1. [ALBU\\_BOVIN](#) Mass: 69248 Score: 64 Matches: 1(1) Sequences: 1(1)  
 Serum albumin OS=Bos taurus GN=ALB PE=1 SV=4  
 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> 1	461.7491	921.4836	921.4807	3.15	0	64	3e-06	1	U	K.AEFVEVTK.L

### Search Parameters

Type of search : MS/MS Ion Search  
 Enzyme : Trypsin  
 Mass values : Monoisotopic  
 Protein Mass : Unrestricted  
 Peptide Mass Tolerance :  $\pm 5$  ppm  
 Fragment Mass Tolerance :  $\pm 0.5$  Da  
 Max Missed Cleavages : 1  
 Instrument type : Default  
 Number of queries : 1

# PFF (Mascot)

## MATRIX SCIENCE MASCOT Search Results

### Protein View: ALBU\_BOVIN

Serum albumin OS=Bos taurus GN=ALB PE=1 SV=4

Database: SwissProt  
Score: 61  
Nominal mass ( $M_r$ ): 69248  
Calculated pI: 5.82  
Taxonomy: [Bos taurus](#)

Sequence similarity is available as [an NCBI BLAST search of ALBU\\_BOVIN against nr.](#)

### Search parameters

MS data file: new 1.txt  
Enzyme: Trypsin: cuts C-term side of KR unless next residue is P.

### Protein sequence coverage: 1%

Matched peptides shown in **bold red**.

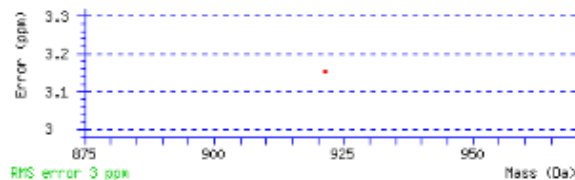
```
1 MKWVTFISLL LLFSSAYSRG VFRDRTHKSE IAHRFKDLGE EHFKGLVLIA
51 FSQYLQQCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
101 VASLRETYGD MADCCERQEP ERNECFLSHK DSDPDLPLK PDPNTLCDEF
151 KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVVFQE CCQAEKDGAC
201 LLPKIETMRE KVLASSARQR LRCASIQKFG ERALKAWSVA RLSQKFFKAE
251 FVEVTKLVTD LTKVHKECOH GDLLECADDR ADLAKYICDN QDTISSKLEK
301 CCKPLLEKLS HCIAEVEKDA IPENLPFLTA DFAEDKDVCK NYQEAKDAFL
351 GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
401 KHLVDEPQNL IKQNCQDFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
451 RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT FVSEKVKKCC
501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV
601 STQTALA
```

Unformatted sequence string: [607 residues](#) (for pasting into other applications).

Sort peptides by  Residue Number  Increasing Mass  Decreasing Mass

Show predicted peptides also

Query	Start - End	Observed	Mr(expt)	Mr(calc)	ppm	M Score	Expect	Rank	U	Peptide
<a href="#">FVEVTKLVTD</a>	249 - 256	461.7491	921.4836	921.4807	3.15	0 61	6.1e-06	1	U	K.AEFVEVTK.L



```
AC P02769; ASPJX3; O02787; P04277; Q3S2R2;
DI 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DI 01-FEB-1996, sequence version 4.
DI 14-OCT-2015, entry version 148.
DE RecName: Full=Serum albumin;
DE AltName: Full=BSA;
DE AltName: Allergen=Bos d 6;
DE FlgAs: Precursor;
```

# PFF (Mascot)

## MATRIX SCIENCE Mascot Search Results

### Peptide View

MS/MS Fragmentation of **AEFVEVTK**

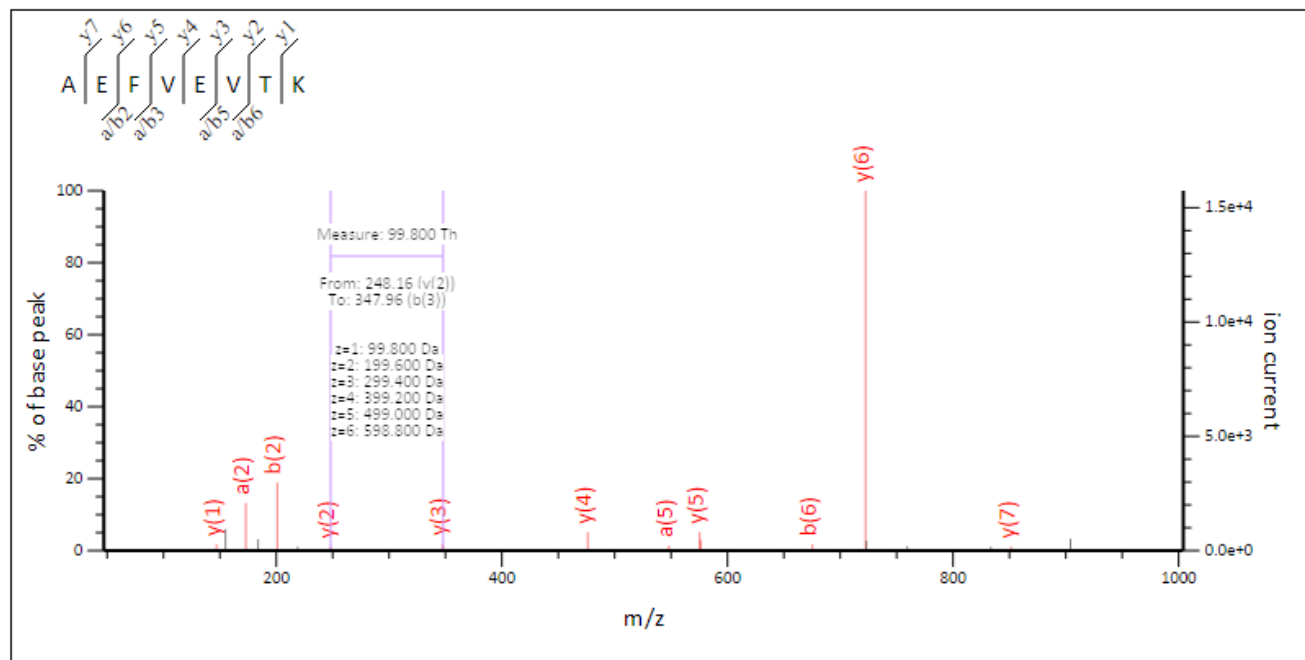
Found in **ALBU\_BOVIN** in **SwissProt**, Serum albumin OS=Bos taurus GN=ALB PE=1 SV=4

Match to Query 1: 921.483648 from(461.749100,2+) intensity(18565.0000) index(0)

Title: Cmpd 7, +MSn(461.8), 16.8 min

Data file new 1.txt

COM=10 pmol digest of Sample BSA  
MASS=Monoisotopic  
USERNAME=Pedro Domingues  
USEREMAIL=p.domingues@ua.pt  
BEGIN IONS  
TITLE= Cmpd 7, +MSn(461.8), 16.8 min  
PEPMASS=461.7491 18565  
CHARGE=2+  
147.05 229  
154.98 787 1+  
172.95 1734 1+  
183.87 410 1+  
200.91 2479 1+  
218.92 144  
248.16 150  
347.22 250  
347.96 200  
476.17 670 1+  
548.13 180  
575.33 672  
576.37 393 1+  
675.33 231  
722.45 13125 1+  
723.14 357 1+  
759.39 159  
833.47 141  
851.45 149  
904.13 432 1+  
END IONS



# 2-PFF (Mascot)

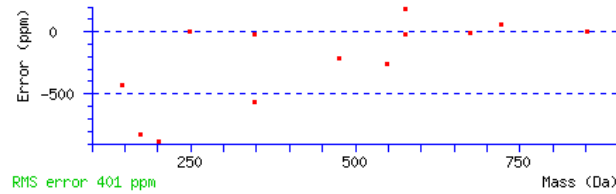
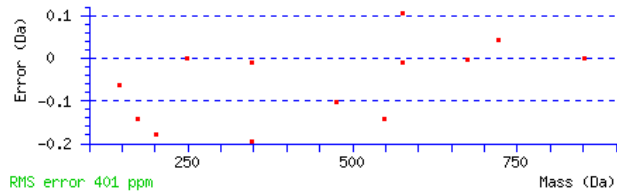
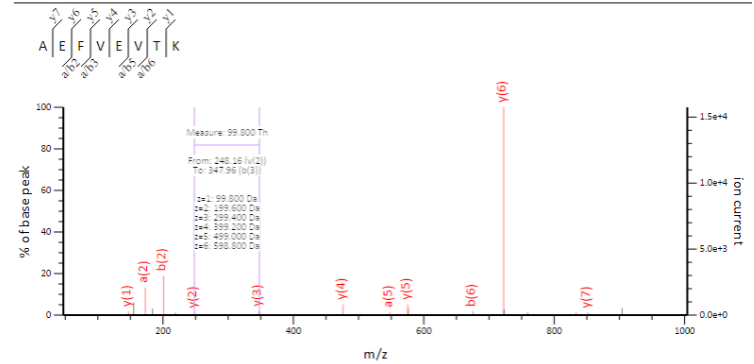
Label all possible matches  Label matches used for scoring

Monoisotopic mass of neutral peptide Mr(calc): 921.4807

Ions Score: 64 Expect: 2.5e-05

Matches : 13/56 fragment ions using 19 most intense peaks [\(help\)](#)

#	a	a <sup>++</sup>	b	b <sup>++</sup>	Seq.	y	y <sup>++</sup>	y <sup>*</sup>	y <sup>+++</sup>	#
1	44.0495	22.5284	72.0444	36.5258	A					8
2	<b>173.0921</b>	87.0497	<b>201.0870</b>	101.0471	E	<b>851.4509</b>	426.2291	834.4244	417.7158	7
3	320.1605	160.5839	<b>348.1554</b>	174.5813	F	<b>722.4083</b>	361.7078	705.3818	353.1945	6
4	419.2289	210.1181	447.2238	224.1155	V	<b>575.3399</b>	288.1736	558.3134	279.6603	5
5	<b>548.2715</b>	274.6394	<b>576.2664</b>	288.6368	E	<b>476.2715</b>	238.6394	459.2449	230.1261	4
6	647.3399	324.1736	<b>675.3348</b>	338.1710	V	<b>347.2289</b>	174.1181	330.2023	165.6048	3
7	748.3876	374.6974	776.3825	388.6949	T	<b>248.1605</b>	124.5839	231.1339	116.0706	2
8					K	<b>147.1128</b>	74.0600	130.0863	65.5468	1



NCBI BLAST search of [AEFVEVTK](#)

(Parameters: blastp, nr protein database, expect=20000, no filter, PAM30)

Other BLAST [web gateways](#)

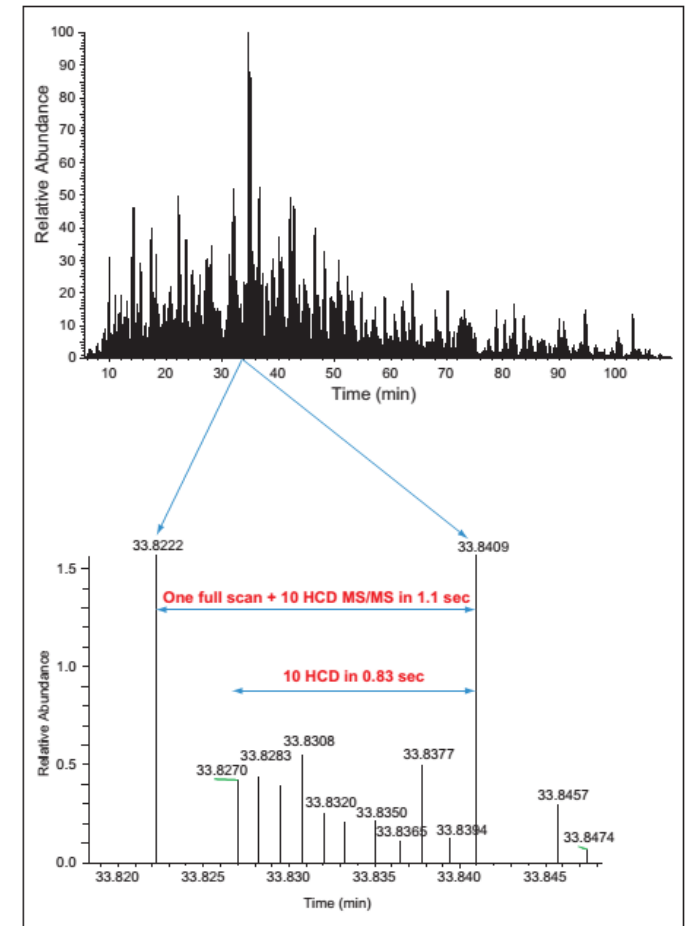
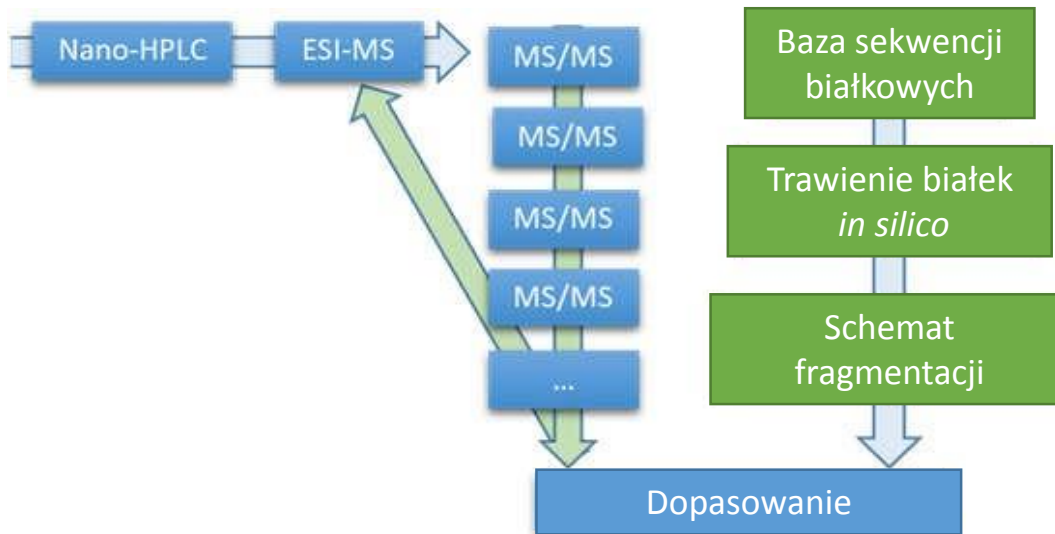
All matches to this query

Score	Mr(calc)	Delta	Sequence
64.2	921.4807	0.0029	<a href="#">AEFVEVTK</a>
14.1	921.4807	0.0029	<a href="#">YIADLEAK</a>
12.8	921.4807	0.0029	<a href="#">EADLFISK</a>
12.8	921.4807	0.0029	<a href="#">VTDFLAEK</a>
11.8	921.4841	-0.0005	<a href="#">VTMETLTK</a>
11.5	921.4807	0.0029	<a href="#">AEALYDIK</a>
11.5	921.4807	0.0029	<a href="#">AEAYLLDK</a>
11.5	921.4807	0.0029	<a href="#">AEYIADIK</a>

# PMF (Ortogonalny zestaw danych i poziom ufności)

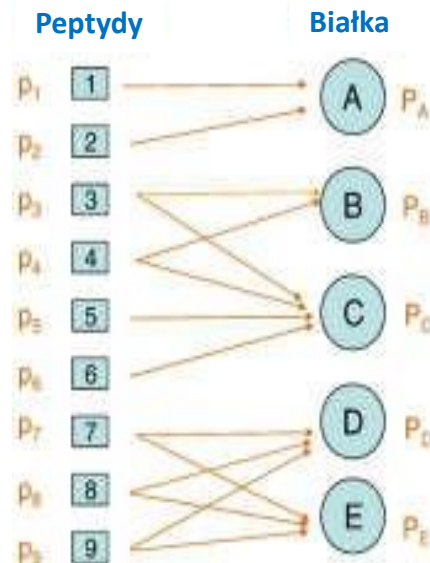
- Ortogonalny zestaw danych i poziom ufności
  - Db : 100 000 sekwencji
  - 500 widm
- Prawdopodobieństwo jednego (dowolnego) widma "przypadkowo" pasujące do sekwencji (niewłaściwe dopasowanie):
  - $1/100\ 000 \times 500 = 5 \cdot 10^{-3}$  (0.005)
- Prawdopodobieństwo błędnego przyporządkowania 2 widm :
  - $5 \cdot 10^{-3} \times 5 \cdot 10^{-3} = 2.5 \times 10^{-5}$
- Znacznie wyższe jest prawdopodobieństwo identyfikacji z co najmniej 2 peptydami pasującymi do tej samej sekwencji białka

# Tandemowa spektroskopia masowa



# 2-PFF

- Identyfikacja peptydu na podstawie danych MS/MS
- Wysoce zautomatyzowany
- Może uzyskać dopasowania z danych marginalnych
- Może być powolny
- Duży zbiór danych

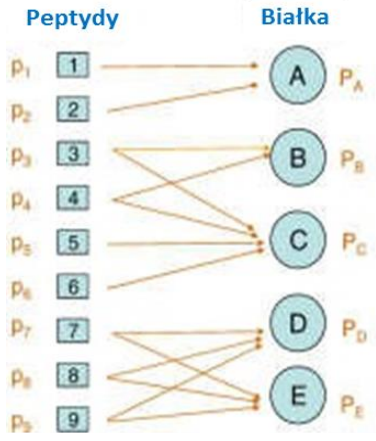


## Lista białek

Lista białek (grup białek) w której zawarte są wszystkie zidentyfikowane peptydy

1. Białko A,  
Prawdopodobieństwo  $P_A$ ,  
Peptydy 1, 2
- 2a. Białko C,  
Prawdopodobieństwo  $P_C$ ,  
Peptydy 3, 4, 5, 6
- 2b. Białko B,  
Prawdopodobieństwo  $P_B = 0$ ,  
Peptydy 3, 4
3. Białka G, E,  
Prawdopodobieństwo  $P_A$ ,  
Peptydy 7, 8, 9

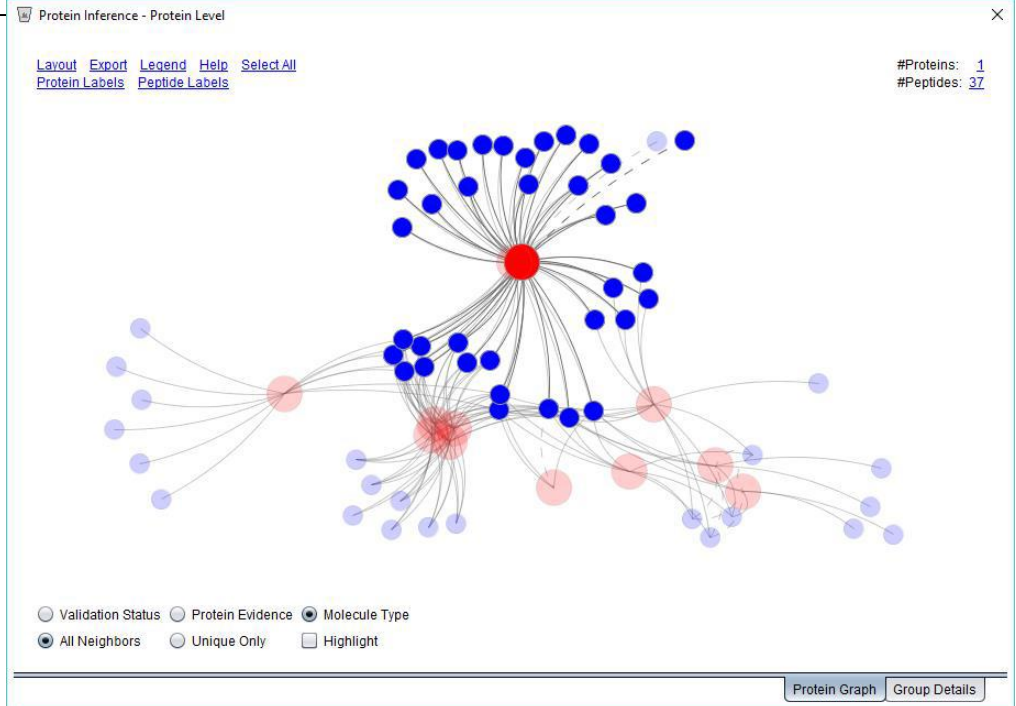




**Lista białek**

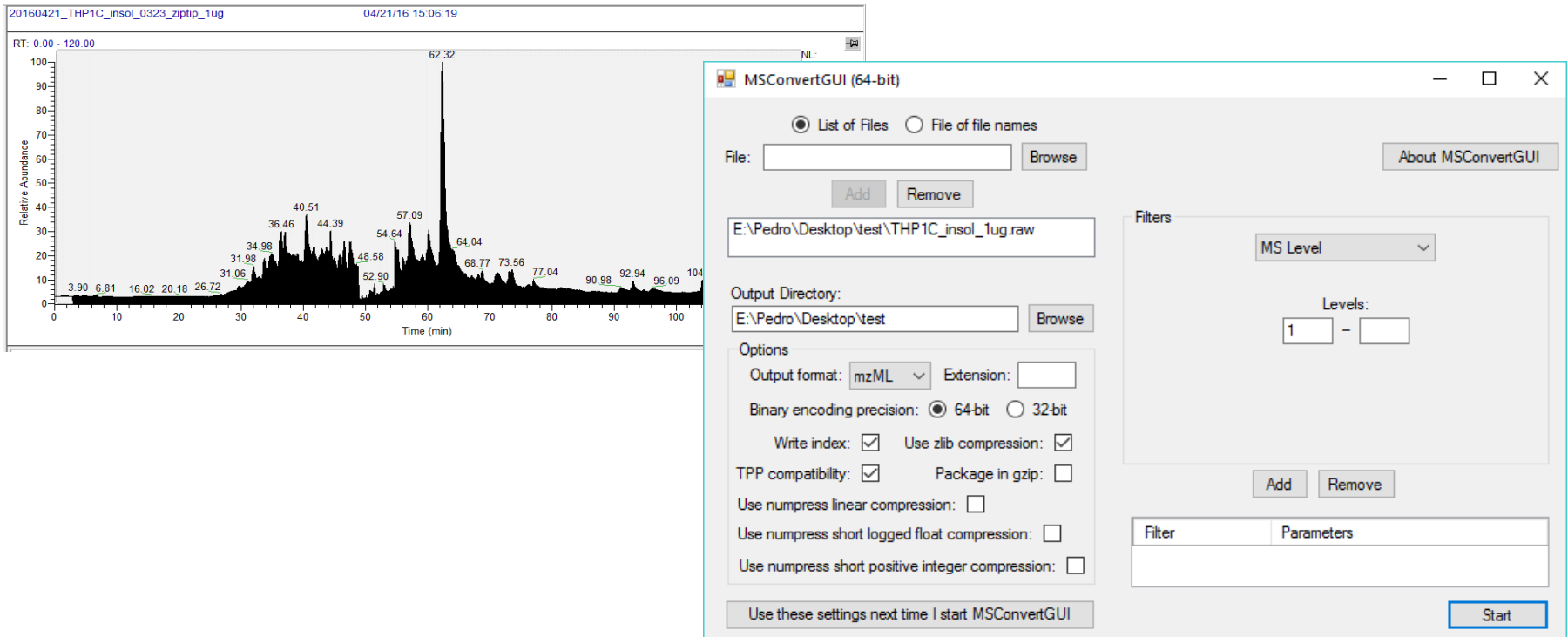
Lista białek (grup białek) w której zawarte są wszystkie zidentyfikowane peptydy

1. Białko A,  
Prawdopodobieństwo  $P_A$ ,  
Peptydy 1, 2
- 2a. Białko C,  
Prawdopodobieństwo  $P_C$ ,  
Peptydy 3, 4, 5, 6
- 2b. Białko B,  
Prawdopodobieństwo  $P_B=0$ ,  
Peptydy 3, 4
3. Białka G, E,  
Prawdopodobieństwo  $P_A$ ,  
Peptydy 7, 8, 9



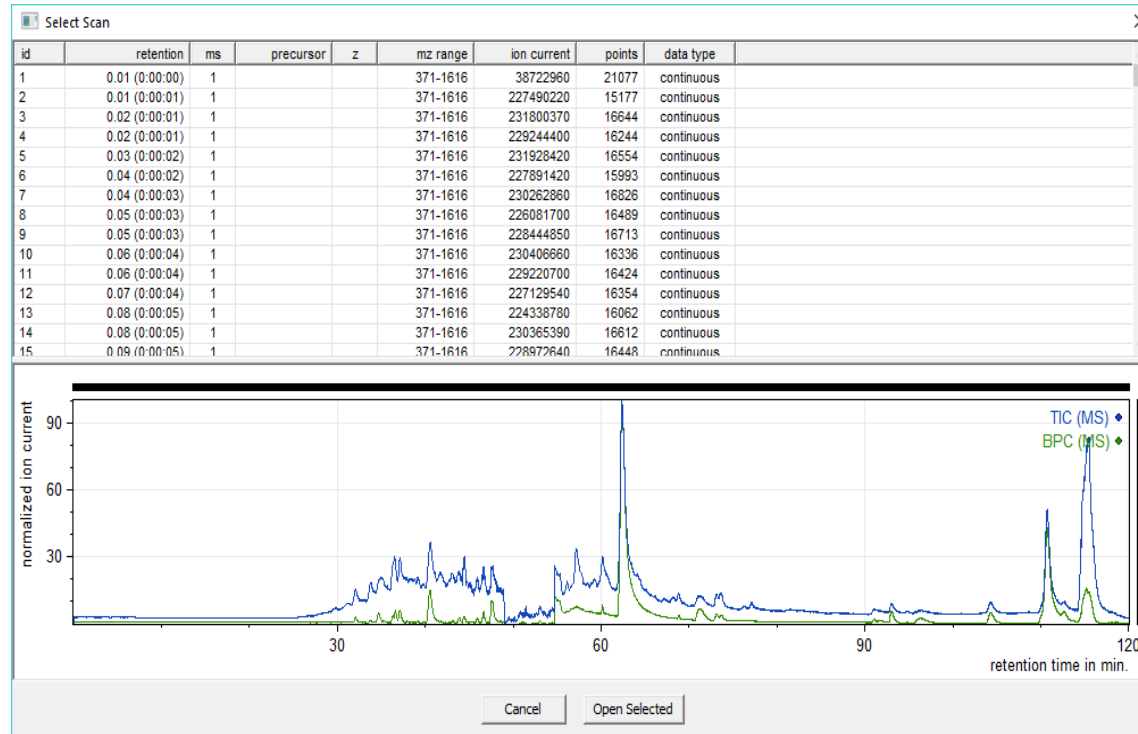
## Zamiana otrzymanych danych na pliki mzML z użyciem MSConverter (ProteoWizard)

Aby obejrzeć dane w aplikacji mMass, należy przekonwertować dane RAW na format mzML.



The image displays two windows. On the left is a chromatogram titled '20160421\_THP1C\_insol\_0323\_ziptip\_1ug' with a date and time of '04/21/16 15:06:19'. The y-axis is 'Relative Abundance' (0-100) and the x-axis is 'Time (min)' (0-104). A major peak is labeled at 62.32 minutes. Other labeled peaks include 3.90, 6.81, 16.02, 20.18, 26.72, 31.06, 31.98, 34.98, 36.46, 40.51, 44.39, 48.58, 52.90, 54.64, 57.09, 64.04, 68.77, 73.56, 77.04, 90.98, 92.94, and 96.09. On the right is the 'MSConvertGUI (64-bit)' window. It has a 'File' field containing 'E:\Pedro\Desktop\test\THP1C\_insol\_1ug.raw'. The 'Output Directory' is 'E:\Pedro\Desktop\test'. Under 'Options', 'Output format' is 'mzML', 'Binary encoding precision' is '64-bit', and 'Write index', 'Use zlib compression', and 'TPP compatibility' are checked. A 'Filters' section shows 'MS Level' set to 1. A 'Start' button is at the bottom right.

# Dane w formacie mMass



- 1) Otwórz plik w formacie mzML
- 2) Przeanalizuj sekwencje otrzymane z widm MS i MS/MS
- 3) Otwórz MS w szczyte piku 45.50.
- 4) Przeanalizuj widmo MS w RT 47.50
- 5) Odszukaj w widmie MS/MS  $m/z = 1216.56(+2)$  w RT 47.53
- 6) Odnajdź wybrane wartości MS/MS w programie Mascot

mMass - MSn spectrum [13720] \*

File View Processing Sequence Tools Libraries Links Window Help



Documents  
\*MSn spectrum [13720]



m/z	int.
185.1644	131033
213.1594	116756
258.0901	187466
282.1801	37448
300.1911	33849
327.1110	108078
369.2132	36687
418.1207	260462
432.1529	16575
435.2018	43404
440.1958	110408
503.2093	67132
531.2040	201657
548.2386	4648
549.2439	83153
579.2233	39620
590.2406	16800
618.2347	43283
620.2822	118666
630.2721	147666
670.2302	20940
687.2548	35384
705.2650	30392
713.3160	42098
721.3287	204016
758.8831	15492
769.3069	29688
776.3419	12510
783.3173	16094
800.3359	28115
804.3346	48481

**Peak Picking**

S/N threshold:

Abs. intensity threshold:

Rel. intensity threshold:  %

Picking height:

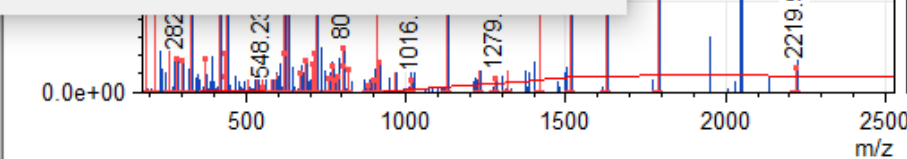
Apply baseline:  See baseline panel

Apply smoothing:  See smoothing panel

Apply deisotoping:  See deisotoping panel

Remove shoulder peaks:

Preview Apply



# Analiza danych MS/MS (Mascot)

Mascot - MS/MS Ion Search

Server: Matrix Science Search

Title: MSn spectrum [13720]

Name: pedro E-mail: p.domingues@ua.pt

Taxonomy: Homo sapiens (human)

Database: SwissProt Enzyme: Trypsin Misc.: 2

Fixed modifications: (0)

- Acetyl (K)
- Acetyl (N-term)
- Acetyl (Protein N-term)
- Amidated (C-term)
- Amidated (Protein C-term)
- Ammonia-loss (N-term C)
- Biotin (K)

Variable modifications: (3)

- ICPL:13C(6)2H(4) (K)
- ICPL:13C(6)2H(4) (N-term)
- ICPL:13C(6)2H(4) (Protein N-term)
- ICPL:2H(4) (K)
- ICPL:2H(4) (Protein N-term)
- iTRAQ4plex (K)
- iTRAQ4plex (N-term)

Show hidden modifications

Precursor m/z: 1216.55735674 Mass type: Monoisotopic

Precursor tolerance: 10.0 ppm Charge: 2+

MS/MS tolerance: 0.05 Da Instrument: Default

Quantitation: None

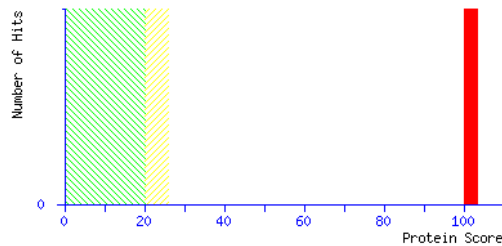
Decoy  Error tolerant Report: AUTO

Modyfikacje:  
Oxidation M  
Carbamidomethyl (C)

# Analiza danych MS/MS (Mascot)

## Mascot Score Histogram

Ions score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event.  
Individual ions scores  $> 20$  indicate peptides with significant homology.  
Individual ions scores  $> 26$  indicate identity or extensive homology ( $p < 0.05$ ).  
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



## Peptide Summary Report

Format As	Peptide Summary	<a href="#">Help</a>			
Significance threshold p<	0.05	Max. number of hits	AUTO		
Standard scoring	<input checked="" type="radio"/> MudPIT scoring <input type="radio"/>	Display non-significant matches	<input type="checkbox"/>	Show sub-sets	0
Show pop-ups	<input checked="" type="radio"/> Suppress pop-ups <input type="radio"/>	Sort unassigned	Decreasing Score	Require bold red	<input type="checkbox"/>
Preferred taxonomy	All entries				

Error tolerant

1. [HS90B\\_HUMAN](#) Mass: 83212 Score: 102 Matches: 1(1) Sequences: 1(1)  
Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4

Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> 1	1216.5574	2431.1002	2431.0970	1.28	0	102	1.4e-09	1	U R.LVSSPCCIVTSTYGWTANMER.I + 2 Carbamidomethyl (C)

### Search Parameters

Type of search : MS/MS Ion Search  
Enzyme : Trypsin  
Variable modifications : [Acetyl \(K\)](#), [Carbamidomethyl \(C\)](#), [Oxidation \(M\)](#)  
Mass values : Monoisotopic  
Protein Mass : Unrestricted  
Peptide Mass Tolerance :  $\pm 10$  ppm  
Fragment Mass Tolerance :  $\pm 0.05$  Da  
Max Missed Cleavages : 2  
Instrument type : Default  
Number of queries : 1

# Analiza danych MS/MS (Mascot)

Sequence Editor

Sequence type: Regular amino acids

Sequence Modifications

Protein Digest

Enzyme: Trypsin

Mass range: 500 - 5000

slice	mis.	m/z
[73-82]	0	1194.647
[360-378]	0	1195.574
[482-502]	1	1196.095
[560-568]	1	1198.595
[339-348]	1	1208.631
[584-604]	0	1216.555
[393-402]	1	1217.692
[457-477]	1	1218.543
[584-604]	0	1224.553
[507-526]	0	1224.576
[457-477]	1	1226.541
[507-526]	0	1232.574
[393-402]	1	1233.687
[457-477]	1	1234.538
[338-347]	1	1236.637
[169-180]	1	1237.628
[96-107]	0	1242.705
[429-438]	1	1246.646

Peptide Fragmentation

Match Fragments

Tolerance: 10.0

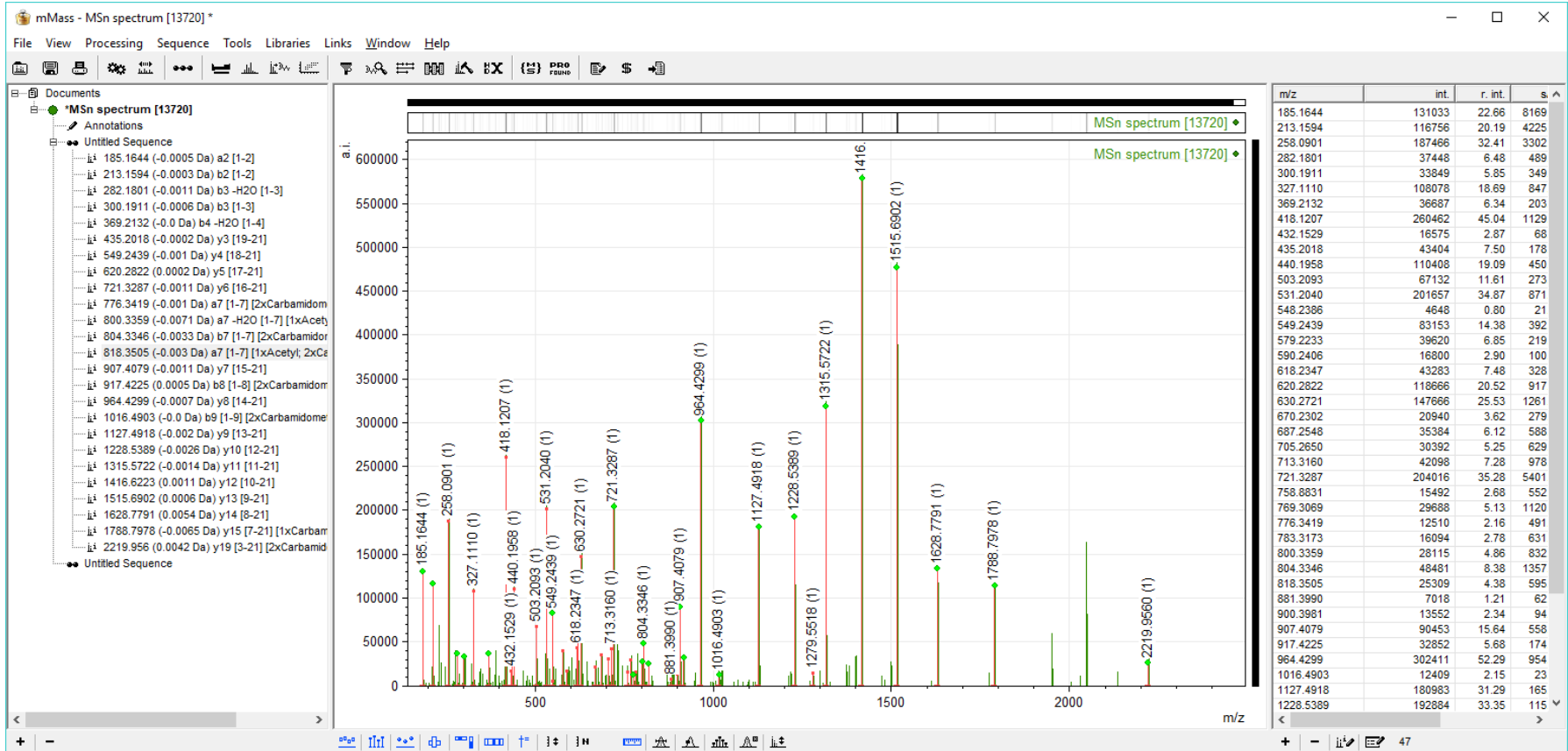
Remove:  Annotated  Matched  Unselected  Isotopes  Unknown

error (ppm)

m/z

Peptide	Charge	m/z
[1-14]	1	1200.120
[1-12]	2	630.3098
[1-12]	1	1246.646

# Analiza danych MS/MS (Mascot)

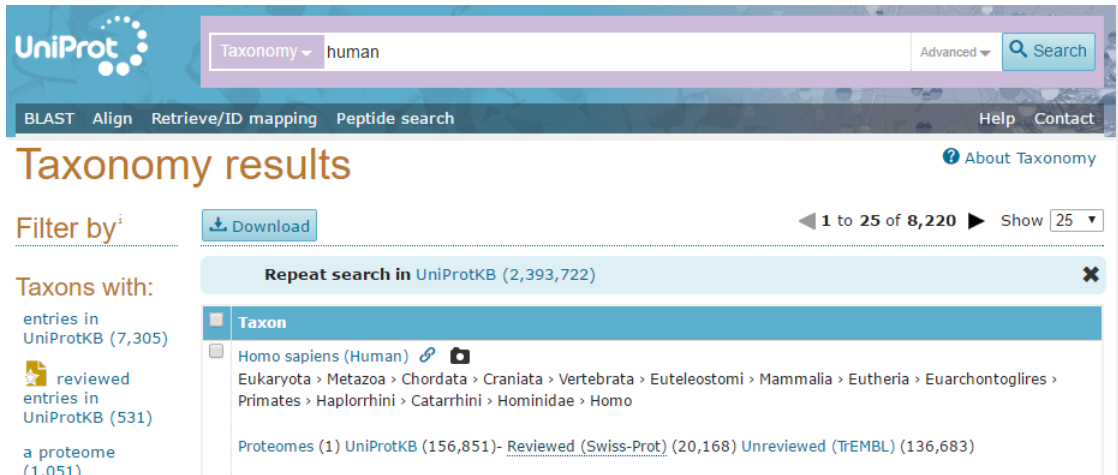




- Dostępne narzędzia
  - Kalkulator mas
  - Przetwarzanie masy na wzory

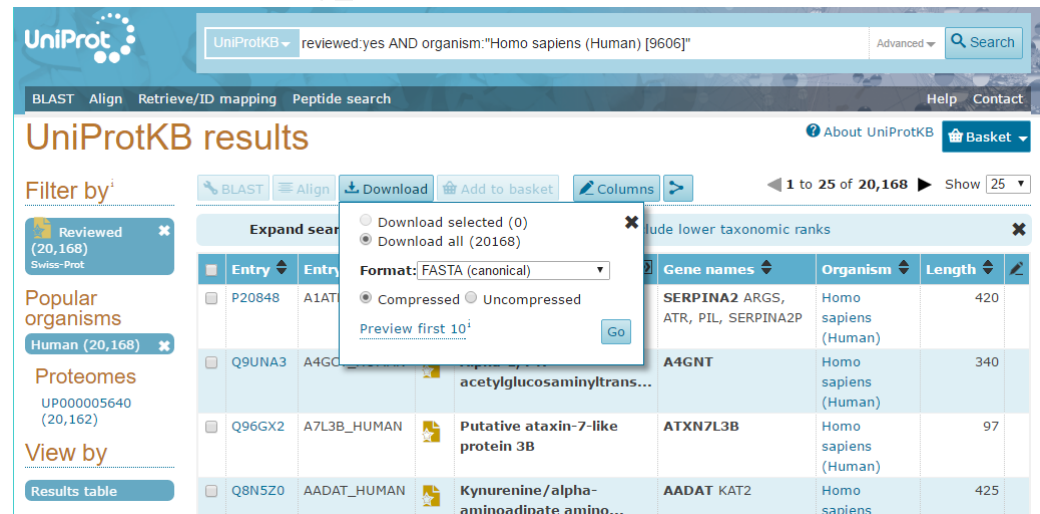
# Generowanie bazy danych FASTA z Uniprot (SwissProt)

1) Należy pobrać plik FASTA proteomu Homo Sapiens ze strony Uniprot.



The screenshot shows the UniProt Taxonomy results page. The search criteria are set to 'Taxonomy' and 'human'. The page displays 'Taxonomy results' with a filter by 'Taxon'. A 'Download' button is visible. The search results show 'Repeat search in UniProtKB (2,393,722)'. The taxonomic path is: Eukaryota > Metazoa > Chordata > Craniata > Vertebrata > Euteleostomi > Mammalia > Eutheria > Euarchontoglires > Primates > Haplorrhini > Catarrhini > Hominoidea > Homo. The results include 1 proteome in UniProtKB (156,851), 20,168 reviewed (Swiss-Prot) entries, and 136,683 unreviewed (TrEMBL) entries.

2) Należy wybrać recenzowaną wersję plików FASTA (SwissProt)

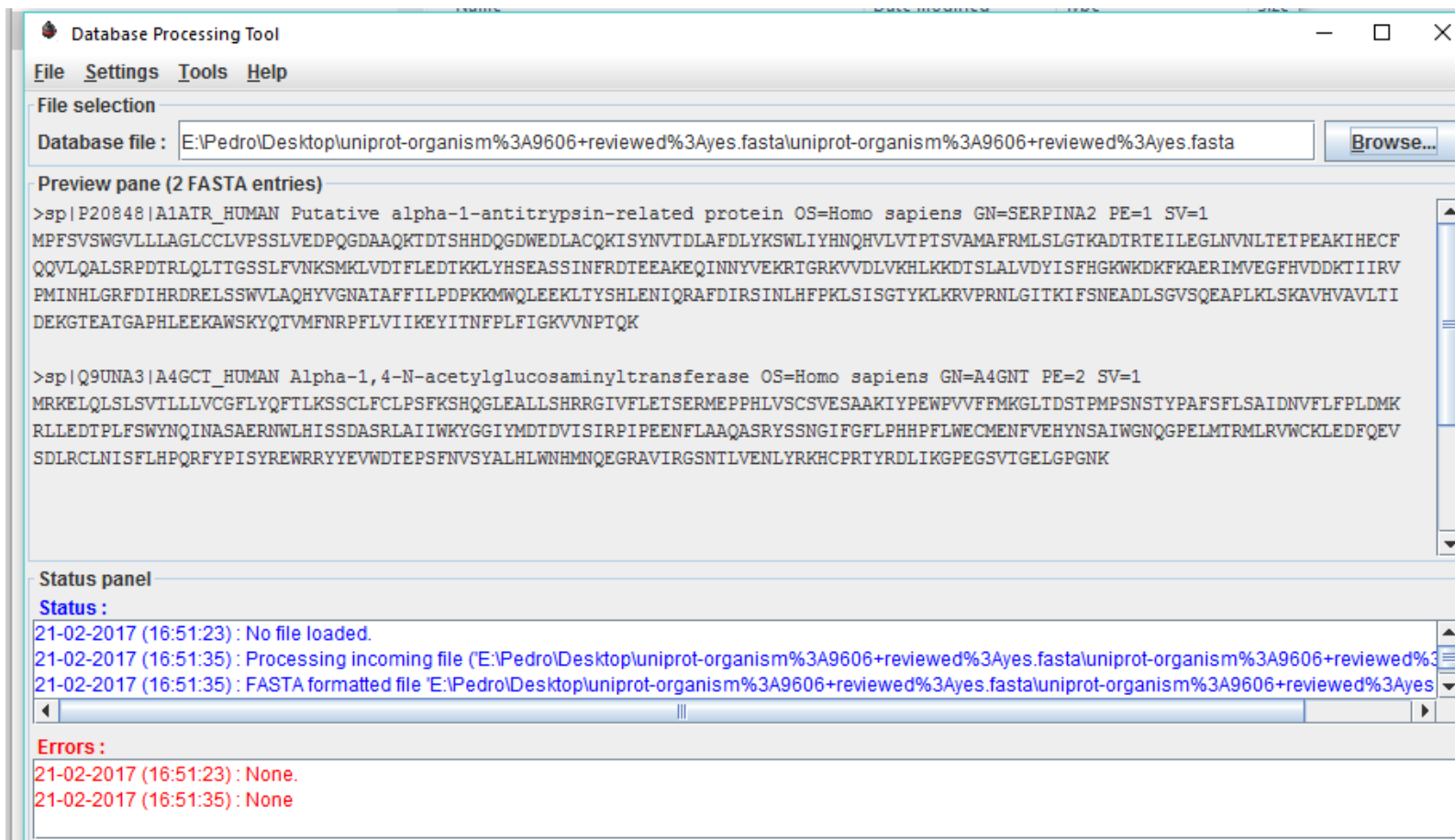


The screenshot shows the UniProtKB results page for the search criteria 'reviewed:yes AND organism:"Homo sapiens (Human) [9606]". The page displays 'UniProtKB results' with a filter by 'Reviewed (20,168) Swiss-Prot'. A 'Download' button is visible. A dialog box is open over the 'Download' button, showing options for 'Format: FASTA (canonical)', 'Compressed', and 'Uncompressed'. The dialog also includes a 'Preview first 10' button and a 'Go' button. The background shows a table of results with columns for 'Entry', 'Gene names', 'Organism', and 'Length'. The table contains entries for SERPINA2, A4GNT, ATXN7L3B, and AADAT KAT2.

Entry	Gene names	Organism	Length
P20848	SERPINA2 ARGs, ATR, PIL, SERPINA2P	Homo sapiens (Human)	420
Q9UNA3	A4GNT	Homo sapiens (Human)	340
Q96GX2	Putative ataxin-7-like protein 3B	Homo sapiens (Human)	97
Q8N5Z0	Kynurenine/alpha-aminoadipate amino...	Homo sapiens	425

# Narzędzie do przetwarzania danych Dbtoolkit

Można przeglądać i edytować informacje w pliku FASTA za pomocą programu Dbtoolkit



**Database Processing Tool**

File Settings Tools Help

File selection

Database file : E:\Pedro\Desktop\uniprot-organism%3A9606+reviewed%3Ayes.fasta

Preview pane (2 FASTA entries)

```
>sp|P20848|A1ATR_HUMAN Putative alpha-1-antitrypsin-related protein OS=Homo sapiens GN=SERPINA2 PE=1 SV=1
MPFSVSWGVLALLAGLCCCLVPSLVEDPQGDAAQKTDTSHHDDQGDWEDLACQKISYNVTDLAFDLYKSWLIYHNQHVLTPTSVAMAFRMLSLGKADTRTEILEGLNVNLTETPEAKIHECF
QQVLQALSRPDTLRLQLTTGSSSLFVNKSMKLVDTFLEDTKKLYHSEASSINFRDTEEAKEQINNYVEKRTGRKVVLDLVKHLKDDTSLALVDYISFHGKWKDKFKAERIMVEGFHVDDKTIIRV
PMINHLGRFDIHRDRELSSWVLAQH YVGNATAFFILPDPKMMQLEEKLTYSHLENIQRAFDIRSINLHFPKLSISGTYKLRVPRNLGITKIFSNEADLSGVSQEAPLKLKSAVHVAVLTI
DEKGTATGAPHLEEKAWSKYQTMFNRPFVLIKEYITNFPLFIGKVVNPTQK

>sp|Q9UNA3|A4GCT_HUMAN Alpha-1,4-N-acetylglucosaminyltransferase OS=Homo sapiens GN=A4GNT PE=2 SV=1
MRKELQLSLSVILLVCGFLYQFTLKSCLFCLPSFKSHQGLEALLSHRRGIVFLETSERMEPHLVSCSVESAAKIYPEWVVFVFMKGLTDSTPMPSNSTYPAFSFLSAIDNVFLFPLDMK
RLLEDTPLF SWYNQINASAERNWLHISDASRLAIWKYGGIYMDTDVISIRPIPEENFLAAQASRYSSNGIFGFLPHHPFLWECMENFVEHYNSAIWGNQGPPELMTRMLRVWCKLEDFQEV
SDLRCLNISFLHPQRFYPISYREWRRYYEVWDTEPSFNVSVALHLWNHMNQEGRAVIRGSNTLVENLYRKHCPRTYRDLIKGPEGSVTGELGPGNK
```

Status panel

Status :

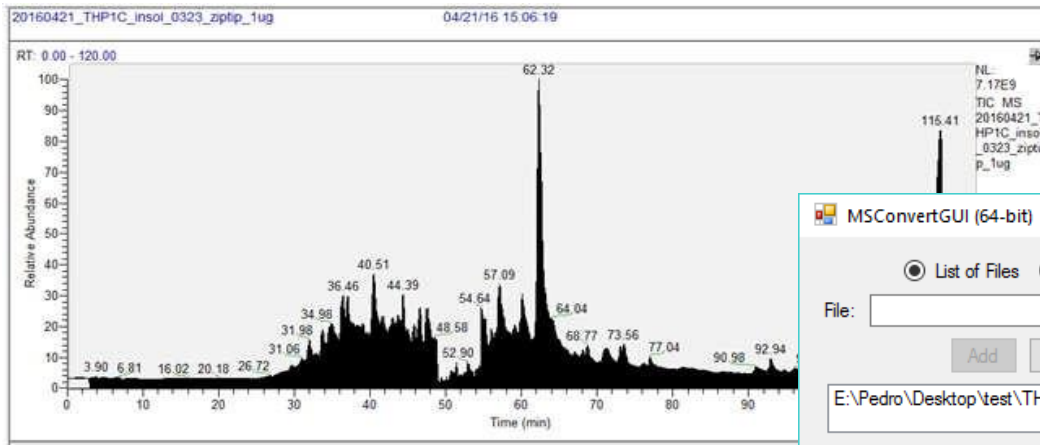
21-02-2017 (16:51:23) : No file loaded.  
21-02-2017 (16:51:35) : Processing incoming file ('E:\Pedro\Desktop\uniprot-organism%3A9606+reviewed%3Ayes.fasta\uniprot-organism%3A9606+reviewed%3Ayes.fasta')  
21-02-2017 (16:51:35) : FASTA formatted file 'E:\Pedro\Desktop\uniprot-organism%3A9606+reviewed%3Ayes.fasta\uniprot-organism%3A9606+reviewed%3Ayes.fasta'

Errors :

21-02-2017 (16:51:23) : None.  
21-02-2017 (16:51:35) : None

# Przetwarzanie danych przez MSConverter (ProteoWizard)

- MSConverter w programie ProteoWizard przekształca dane HPC-MS uzyskane podczas pracy spektrometru mas (plik RAW) na format otwierany przez SearchGUI (pliki MGF).



MSConvertGUI (64-bit)

List of Files  File of file names

File:  Browse

Add Remove

E:\Pedro\Desktop\test\THP1C\_insol\_1ug.raw

About MSConvertGUI

Output Directory:  Browse

Options

Output format:  Extension:

Binary encoding precision:  64-bit  32-bit

Write index:  Use zlib compression:

TPP compatibility:  Package in gzip:

Use numpress linear compression:

Use numpress short logged float compression:

Use numpress short positive integer compression:

Use these settings next time I start MSConvertGUI

Filters

MS Level  ▾

Levels:  -

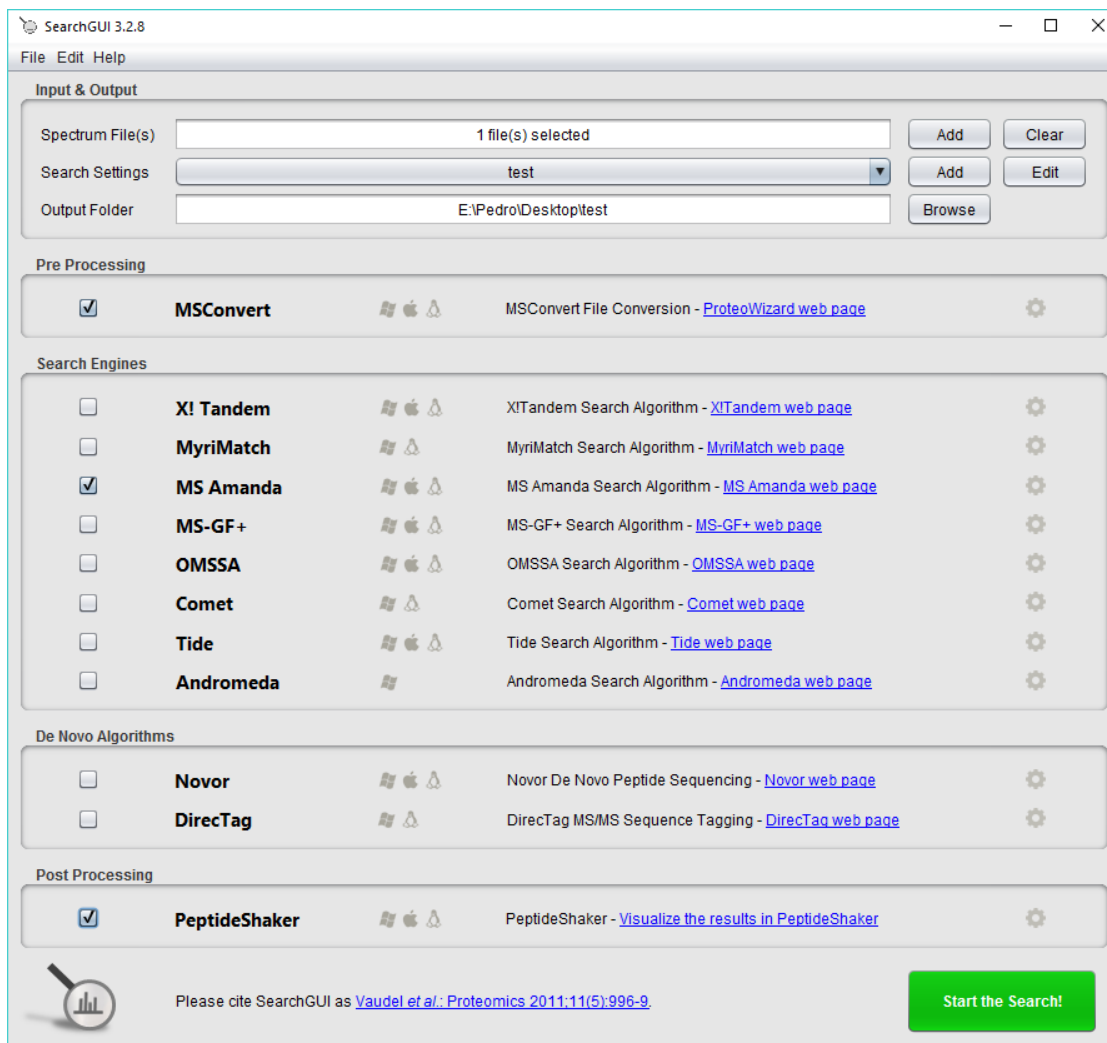
Add Remove

Filter	Parameters

Start

# Wyszukiwarka SearchGUI

- 1) Otwórz SearchGUI.
- 2) W ustawieniach wyszukiwania wybierz edycję i wypełnij formularz, jak pokazano.
- 3) Wybrano modyfikacje zmienne (dlaczego?).
- 4) Możesz także skonfigurować PeptideShaker, aby otworzyć plik wyników, jak pokazano poniżej.
- 5) Wyszukiwarka SearchGUI zapyta, czy chcesz utworzyć plik fasta concatenated\_target\_decoy. Wybierz „yes” (dlaczego?)



# Wyszukiwarka SearchGUI

Wybierz opcje wyszukiwania  
i baze danych FASTA

Wybrano modyfikacje zmienne.

Dlaczego?

Search Settings - test

Database

Database (FASTA)

Modifications

Fixed Modifications (0)

Name	Mass
------	------

Variable Modifications (3)

Name	Mass
<input checked="" type="checkbox"/> Oxidation of M	15.99
<input checked="" type="checkbox"/> Carbamidomethylation of C	57.02
<input checked="" type="checkbox"/> Acetylation of K	42.01

Most Used Modifications

Name	Mass
<input checked="" type="checkbox"/> Acetylation of protein N-term	42.01
<input checked="" type="checkbox"/> Deamidation of N	0.98
<input checked="" type="checkbox"/> Deamidation of Q	0.98
<input checked="" type="checkbox"/> Phosphorylation of S	79.97
<input checked="" type="checkbox"/> Phosphorylation of T	79.97
<input checked="" type="checkbox"/> Phosphorylation of Y	79.97
<input checked="" type="checkbox"/> Pyroglutamate from E	-18.01
<input checked="" type="checkbox"/> Pyroglutamate from Q	-17.03
<input checked="" type="checkbox"/> Pyroglutamate from carbamidomethylated C	-17.03
<input checked="" type="checkbox"/> TMT 10-plex of K	229.16
<input checked="" type="checkbox"/> TMT 10-plex of peptide N-term	229.16
<input checked="" type="checkbox"/> TMT 6-plex of K	229.16
<input checked="" type="checkbox"/> TMT 6-plex of peptide N-term	229.16
<input checked="" type="checkbox"/> ITRAQ 4-plex of K	144.10

Protease & Fragmentation

Digestion

Enzyme

Specificity

Max Missed Cleavages

Fragment Ion Types

Precursor m/z Tolerance

Fragment m/z Tolerance

Precursor Charge  -

Isotopes  -

## Wybierz opcje programu PeptideShaker

PeptideShaker Settings

**Peptide Shaker**

Location

**Project Details**

Project Name

Sample Name  Replicate

**Output**

Output File

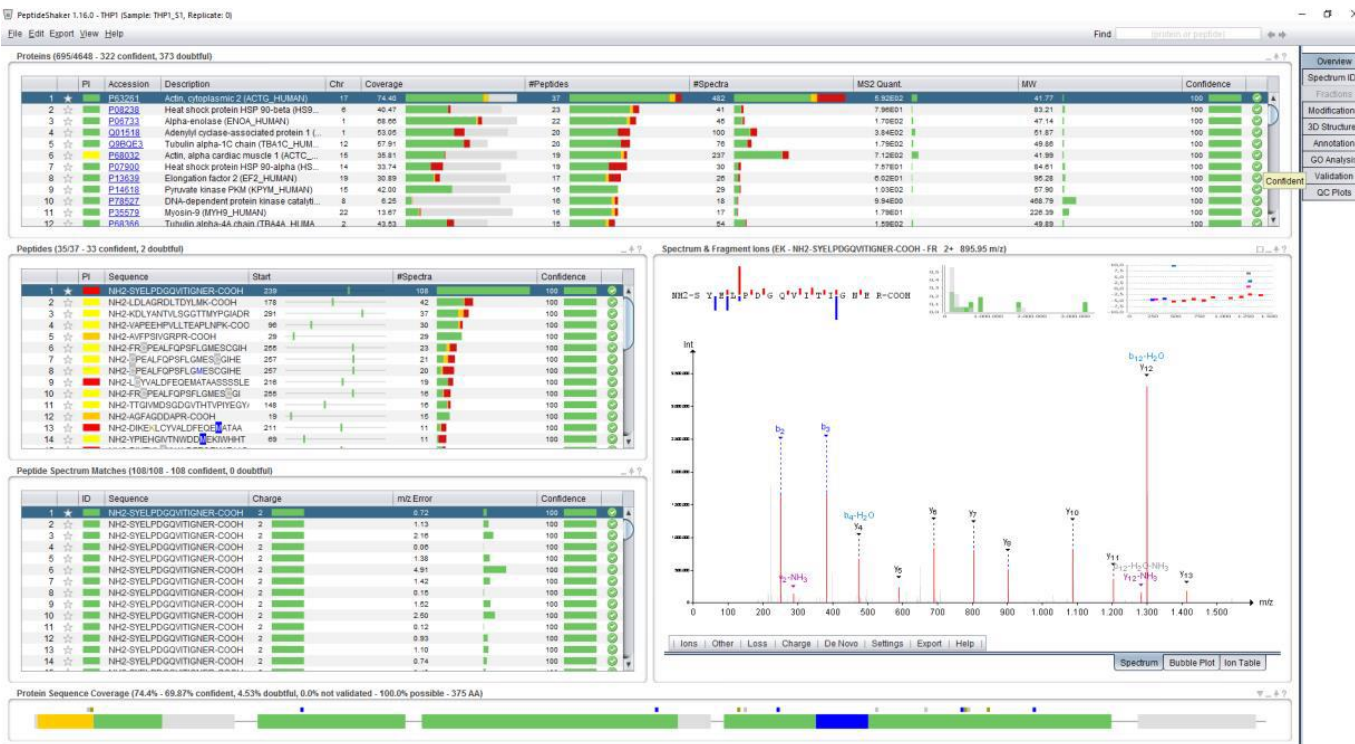
**Advanced Settings (see help for details)**

Project

Mascot Files

# Generowanie i ocena wyników. PeptideShaker do wizualizacji peptydów i białek oraz do walidacji. Analiza PTM

- 1) Po przetworzeniu danych przez SearchGUI (~3 minuty) wyniki zostaną otworzone w oknie PeptideShaker.
- 2) Jesteśmy w stanie zobaczyć ok 700 zidentyfikowanych białek, chociaż 373 zostały sklasyfikowane jako „niepewne” (dlaczego?).
- 3) Można również zobaczyć informacje o peptydach zidentyfikowanych dla każdego białka i widmach masowych z opisanym wzorem fragmentacji.





# Dopasowanie widma do peptydu (peptide-spectrum match - PSM)

Quality Control (beta)

**General Settings**

Mark as Doubtful

- Hits obtained on small databases (<1000 protein sequences)
- Datasets with a low number of target hits
- Hits near the confidence threshold (margin= 1 x resolution)

**Protein Filters**

	Name	Description
1	>=2 confident peptides	Number of confident peptides filter
2	>=2 confident spectra	Number of confident spectra filter

**Peptide Filters**

	Name	Description
1	One confident PSM	Number of confident PSMs filter

**PSM Filters**

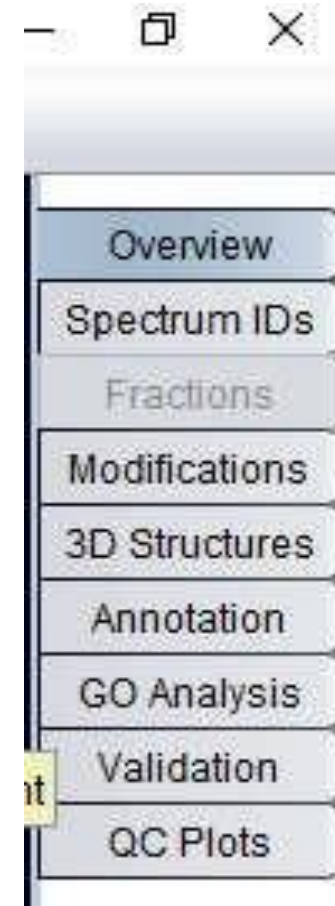
	Name	Description
1	Fragment Ion Sequence Coverage	Sequence coverage filter by fragment ions
2	Mass deviation	Precursor m/z deviation probability

Right-click in the tables to edit the filters.

OK Cancel

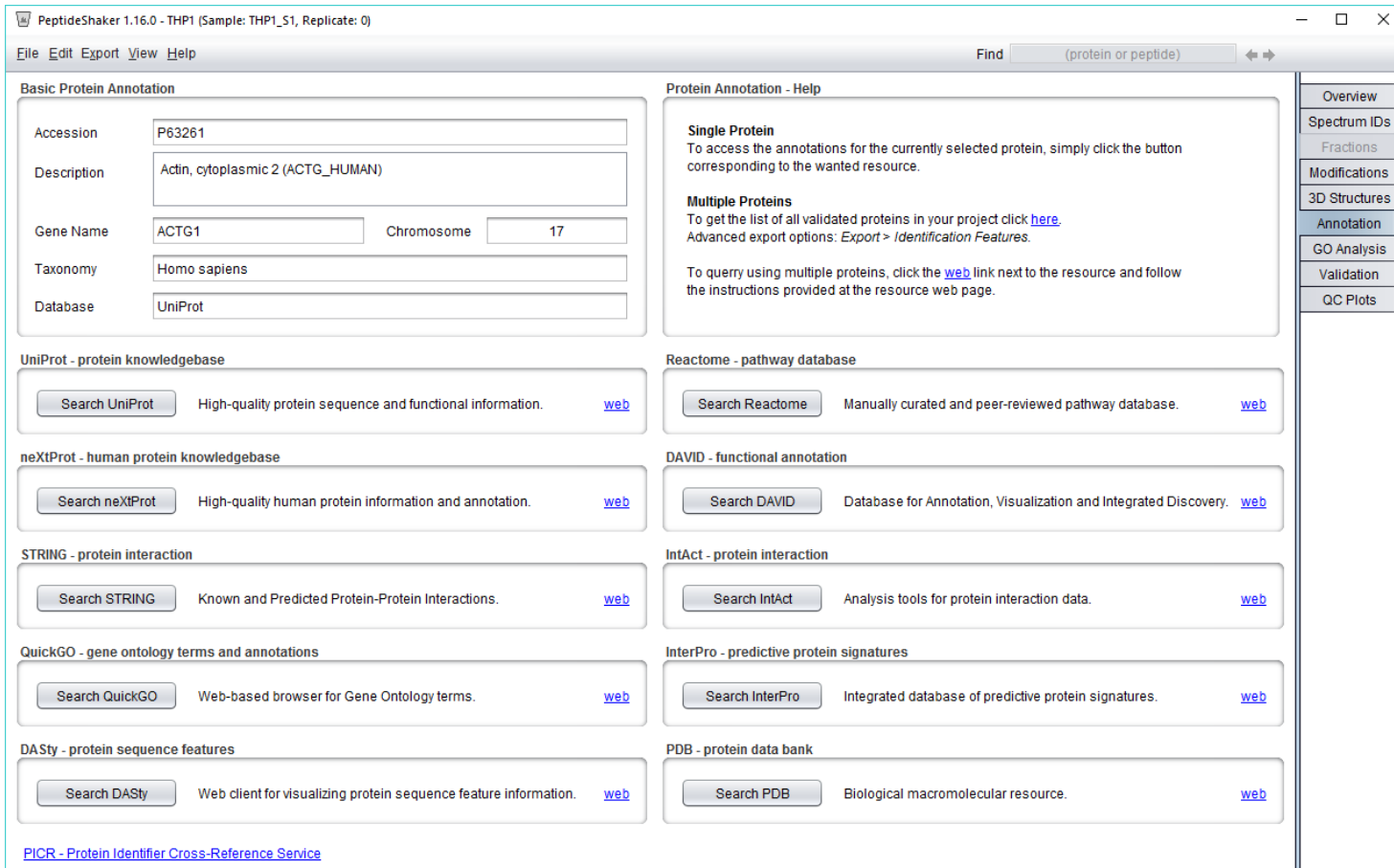
# Analiza danych: informacje o białku, analiza szlaków i ontologia genowa

Zapoznaj się z zaawansowanymi opcjami analizy danych modułu PeptideShaker otwierając kartę modyfikacji (GO analysis tab).



# Analiza danych: informacje o białku, analiza szlaków i ontologia genowa

W zakładce adnotacji można dodawać komentarze do każdego białka. Jeśli jednak chcesz opisać wiele białek, musisz wyeksportować wyniki (domyślny raport białkowy) i klikając link do strony internetowej obok zasobu, postępować zgodnie z instrukcjami podanymi na stronie internetowej zasobów.

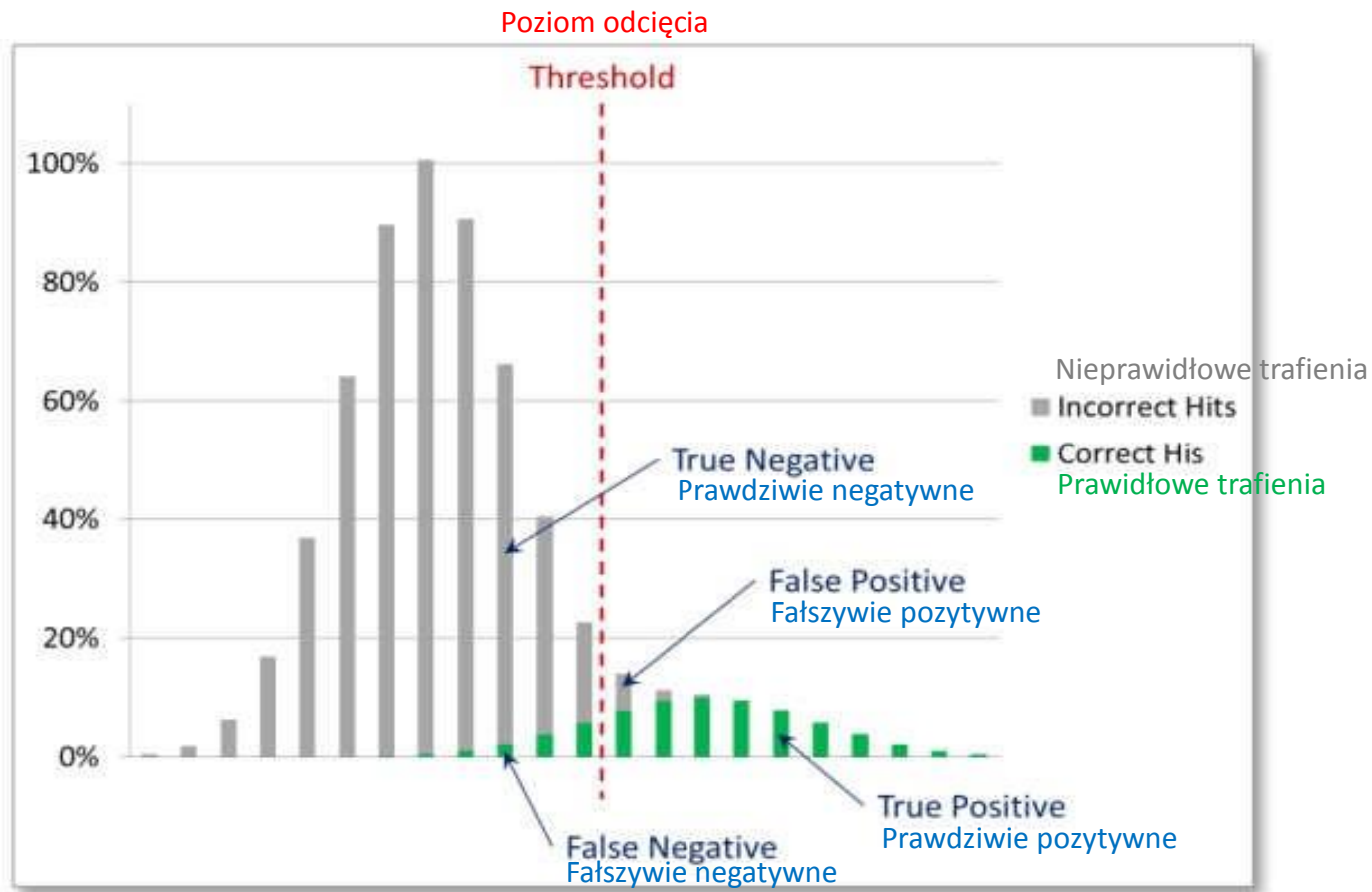


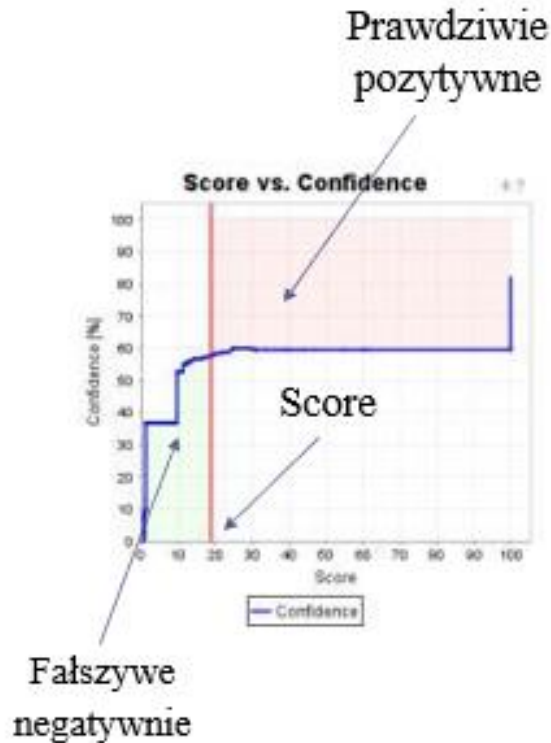
The screenshot displays the PeptideShaker 1.16.0 interface for a sample labeled 'THP1 (Sample: THP1\_S1, Replicate: 0)'. The main window is divided into several sections:

- Basic Protein Annotation:** A form with fields for Accession (P63261), Description (Actin, cytoplasmic 2 (ACTG\_HUMAN)), Gene Name (ACTG1), Chromosome (17), Taxonomy (Homo sapiens), and Database (UniProt).
- Protein Annotation - Help:** A text box providing instructions for 'Single Protein' and 'Multiple Proteins' annotations, including links for advanced export options and querying multiple proteins.
- Database Search Panels:** A grid of search buttons for various databases, each with a brief description and a 'web' link:
  - UniProt - protein knowledgebase:** High-quality protein sequence and functional information.
  - Reactome - pathway database:** Manually curated and peer-reviewed pathway database.
  - neXtProt - human protein knowledgebase:** High-quality human protein information and annotation.
  - DAVID - functional annotation:** Database for Annotation, Visualization and Integrated Discovery.
  - STRING - protein interaction:** Known and Predicted Protein-Protein Interactions.
  - IntAct - protein interaction:** Analysis tools for protein interaction data.
  - QuickGO - gene ontology terms and annotations:** Web-based browser for Gene Ontology terms.
  - InterPro - predictive protein signatures:** Integrated database of predictive protein signatures.
  - DASy - protein sequence features:** Web client for visualizing protein sequence feature information.
  - PDB - protein data bank:** Biological macromolecular resource.

At the bottom left, there is a link to 'PICR - Protein Identifier Cross-Reference Service'. On the right side of the interface, a vertical sidebar contains navigation buttons for 'Overview', 'Spectrum IDs', 'Fractions', 'Modifications', '3D Structures', 'Annotation', 'GO Analysis', 'Validation', and 'QC Plots'.

# Walidacja





Liczba wykrytych białek docelowych i „wabików”



Aby uzyskać 100% pokrycia, potrzebujemy 50% FRD

Pokaż wyniki tej analizy:

- Liczba zidentyfikowanych białek
- 3 białka o największej intensywności pików
- Wybierz grupę na podstawie analizy GO
  - Pokaż analizę sieci skojarzeń białka funkcjonalnego w tej grupie
  - Opisz jedno białko stanowiące węzeł centralny



Ten projekt został sfinansowany przy wsparciu Komisji Europejskiej.

Niniejsze opracowanie odzwierciedla jedynie poglądy autorów,  
a Komisja nie ponosi odpowiedzialności za jakiegokolwiek wykorzystanie  
informacji zawartych w tej publikacji.

Niniejsze e-materiały udostępnione są na licencji Creative Commons  
Attribution-NonCommercial 3.0 License CC BY-NC 3.0.

