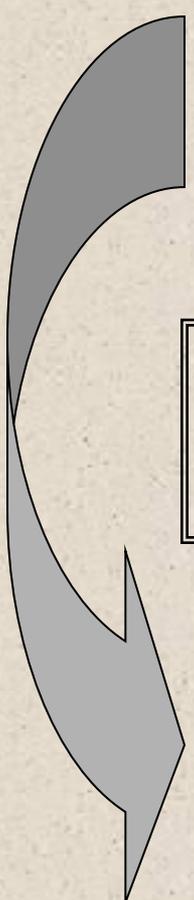


# **МАСС-СПЕКТРОМЕТРИЯ В ПРОТЕОМНЫХ ИССЛЕДОВАНИЯХ**

**Часть 2: Приложения**

# ПРОТЕОМИКА – набор высокотехнологичных методов установления состава сложных смесей белков

- 1) определение количества того или иного белка в образце;
- 2) идентификация белка;
- 3) уточнение первичной структуры;
- 4) определение пост-трансляционных модификаций.



2D-электрофорез  
Специфический гидролиз  
MALDI-MS

Специфический гидролиз  
Хроматография  
ESI-MS-MS, микросеквенирование

Белковые чипы

# Масс-спектрометрия целых белков

В базе данных  
NCBI **2** более  
500 000 записей

Для идентификации белка по его молекулярной массе необходима точность ее измерения  $< 2\text{ppm}$

Точность недостижима при использовании времяпролетной масс-спектрометрии, ионных ловушек

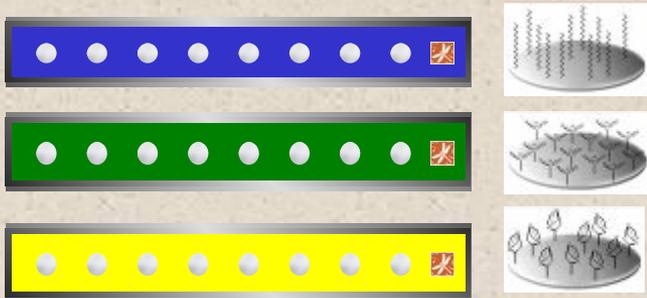


Такая точность доступна при использовании масс-спектрометрии ИЦР

Исучаемый белок как правило отличен от имеющегося в базе данных

Пример: при анализе всех цитозольных белков бактерии *Rhodobacter* с использованием масс-спектрометрии ИЦР идентифицировано около 10% белков, пики которых наблюдались в спектре. Массы остальных белков были отличны от рассчитанных по гену из-за наличия пост-трансляционных модификаций.

# Профилирование белков



Подложки с разными типами поверхности:  
обращенная фаза, анионообменная,  
катионообменная.

## Биомаркер рака яичников

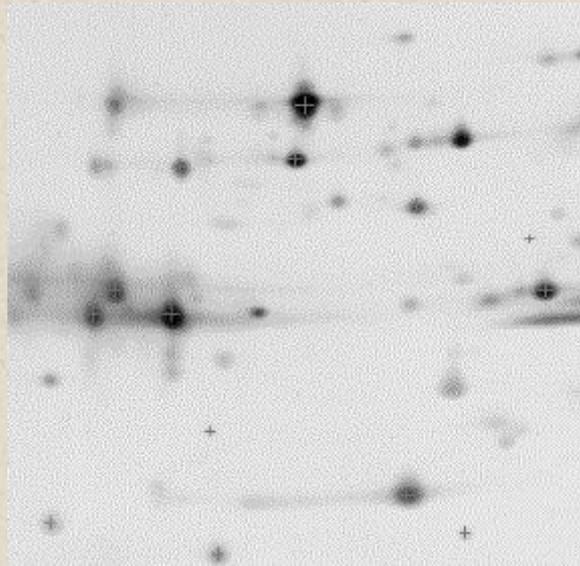


Для идентификации белков отличий необходимо их выделение.

Детектируются, как правило, хорошо представленные белки.

В данном случае –  
сывороточный амилоид А1

# Стратегия анализа: 2DE + MALDI-MS



2D электрофорез белков



Компьютерный анализ изображений гелей



Вырезание и трипсинолиз белков в  
фрагментах геля



Экстракция пептидов и получение  
MALDI масс-спектра суммарного  
гидролизата



Поиск в базе данных  
измеренных масс пептидов  
(пептидный фингерпринт)



Post Source Decay-MS, TOF-TOF  
фрагментация отдельных пептидов



Поиск в базе данных  
измеренных масс фрагментов

# 2D-Электрофорез

Первое направление – изоэлектрическая фокусировка

Второе направление – разделение по массам в полиакриламидном геле

Белки, растворенные в неионном детергенте, вносятся в гель с фиксированным градиентом pH

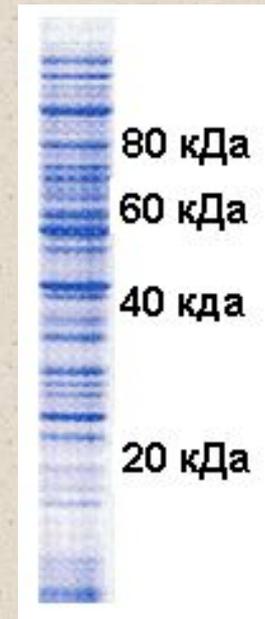
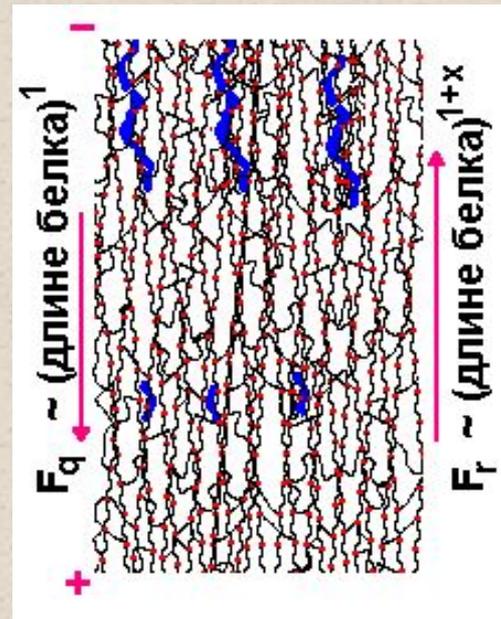


Эффективное разделение по массам белков 10-200 кДа



Электрофорез по Лэмбли:

Добавляется додецилсульфат Na: сильный детергент вносит отрицательный заряд



# 2D-Электрофорез

## Достоинства

Реальные параметры разделения  
Эффективное разделение  
Полуколичественный метод

## Ограничения

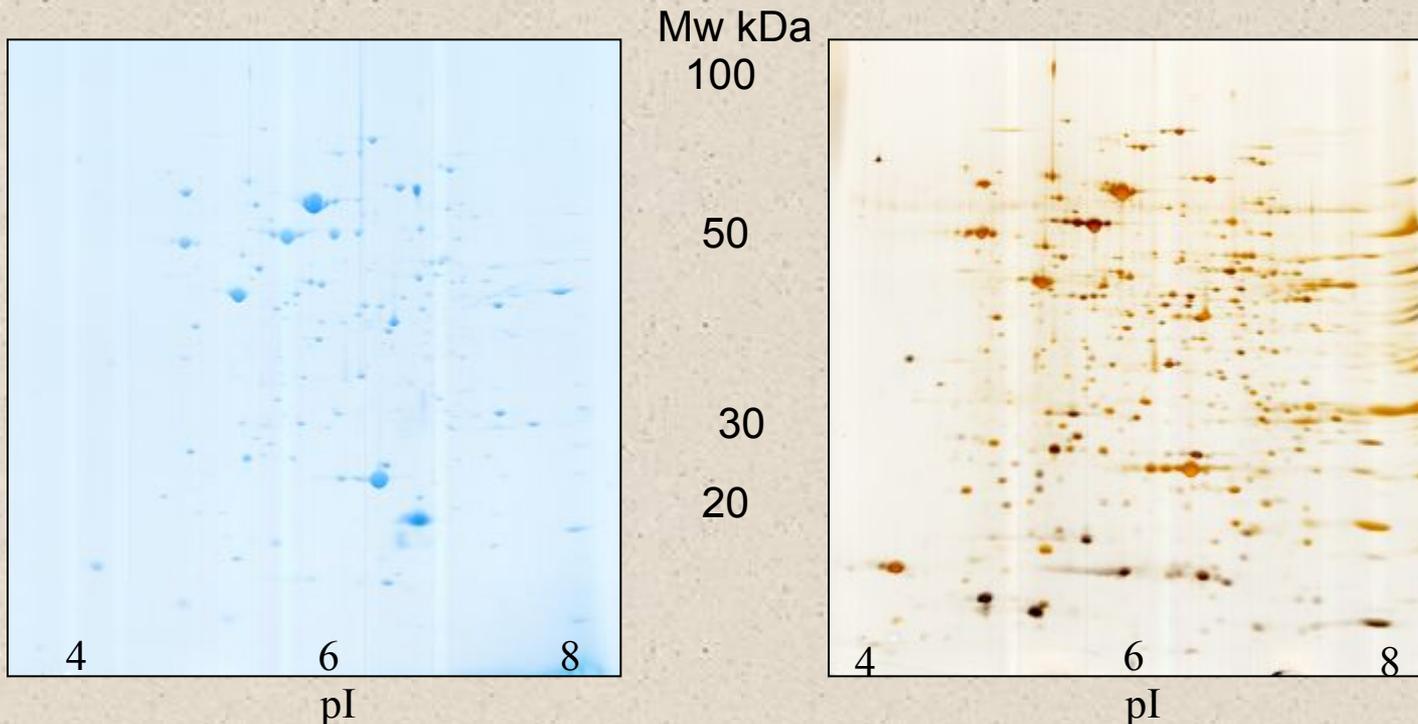
Проблемы с кислыми, щелочными  
и гидрофобными белками  
Невысокая воспроизводимость

Окраска:

- Кумасси голубым
- Серебром, совместимая с MS анализом
- Серебром, классическая

Чувствительность:

50 нг/пятно  
1 нг/пятно  
0.1 нг/пятно



# 2D-Электрофорез

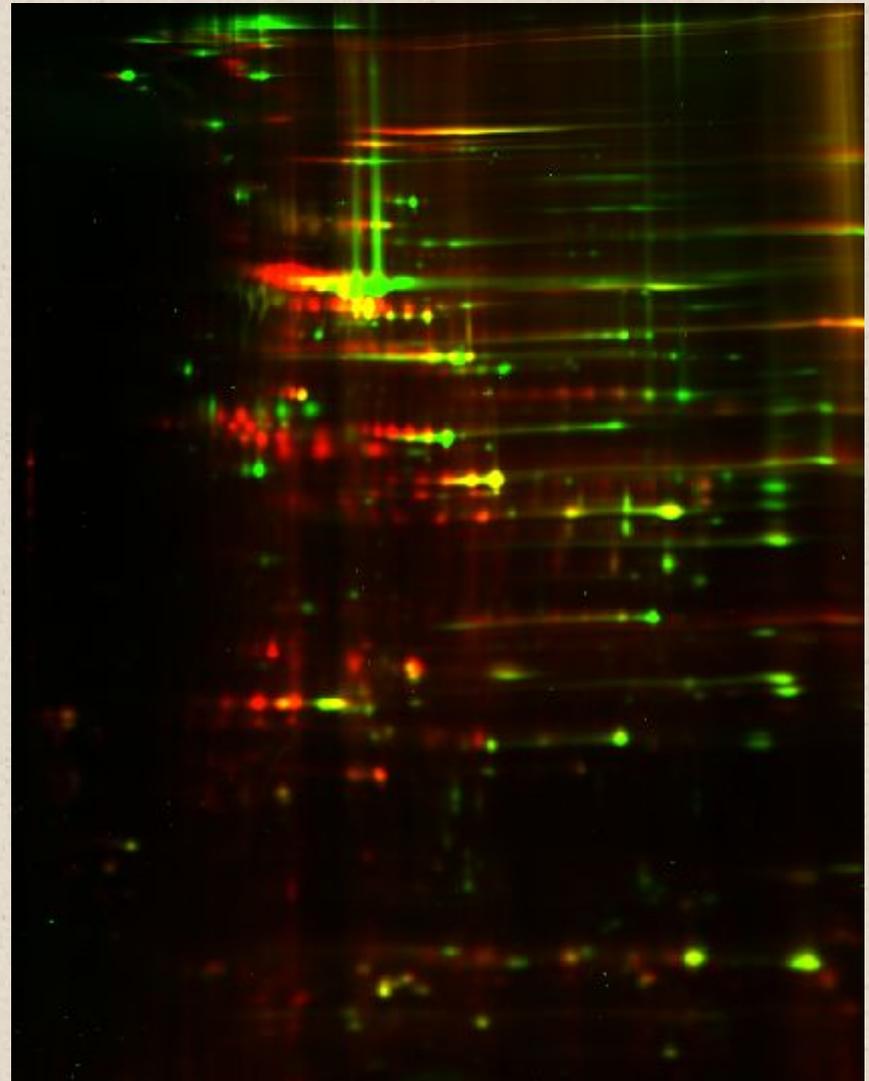
## Окраска DIGE Cy3/Cy5

Электрофорез 2-х предварительно окрашенных белков *Mycoplasma gallisepticum* проводится на одном геле.

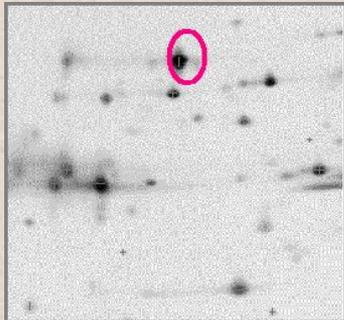
- Зеленым (Cy3) – белки из контрольных клеток,
- Красным (Cy5) – белки из обработанных клеток.

**Достоинства:** Очень чувствителен, прекрасное сопоставление пятен, очень точен и удобен для оценки экспрессии белка

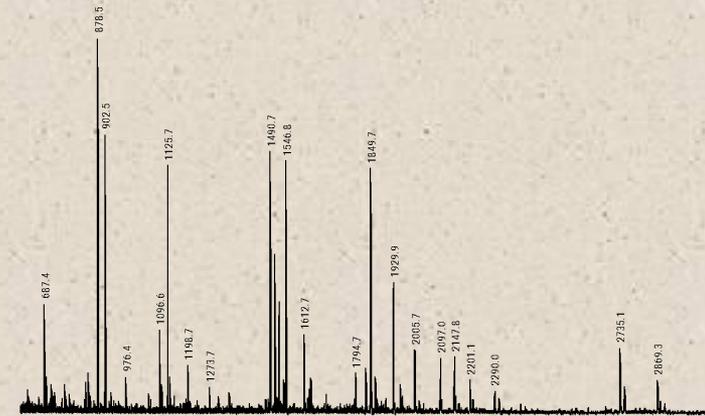
**Недостатки:** Недолговечность флуоресценции (сутки), требует дорогих сканеров, непригоден для вырезания пятен без дополнительной окраски серебром



# Схема проведения идентификации белка по его пептидному фингерпринту



Вырезание отдельных белковых пятен (>20 нг/мм<sup>2</sup>)  
Трипсинолиз белков в фрагментах геля  
Получение масс-спектра



### Mascot: Peptide Mass Fingerprint

[Your name](#)  [Email](#)

[Search title](#)

[Database](#) NCBInr **выбор базы данных**

[Taxonomy](#) ..... Homo sapiens (human) **ограничение поиска по видовой принадлежности**

[Enzyme](#) Trypsin [Allow up to](#) 1 missed cleavages

[Fixed modifications](#) Acetyl (K), Acetyl (N-term), Amide (C-term), Biotinylated (K), Biotinylated (N-term)

[Variable modifications](#) Oxidation (M), PEO Biotin (C), Phospho (ST), Phospho (Y), Propionamide (C) **учет возможных модификаций пептидов**

[Protein mass](#)  kDa [Peptide tol. ±](#) 100 ppm **точность измеренных масс**

[Mass values](#)  MH<sup>+</sup>  M<sub>r</sub> [Monoisotopic](#)  [Average](#)

[Data file](#) G:\bmcch\petr\micros\09\_01\_04\1D\_28\0 [Browse...](#)

[Query](#)  **список масс пептидных фрагментов**

[Overview](#)  [Report top](#) 20 hits **количество кандидатов**

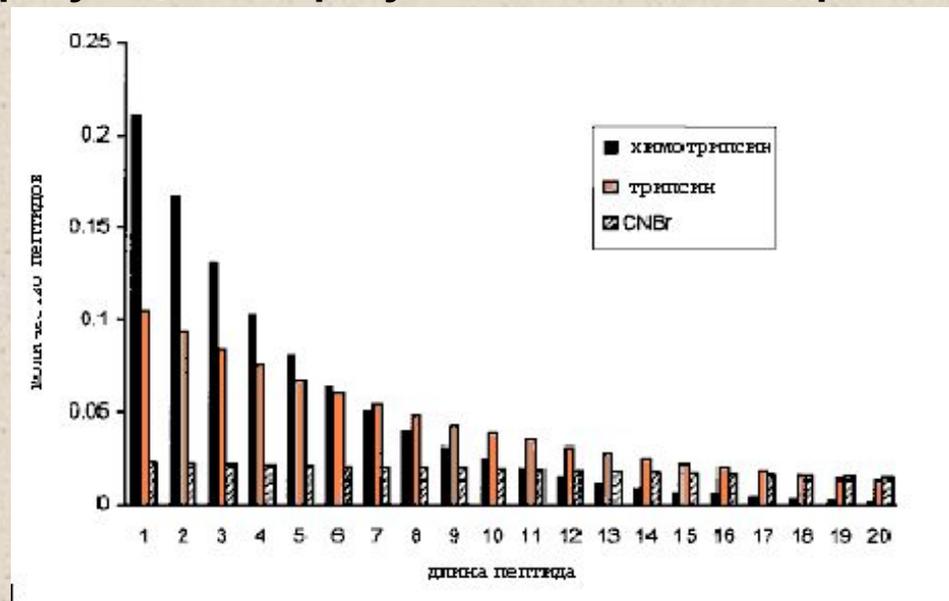
# Немного статистики:

В базе данных  
NCBI nr более 2  
500 000 записей

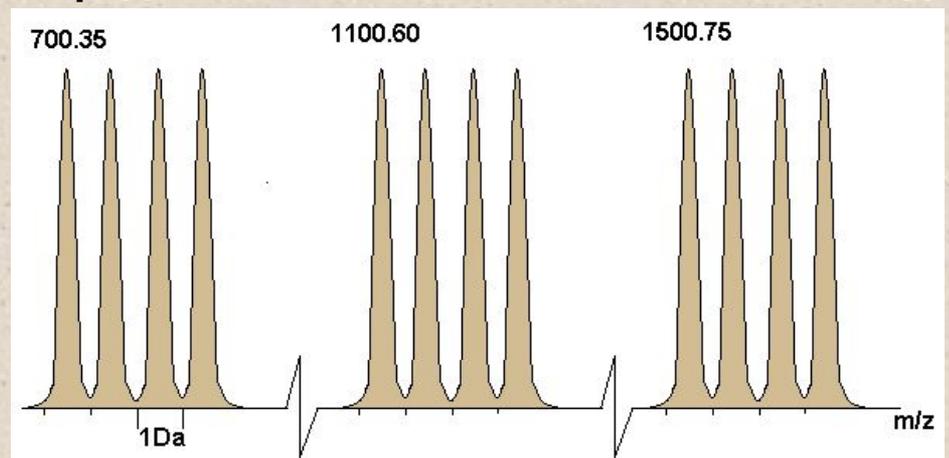
Триптические пептиды 1000-2000Da

точность (%)	кол-во кандидатов
0.1	>1000
0.01	100-1000
0.001	5-500

Среднее количество пептидов разной длины, образующихся в результате полного гидролиза

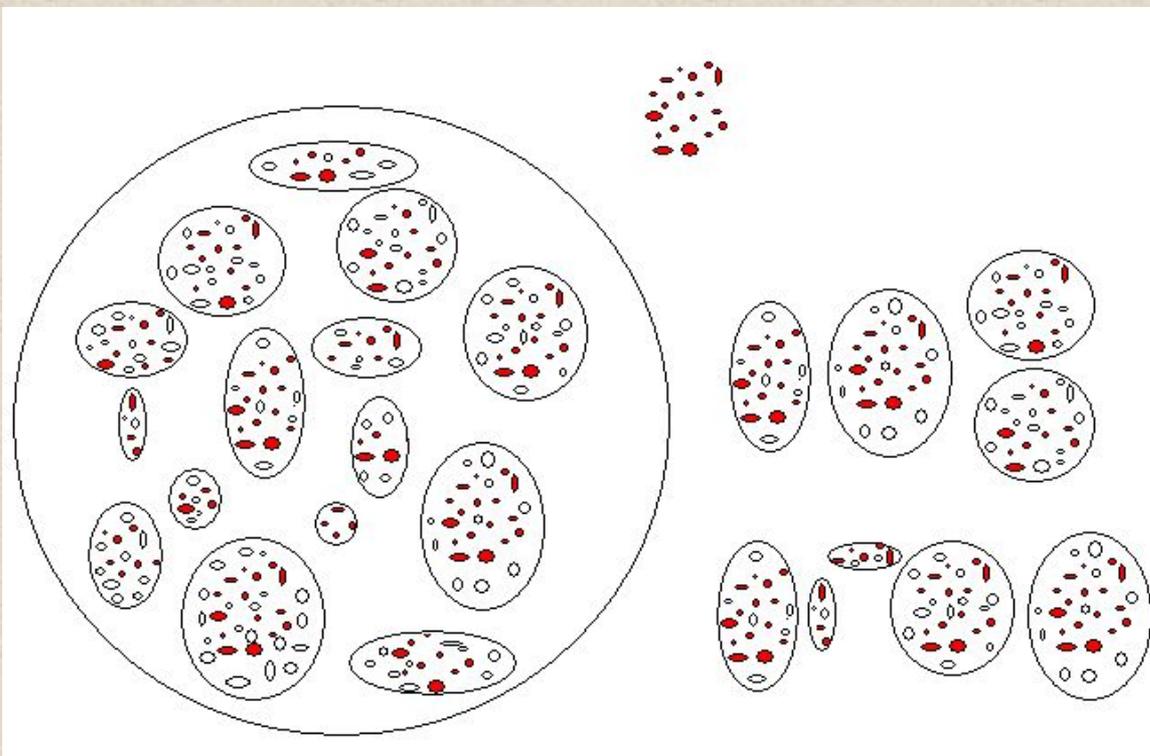


Распределение моноизотопных масс пептидов



# Критерии достоверности поиска белков в базах данных: peptide fingerprint

1. количество «совпавших» пептидов, в предположении одного белка
2. количество «совпавших» пептидов, в предположении нескольких белков
3. распределение «совпавших» пептидов по точности
4. «уникальность» пептидов
5. перекрытие «совпавшими» пептидами последовательности белка
6. паттерны протеолиза
7. отклонение массы белка от предполагаемой и т. д.

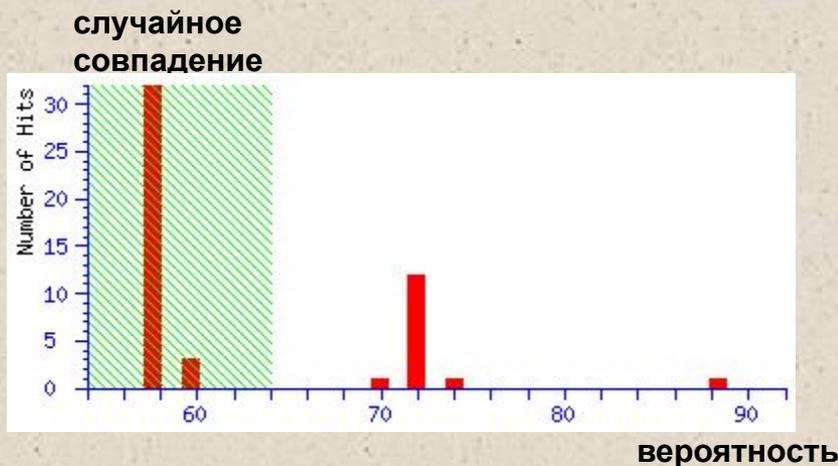


Приоритет критериев  
2, 4, 7  
программа [Profound](#)

Приоритет критериев  
1, 5, 6  
программа [Mascot](#)

# Идентификация белка заменяется нахождением его ближайшего гомолога

Исследуемый белок обычно отличен от имеющегося в базе данных



массы

20 кандидатов

Hit:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
✓ 625.35 (1+)																				
✓ 878.50 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
✓ 1068.61 (1+)																				
✓ 1078.61 (1+)																				
✓ 1125.79 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
✓ 1250.65 (1+)																				
✓ 1273.78 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
✓ 1289.77 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
✓ 1329.84 (1+)																				
✓ 1422.76 (1+)																				
✓ 1490.82 (1+)	●	●	●	●	●	●	●	●							●		●	●		
✓ 1506.83 (1+)	●	●	●	●	●	●	●	●							●		●	●		
✓ 1546.86 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
✓ 1635.96 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●					●
✓ 1823.95 (1+)																				
✓ 1878.88 (1+)															●					
✓ 2383.02 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
✓ 2399.02 (1+)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
✓ 2478.31 (1+)																				
✓ 2685.47 (1+)	●									●	●	●	●							●
✓ 2841.56 (1+)																				
✓ 3110.53 (1+)																				

20 кандидатов в порядке убывания достоверности поиска

	N	Масса	Вероятность	Описание
1.	gi 8486123	27875	88	reading frame [Influenza A virus]
2.	gi 4996872	26451	74	M1 [influenza A virus (A/Taiwan/96/1769)]
3.	gi 324260	27872	73	M1 subunit [Influenza A virus]
4.	gi 75116	27866	73	M1 - influenza A virus (strain A/WSN/33)
-----				
20.	gi 6048806	27218	58	M1 [Influenza A virus (A/Duck/Hong Kong/698/79 (H5N3))]

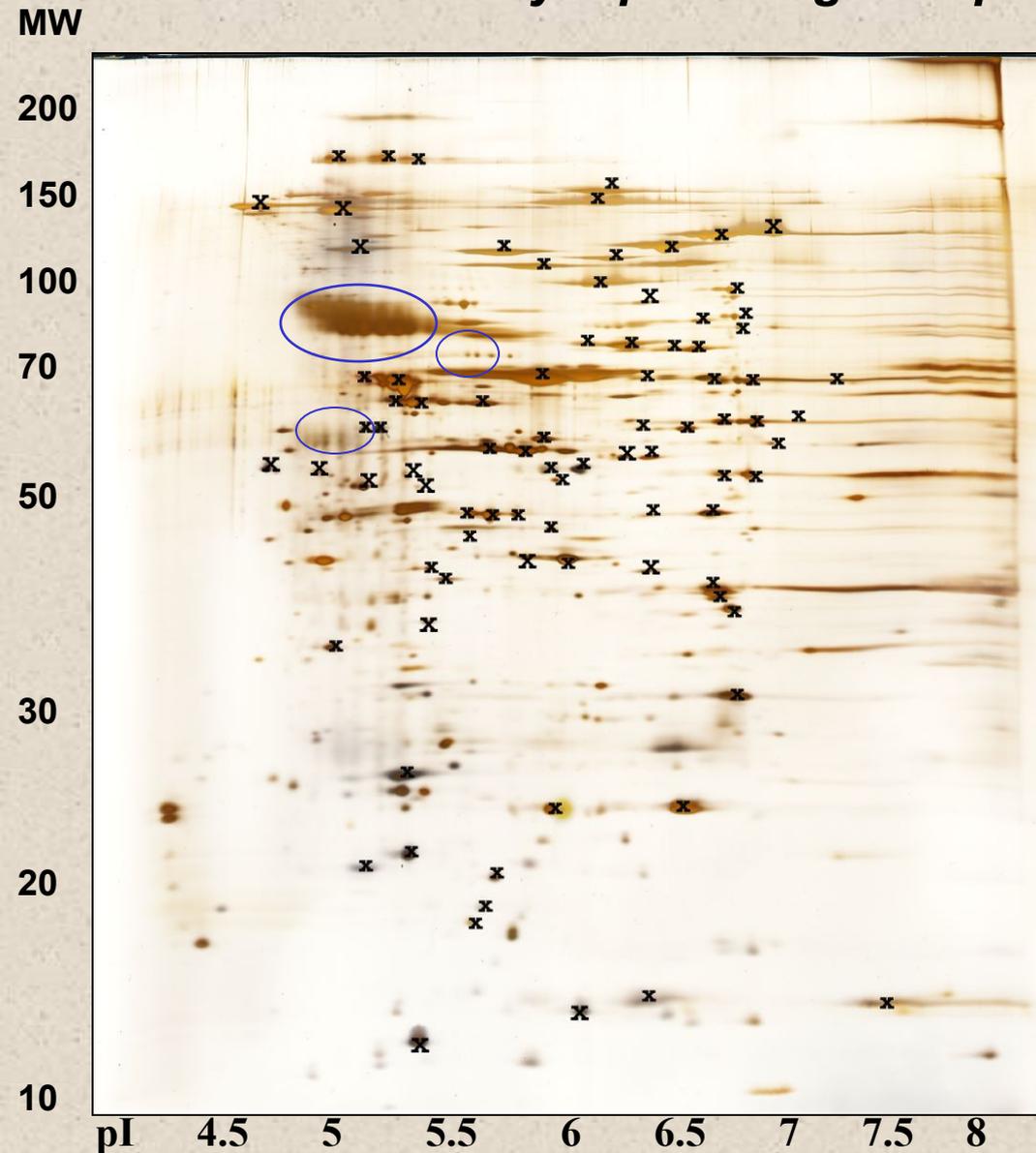
Sequence Coverage: 35%

IVPSGPLKAE IAQRLEDVFA GKNTDLEVLV EWLKTRPILS PLTKGILGFV  
 FTLTVPSEK LQRRRFVQNA LNGNGDPNNM DKAVKLYRKL KREITFHGAK  
 EIALSYSAGA LASCMLIYN RMGTVTTEVA FGLVCATCEQ IADSQHRSHR  
 QMVTITNPLI RHENRMVLAS TTAKAMEQMA GSSEQAAEAM DIASQARQMV  
 QAMRTIGTHH SSSAGLKDDL LENLQAYQKR MGVQMQRFK



# 2DE + MALDI-MS - результаты

## *Mycoplasma gallisepticum* strain R



Protein coding genes - 726

Белковых пятен > 400

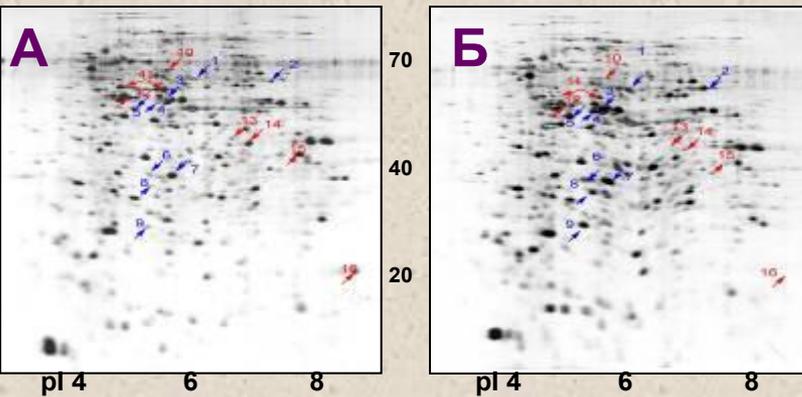
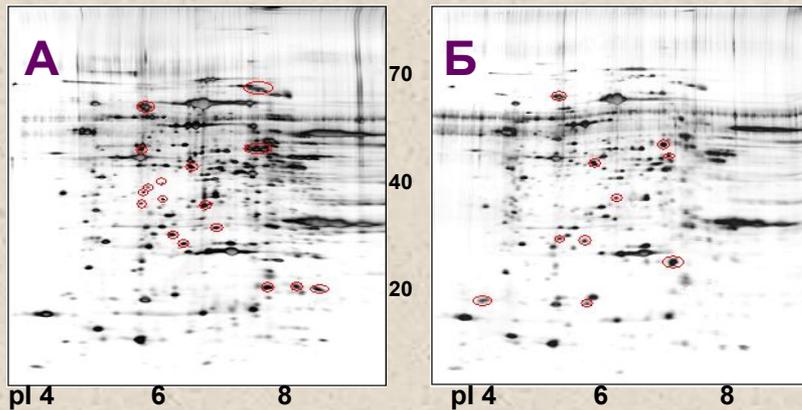
Проанализировано >150

Индивидуальных белков  
*Mycoplasma gallisepticum* 105

# Примеры анализа 2DE карт

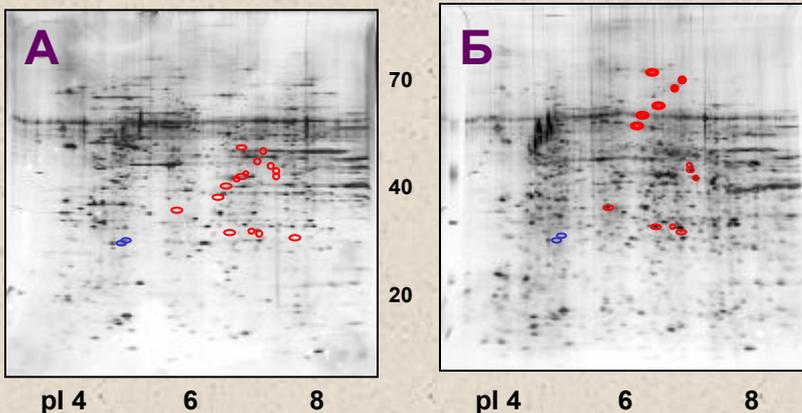
## 2DE карты клеточных белков двух штаммов *Helicobacter pylori*

Всего видно на геле ~ 500 белковых пятен  
Видимых различий ~ 200  
Идентифицировано ~ 150 разных белков  
Из них белки отличия ~ 100



## 2DE карты белков штамма клеток *E.coli* (А) и штамма-продуцента треонина (Б)

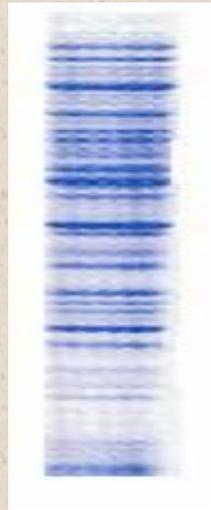
Всего видно на геле ~1200 белковых пятен  
Видимых различий ~ 50  
Идентифицировано ~ 30 белков отличия



## 2DE карты белков клеточной линии аденокарциномы молочной железы MCF7: контрольной (А) и устойчивой к Ноеchst33342 (Б)

Всего видно на геле ~ 2500 белковых пятен  
Видимых различий ~ 30  
Идентифицировано ~ 10 белков отличия

# Стратегия анализа: 1D - LC + ESI-MS/MS



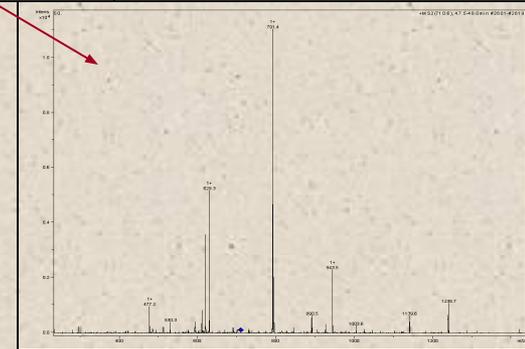
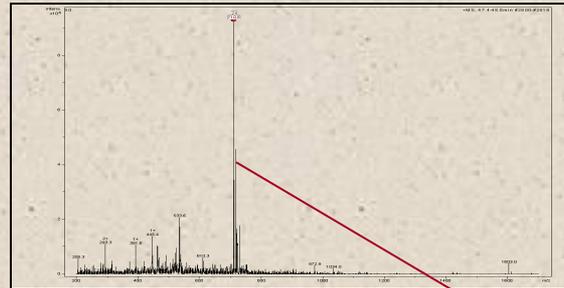
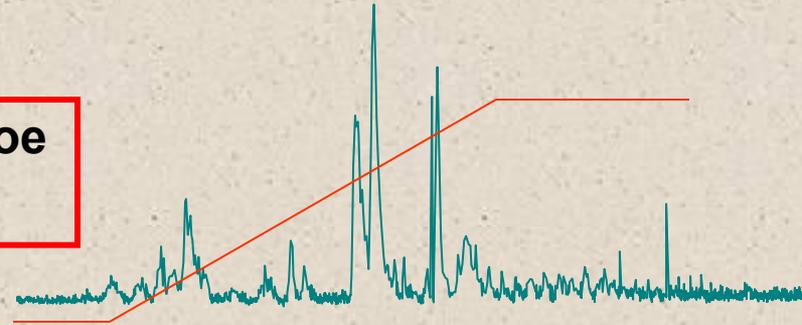
1D электрофорез белков

Вырезание и трипсинолиз белков в полосах геля

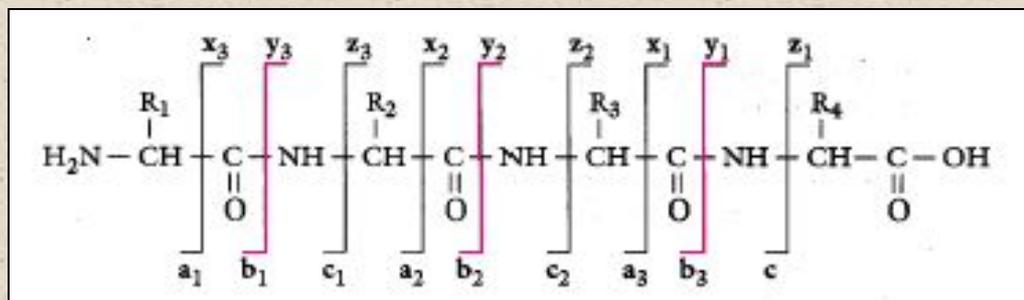
Экстракция и хроматографическое  
разделение пептидов

Получение масс-спектров  
пептидов и их  
фрагментации

Объединение данных  
и анализ



# Способы фрагментации пептидов



## LID – фрагментация, индуцированная лазером

Пептид поглощает энергию лазера, захваченная энергия перераспределяется по молекуле, разрушается ближайшая слабая связь.

Основной тип образовавшихся пептидных фрагментов - b-ионы и y-ионы.

Может происходить отщепление модификаций.

## CID – фрагментация, индуцированная столкновениями

Пептид приобретает энергию от столкновений, захваченная энергия перераспределяется по молекуле, разрушается ближайшая слабая связь.

Основной тип образовавшихся пептидных фрагментов - b-ионы и y-ионы.

Дополнительно образуются a-ионы и x-ионы.

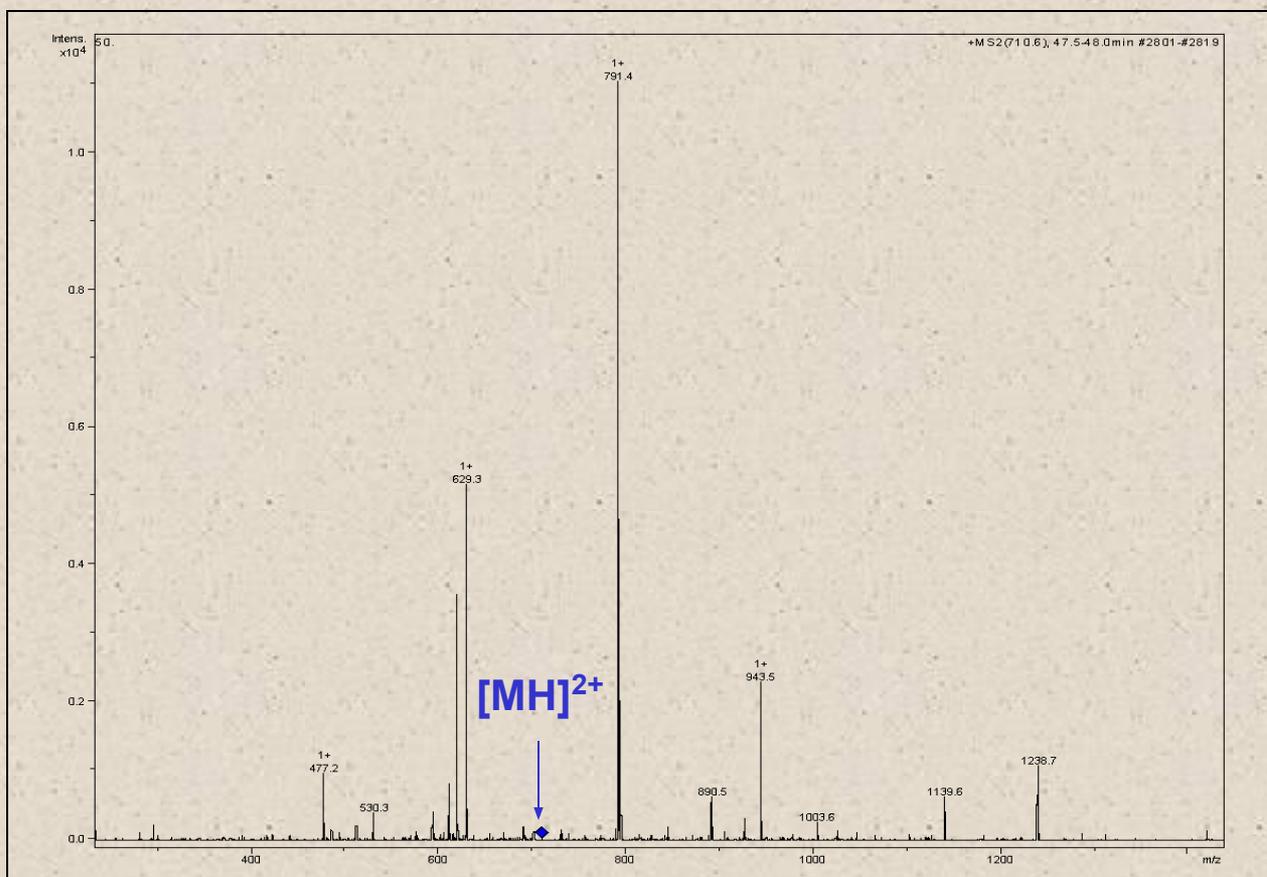
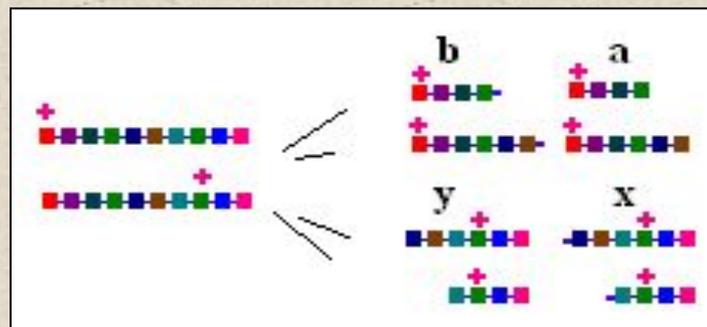
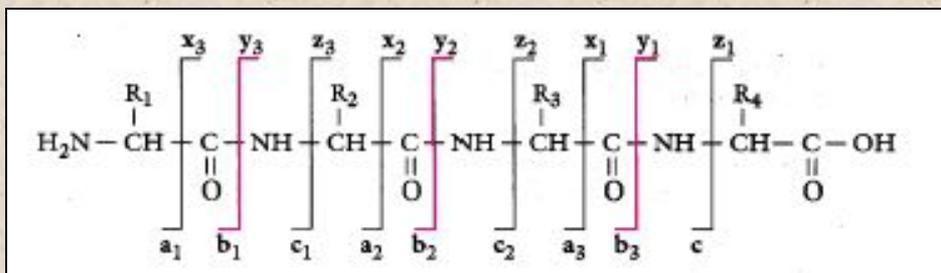
Часто происходит отщепление модификаций.

## ECD – фрагментация, индуцированная захватом медленных электронов

Точный механизм неизвестен. Пептид захватывает медленные электроны на азот пептидной связи, происходит мгновенный разрыв пептидной цепи с образованием z-ионов.

Не происходит отщепления модификаций.

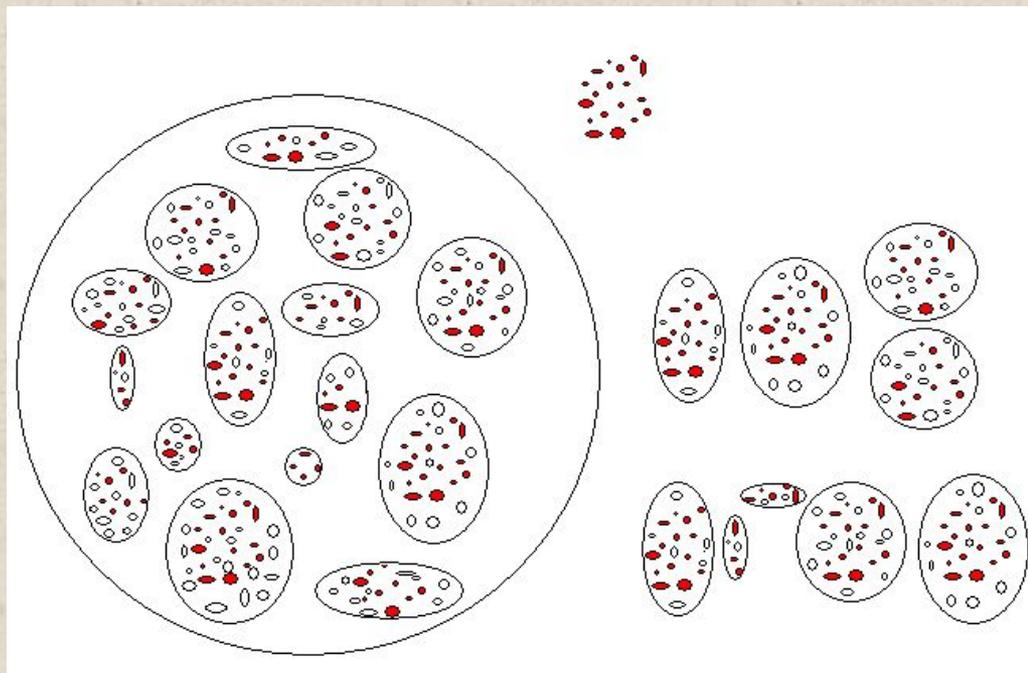
# Интерпретация спектров распада пептидов



**TIGTHPSSSAGLK**

b		Res.		y
102.0	1	T	13	
215.1	2	I	12	1154.6
272.2	3	G	11	1041.5
373.2	4	T	10	984.5
510.3	5	H	9	883.5
607.3	6	P	8	746.4
694.4	7	S	7	649.4
781.4	8	S	6	562.3
868.4	9	S	5	475.3
939.5	10	A	4	388.3
996.5	11	G	3	317.2
1109.6	12	L	2	260.2
1237.7	13	K	1	147.1

# Критерии достоверности поиска белков в базах данных: peptide fingerprint + MS/MS



1. Отбор из базы данных всех кандидатов, содержащих триптические пептиды указанных  $m/z$

2. «Примерка» спектров фрагментации (MS/MS) :  
количество «совпавших» фрагментов пептидов  
«уникальность» спектров фрагментации пептидов

Score пептидов

3. Score белка =  $\sum$  Score пептидов

# 1DE + LC-ESI-MS/MS - результаты

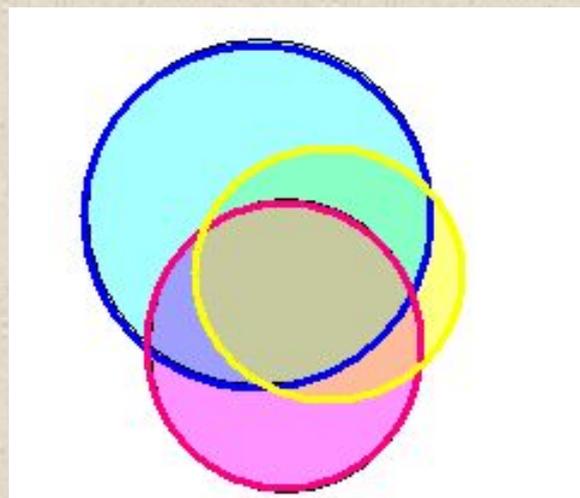
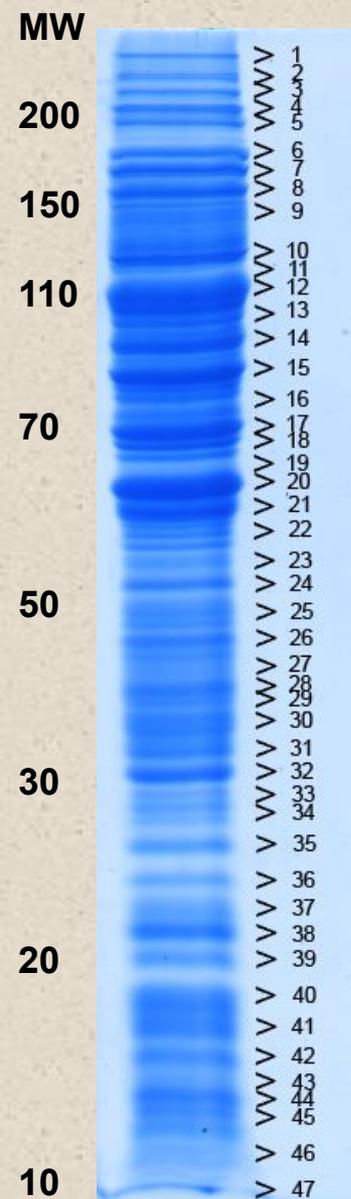
*Mycoplasma gallisepticum* strain R

Проведено 3 эксперимента

Белковых полос - 50

Индивидуальных белков

*Mycoplasma gallisepticum* - 427



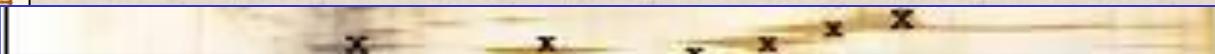
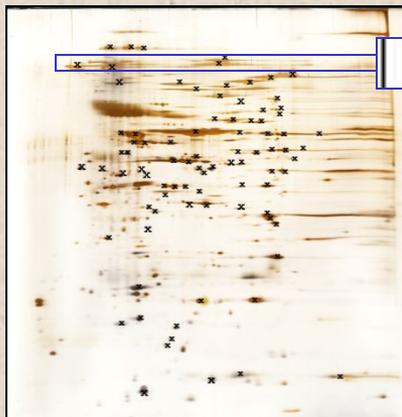
Белки, которые обнаружены хотя бы в 1 эксперименте - 427

Белки, которые обнаружены хотя бы в 2 экспериментах - 285

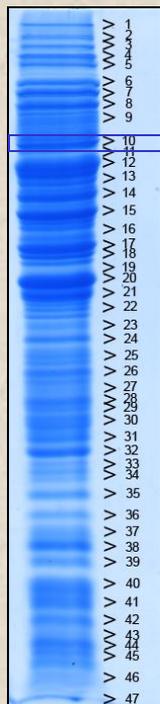
Белки, которые обнаружены во всех 3 экспериментах - 155

Белки, которые обнаружены при проведении 2DE+MALDI-MS - 90

# СОПОСТАВЛЕНИЕ РЕЗУЛЬТАТОВ: 2DE + MALDI-MS и 1DE + LC-ESI-MS/MS



Protein	№ NCBI	mW	score
<u>conserved hypothetical lipoprotein</u>	<u><a href="#">gi 31541381</a></u>	85818	<b>119</b>
conserved hypothetical lipoprotein	<a href="#">gi 31541566</a>	100472	<b>130</b>
conserved hypothetical lipoprotein	<a href="#">gi 31541571</a>	120325	<b>273</b>
<u>conserved hypothetical lipoprotein</u>	<u><a href="#">gi 31541607</a></u>	130197	<b>205</b>
conserved hypothetical lipoprotein	<a href="#">gi 31541608</a>	133826	<b>168</b>



Protein	№ NCBI	mW	score
<u>conserved hypothetical lipoprotein</u>	<u><a href="#">gi 31541381</a></u>	85818	<b>127</b>
conserved hypothetical lipoprotein	<a href="#">gi 31541566</a>	100472	<b>817</b>
HepA/SNF	<a href="#">gi 31541753</a>	132618	<b>341</b>
conserved hypothetical lipoprotein	<a href="#">gi 31541571</a>	120325	<b>2152</b>
<u>conserved hypothetical lipoprotein</u>	<u><a href="#">gi 31541607</a></u>	130197	<b>1781</b>
lipoprotein	<a href="#">gi 45453612</a>	18284	<b>313</b>
lipoprotein	<a href="#">gi 45453638</a>	18339	<b>362</b>
lipoprotein	<a href="#">gi 33327196</a>	25889	<b>436</b>
conserved hypothetical lipoprotein	<a href="#">gi 31541608</a>	133826	<b>1008</b>

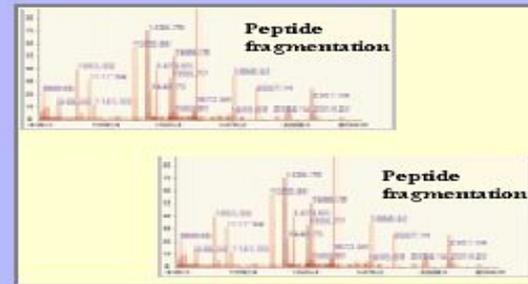
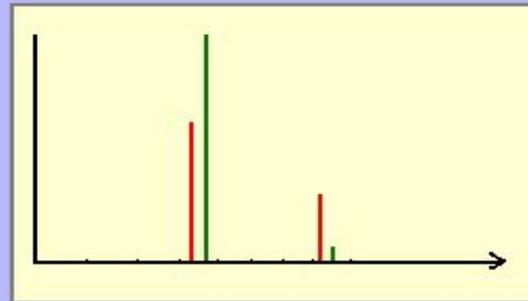
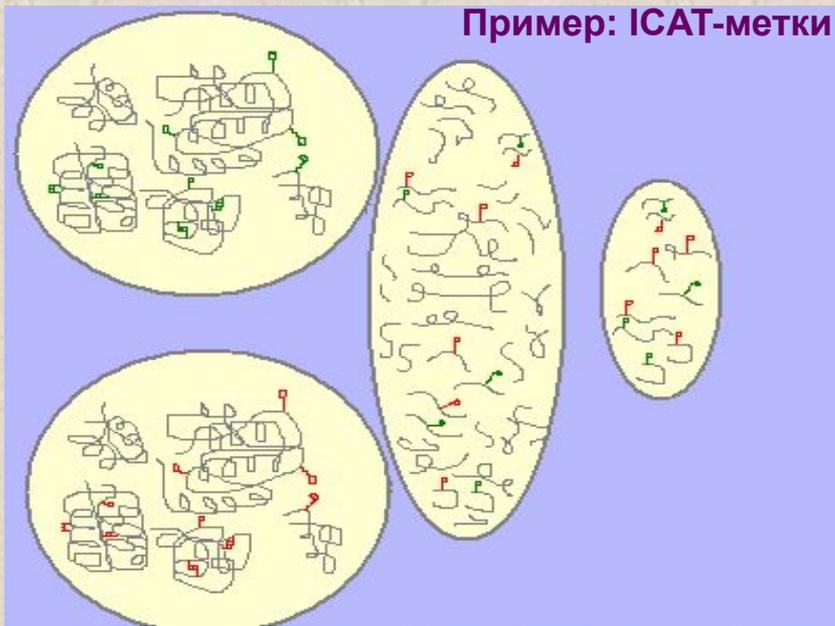
Специфический гидролиз  
суммарного белка + мечение ICAT

# Стратегия анализа: LC + ESI-MS/MS

Выделение цистеинсодержащих пептидов  
(аффинное связывание с авидином)

ICAT **Cys + acetamide-CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>-biotin** контроль  
**Cys + acetamide-CD<sub>2</sub>-CD<sub>2</sub>-CD<sub>2</sub>-biotin** опыт

Хроматографическое  
разделение пептидов + MS/MS



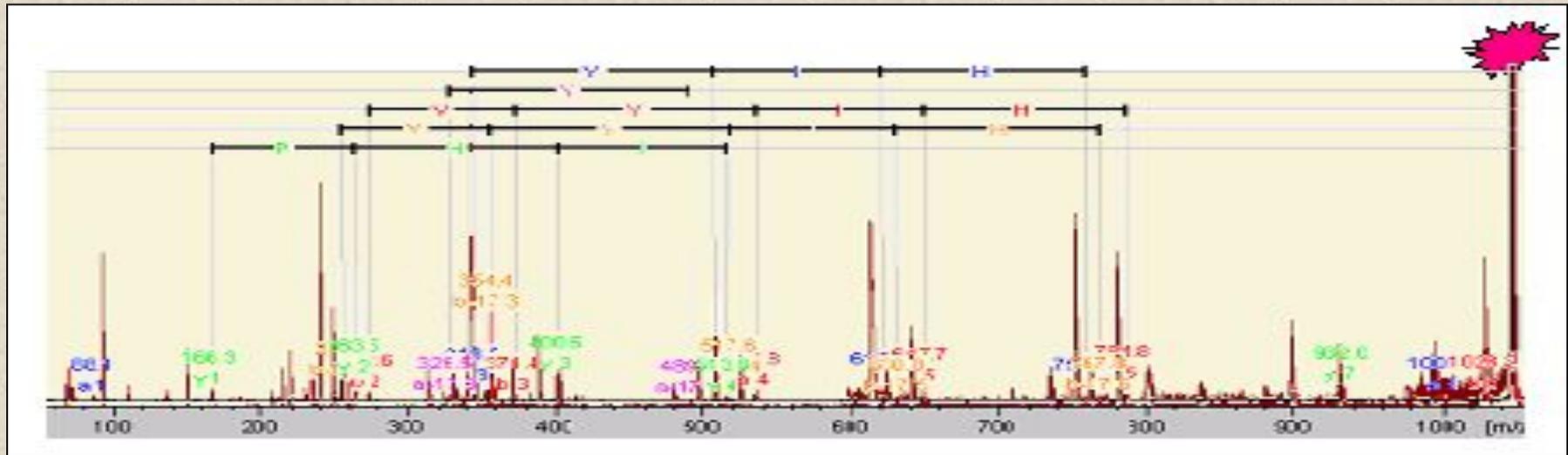
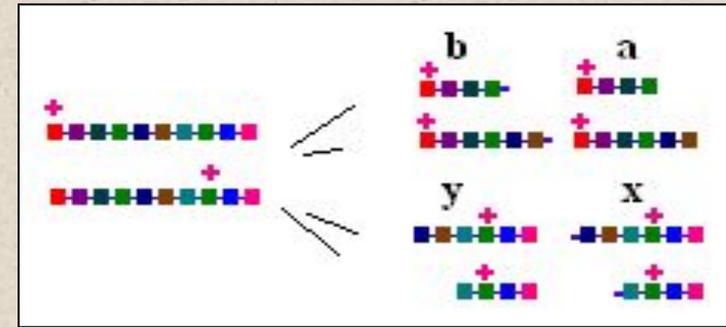
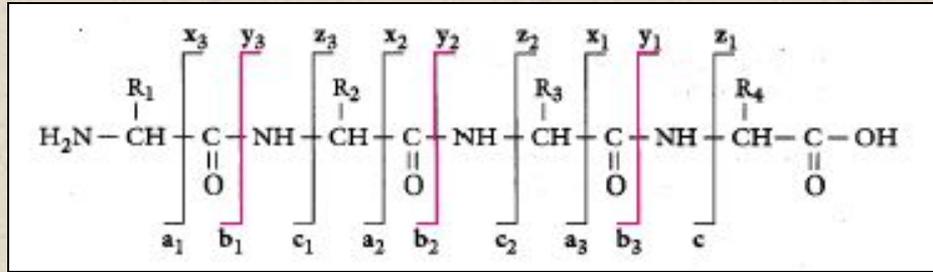
## Достоинства:

Анализируются все меченные белки  
Анализ содержания одного белка в  
контроле и опыте

## Ограничения:

Не видно взаимных количеств  
разных белков в образце  
Артефакты в интерпретации MS/MS

# Интерпретация спектров распада пептидов (LID +CID)



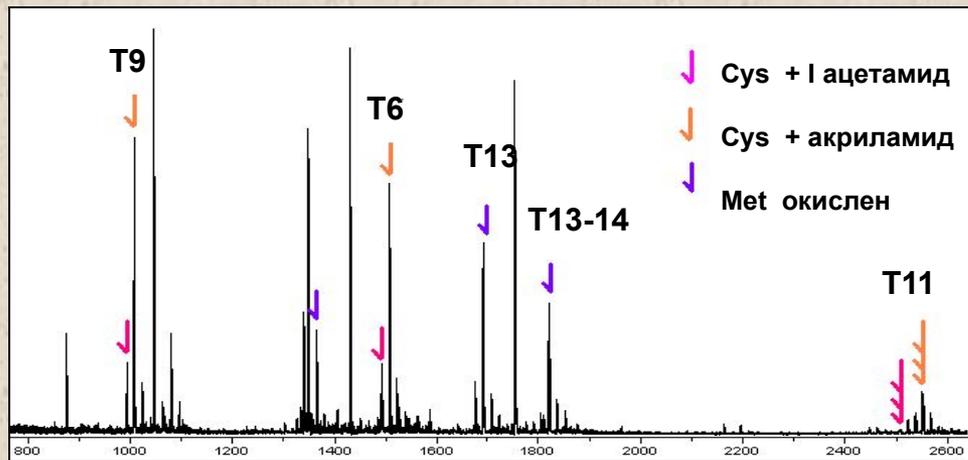
## Сиквенирование *de-novo*:

- определение пар пиков, в сумме дающих массу родительского иона
- по наличию соседних пиков  $\pm 28$ ,  $\pm 15$  определение типа ионов
- по разнице масс соседних однотипных ионов построение возможных последовательностей



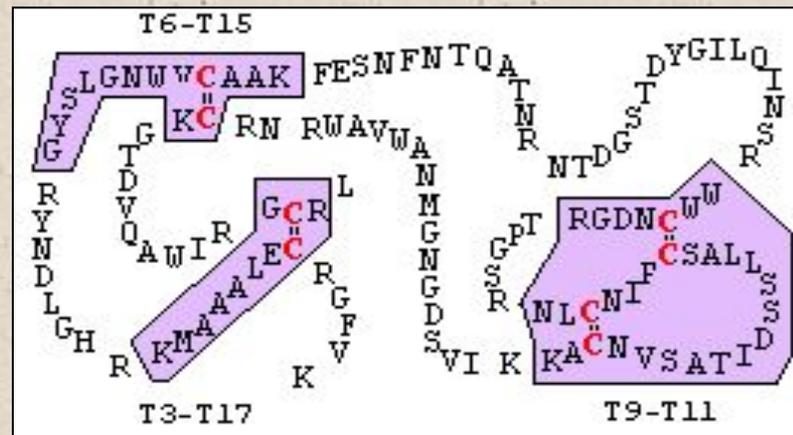
# Определение пост-трансляционных модификаций белков

Модификации *in vitro*: Модификации белков после их разделения в ПААГ

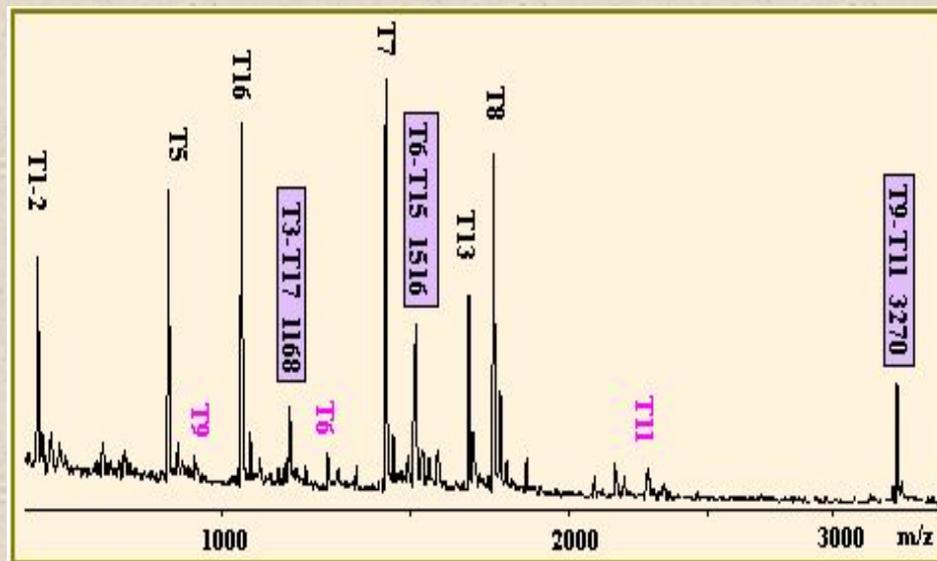


цистеины алкилированы йодоацетамидом (в геле)

Лизоцим



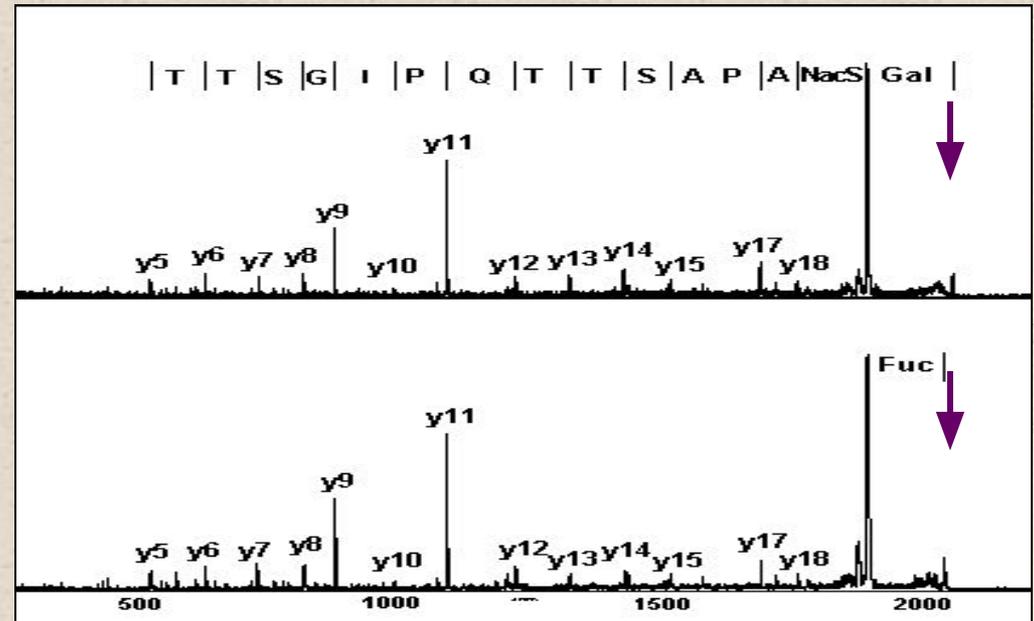
Модификации *in vivo*: Изучение S-S мостов



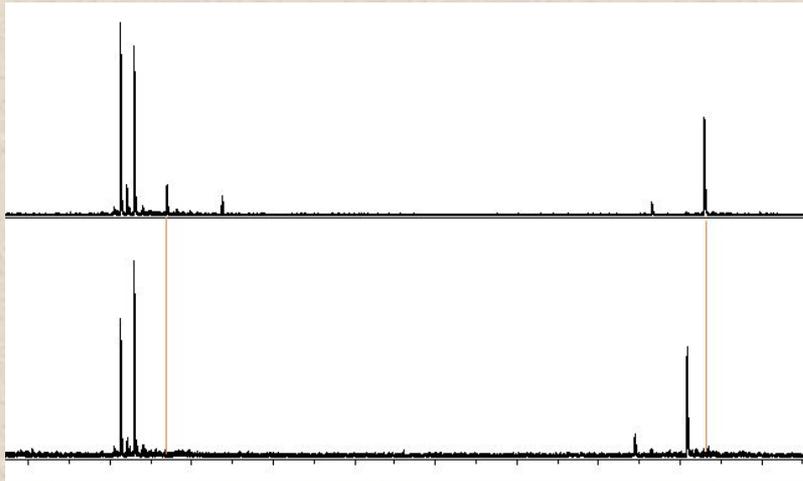
# Стабильность различных модификаций белков в условиях получения масс-спектров (LID + CID)

Связи менее устойчивые, чем пептидная:  
фосфоэфирная  
гликозидная  
дисульфидная (иногда)

Сахарный остаток локализован на N-концевом серине ?



Такова степень фосфорилирования ?

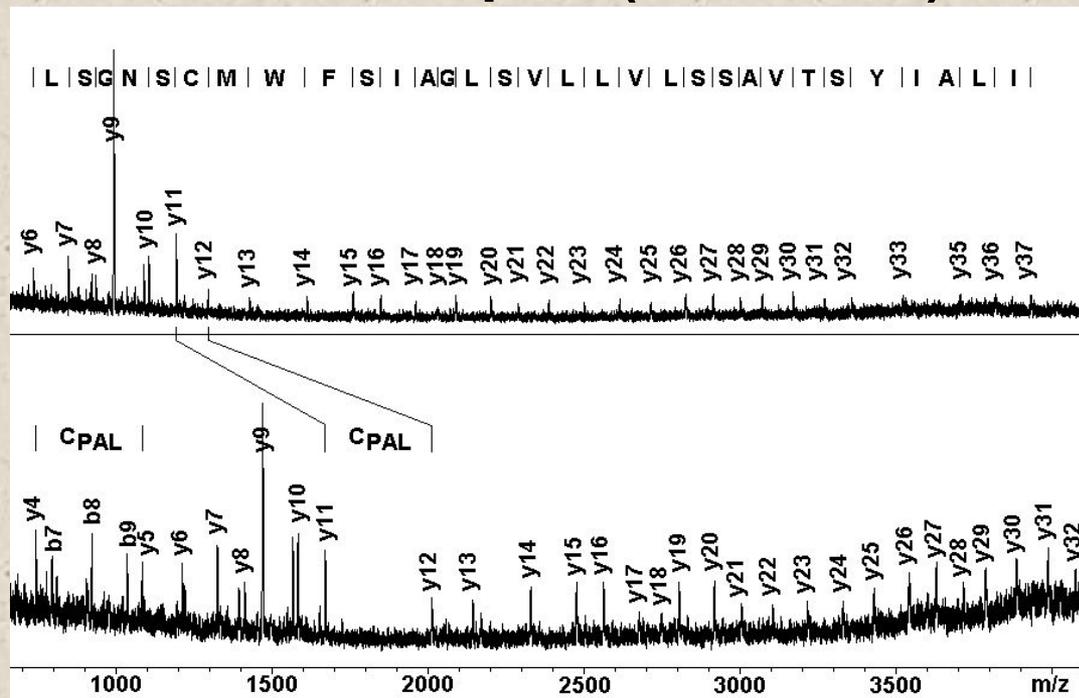


О-гликозилированный пептид SAPASTTQPIG**STTSTTTK**  
сахарный остаток галактоза (верх) либо фукоза (низ)

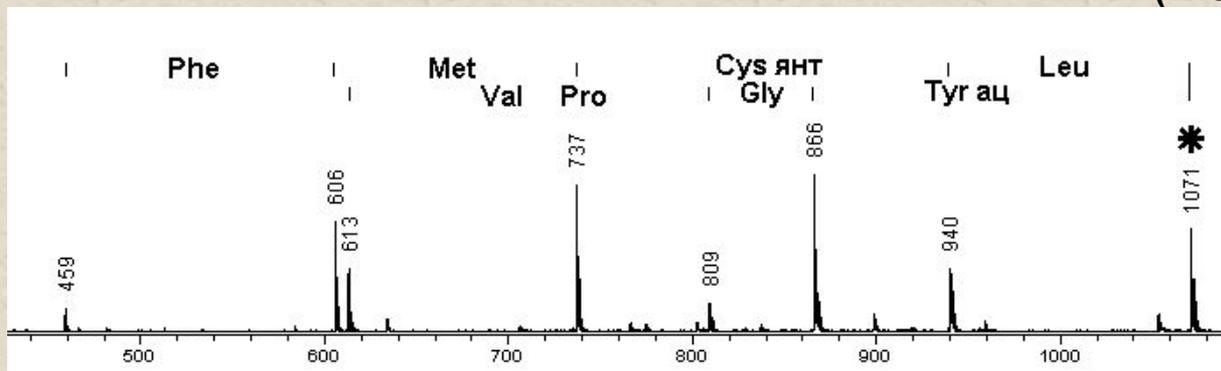
Фрагмент спектра трипсинолиза природного (верх)  
и генно-инженерного (низ) белка

# Стабильность различных модификаций белков в условиях получения масс-спектров (LID + CID)

Связи сопоставимые по устойчивости с пептидной:  
 тиоэфирная  
 любая амидная -  
 N-ацелирование  
 дисульфидная (иногда)



Пептид, модифицированный тремя пальмитатами (тиоэфирная связь)



Пептид N- ацелирован, а цистеин модифицирован янтарным ангидридом (тиоэфирная связь)

# Сравнение подходов

при определении белкового состава сложной многокомпонентной смеси

	MS целых белков	LC-ESI-MS/MS	2DE -MALDI-TOF-MS	1DE-MALDI-TOF-MS
Определение количества одного белка в разных образцах	++	++ - +++++	+++	+
Определение количеств разных белков в одном образце	++	+	+++	+
Идентификация сильно представленных в образце белков		++++	+++++	+++
Идентификация мало представленных в образце белков		+	+++	+
Идентификация мембранных белков		+++		+++
Определение возможных модификаций белка		+	++	

## Стратегия определения модификаций:

Выделение белка

Гидролиз специфическими протеазами -

соотнесение значений масс пиков в спектре с первичной структурой.

Определение пептидов отличных по массе от рассчитанной –

гипотезы о природе модификации

Индивидуальный подход к получению структурной информации из спектров фрагментации –

выбор типа фрагментации, использование ферментов для дальнейшего расщепления пептида, удаления модификации, химические подходы...

# Литература:

- \* Говорун В.М., Арчаков А.И. *Протеомные технологии в современной биомедицинской науке.* Biochemistry (Mosc). 2002 Oct;67(10):1109-23.
- \* Lottspeich F. *Proteome Analysis: A Pathway to the Functional Analysis of Proteins.* Angew Chem Int Ed Engl. 1999 Sep;38(17):2476-2492
- \* M.Mann, R.C.Hendrickson, A.Pandey, *Analysis of proteins and proteomes by mass spectrometry,* Annu. Rev. Biochem., 70, 437-473 (2001).
- \* O.N.Jensen, M.Wilm, A.Shevchenko, M.Mann, *Sample preparation methods for mass spectrometric peptide mapping directly from 2-DE gels,* Methods Mol. Biol., 112, 513-530 (1999).
- \* B.Kuster, T.N.Krogh, E.Mortz, D.J.Harvey, *Glycosylation analysis of gel-separated proteins,* Proteomics, 1, 350-361 (2001).
- \* Aebersold R, Goodlett DR *Mass spectrometry in proteomics* Chem Rev 2001 Feb;101(2):269-95
- \* Nordhoff E, Egelhofer V, Giavalisco P, Eickhoff H, Horn M, Przewieslik T, Theiss D, Schneider U, Lehrach H, Gobom J. *Large-gel two-dimensional electrophoresis-matrix assisted laser desorption/ionization-time flight-mass spectrometry: an analytical challenge for studying complex protein mixtures.* Electrophoresis. 2001 Aug;22(14):2844-55
- \* Roepstorff P *MALDI-TOF mass spectrometry in protein chemistry..* EXS 2000;88:81-97
- \* Gygi SP, Aebersold R. *Using mass-spectrometry for quantitative proteomics* Proteomics: A Trands Guide 2000, 31-36
- \* Andersen JS, Mann M *Functional genomics by mass spectrometry.* FEBS Lett. 2000 Aug 25;480(1):25-31
- \* Kussmann M, Roepstorff P *Sample preparation techniques for peptides and proteins analyzed by MALDI-MS.* Methods Mol Biol. 2000;146:405-24
- \* Chaurand P, Luetzenkirchen F, Spengler B. *Peptide and protein identification by matrix-assisted laser desorption ionization (MALDI) and MALDI-post-source decay time-of-flight mass spectrometry.* J Am Soc Mass Spectrom. 1999 Feb;10(2):91-103
- \* Spengler, B, *Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules.* J. Mass Spectrom., 32(10) 1019-36 (1997).
- \* Jonscher, KR and Yates, JR, 3rd, *The quadrupole ion trap mass spectrometer--a small solution to a big challenge.* Anal Biochem, 244(1) 1-15 (1997).
- \* Qin, J and Chait, BT, *Identification and characterization of posttranslational modifications of proteins by MALDI ion trap mass spectrometry.* Anal Chem, 69(19) 4002-9 (1997).
- \* Papayannopoulos, IA, *The interpretation of collision-induced dissociation tandem mass spectra of peptides.* Mass Spectrom. Rev., 14(1) 49-73 (1995).
- \* Hillenkamp, F, Karas, M, Beavis, RC and Chait, BT, *Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers.* Anal. Chem., 63(24) 1193A-203A (1991).
- \* Fenn, JB, Mann, M, Meng, CK, Wong, SF and Whitehouse, CM, *Electrospray ionization-principles and practice.* Mass Spectrom. Rev., 9(1) 37-70 (1990)