ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ УЧРЕЖДЕНИЕ «НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ БИОМЕДИЦИНСКОЙ ХИМИИ ИМЕНИ В.Н. ОРЕХОВИЧА»

На правах рукописи

КУЗНЕЦОВА КСЕНИЯ ГЛЕБОВНА

ВЛИЯНИЕ ХИМИЧЕСКОЙ ОБРАБОТКИ ОСТАТКОВ ЦИСТЕИНА В БЕЛКАХ НА РЕЗУЛЬТАТЫ ПРОТЕОГЕНОМНОГО АНАЛИЗА

03.01.04 - биохимия

Диссертация на соискание ученой степени кандидата биологических наук

Научный руководитель: д.б.н., проф. Мошковский Сергей Александрович

Москва 2020

СПИСОК СОКРАЩЕНИЙ	4
ВВЕДЕНИЕ	6
ГЛАВА №1. ОБЗОР ЛИТЕРАТУРЫ	. 10
1.1. Историческое развитие методов протеомного анализа	. 10
1.2. Основные методы протеомного анализа	12
1.3. Биоинформатический анализ протеомных данных	. 14
1.3.1. Протеомный поиск	. 14
1.3.2. Протеогеномный подход	. 14
1.3.3. Расчет уровня ложноположительных результатов	. 15
1.4. Основные этапы пробоподготовки в протеомном анализе	. 16
1.4.1. Лизис образца и выделение смеси белков	. 16
1.4.2. Обработка дисульфидных связей остатков цистеина для протеомного анализа	. 19
1.5. Алкилирование остатков цистеина: побочные реакции	23
1.6. Нежелательные модификации алкилирующими агентами остатков метионина	24
1.7. Выбор алкилирующих агентов при исследовании посттрансляционных модификаций	27
1.8. Обработка остатков цистеина в протеогеномных исследованиях	
1.9. Обработка остатков цистеина в редокс-протеомике	
1.10. Современные решения для поиска модификаций аминокислотных остатков.	
«Открытый» поиск	. 29
1.11. Определение изомеров аминокислотных остатков в пептидах при помощи масс-	
спектрометрии	. 33
ГЛАВА №2 МАТЕРИАЛЫ И МЕТОЛЫ	35
21 Kultupa kultu	35
2.1. Культура клоток	. 35
2.2. Подготовка образцов к масе спектрометри некому апализу	35
2.2.1. Подготовка образцов сы пето сыворото пого альсумина	36
2.2.2. Подготовка образцов культур клеток	36
2.2.5. Сиптер пентидор	37
2.3.1 Хроматомасс-спектрометрия образцов бычьего сывороточного альбумина	37
2.3.2. Хроматомасс-спектрометрия образцов сы пете сысерете него шесумний ините	38
2.3.2. Проматомае сполгромотрии сораздов культур настоя	39
2.4. Анализ ланных	. 40
2.4.1. Протеогеномный поиск при помощи программы X! Tandem	. 40
2.4.2. Количественный анализ без использования метки при помоши программы	
MaxOuant	41
2.4.3. Протеомный поиск при помощи программы IdentiPy	
2.4.4. Анализ частоты преобразования Мет в изоТре в панорамных протеомных данных	x 42
2.4.5. Оценка влияния соседствующих аминокислотных остатков на частоту	
преобразования Мет в изоТре	43
2.4.6. Открытый поиск и анализ с использованием программы AA_stat	43
2.4.7. Анализ масс-спектров синтетических пептидов	44
2.4.8. Машинное обучение	45
ΓΠΑΡΑ Μ.2 ΒΕΡΥΠΙΤΑΤΙΙ Η ΟΓΟΥΜΠΕΙΗΗΕ	16
1 ЛАДА №3. ГЕЗУЛІГАТЫ И ОВСУЖДЕНИЕ	- 40 7
5.1. замена метионина в от позиции в ослке 1150 / о произошла в 6 из 9 клеточных линиях понеци NCI60	16
наполи тустой. 3.2. Замена метионина в положении 61 белиз НССТО на заколирована в ганома.	. 4 0 //7
3.3. Гипотеза о преобразорации метионица в изотреоции в процессе проболототории	.+/ /7
3.4 Анализ масс-спектров синтетицеских стандартов пелтила $57(NOV \Delta TNPTNTVFD \Lambda K)$)71
и его варианта с изотреонином	.48
· · · · · · · · · · · · · · · · · · ·	

3.5. Количественный анализ пептида 57(NQVAMNPTNTVFDAK)71 и его варианта с	
изотреонином в панорамных протеомных данных клеточных линий панели NCI60	50
3.6. Моделирование протеомной пробоподготовки на бычьем сывороточном альбумине дл предсказания преобразования Мет в изоТре	ля 50
3.7. Анализ масс-спектров синтетических пептидов, содержащих Тре и изоТре	53
3.7.1. Синтетические аналоги с треонином и изотреонином различаются по времени выход	дa
с хроматографической колонки	53
3.7.2. Анализ спектров фрагментации пептидов с треонином и изотреонином	56
3.8. Сравнение методов алкилирования остатков цистеина при пробоподготовке к	
протеомному анализу	60
3.8. Анализ частоты преобразования метионина в изотреонин в протеомных данных	66
3.9. «Открытый» поиск и статистический анализ сдвигов молекулярной массы	71
3.10. Окисление и карбамидометилирование метионина	78
ЗАКЛЮЧЕНИЕ	78
выводы	.79
ФИНАНСИРОВАНИЕ	80
БЛАГОДАРНОСТИ	80
ЛИТЕРАТУРА	81
ПРИЛОЖЕНИЕ	92

СПИСОК СОКРАЩЕНИЙ

- а.о. аминокислотный остаток
- ВВ время выхода с хроматографической колонки
- ВЖЭХ высокоэффективная жидкостная хроматография

МС – масс-спектрометрия

МС/МС – масс-спектр второго порядка (масс-спектр фрагментации)

MC³ – масс-спектр третьего порядка

- Трис-HCl буферный раствор трис(гидроксиметил)аминометана и соляной кислоты
- 4-VP (англ. 4-vinylpyridine) 4-винилпиридин
- АА (англ. acrylamide) акриламид
- ACN (англ. acetonitrile) ацетонитрил

```
ADAR – (англ. double-stranded RNA-specific adenosine deaminase) РНК-зависимая
```

аденозиндезаминаза

AGC – (англ. automatic gain control) автоматическая регулировка усиления

BCA – (англ. bicinchoninic acid) бицинхониновая кислота

BSA – (англ. bovine serum albumin) бычий сывороточный альбумин

САМ – (англ. chloroacetamide) хлорацетамид

СНАРЅ – (англ. 3-[(3-cholamidopropyl)dimethyl-ammonio]-1-propane sulfonate) 3-[(3-

холамидопропил)диметиламмонио]-1-пропана сульфонат

CID – (англ. collision-induced dissociation) диссоциация, индуцированная соударением

DDA – (англ. data dependent acquisition) зависимый от данных режим

DMF – (англ. dimethylformamide) диметилформамид

DTT – (англ. dithiothreitol) дитиотреитол

ECD – (англ. electron capture dissociation) диссоциация при захвате электрона

ETD – (англ. electron transfer dissociation) диссоциация при переносе электрона

FA – (англ. formic acid) муравьиная кислота

FAB – (англ. fast atom bombardment) бомбардировки быстрыми атомами

FASP (англ. *filter-aided sample preparation*) пробоподготовка с использованием центрифужных фильтров

FBS – (англ. fetal bovine serum) фетальная бычья сыворотка

FDR – (англ. false discovery rate) уровень ложноположительных результатов

Fmoc-L-Arg – (англ. N_{α} -Fmoc- N_{ω} – (2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)-L-

arginine) N_α-Fmoc-N_ω – (2,2,4,6,7-пентаметилдигидробензофуран-5-сульфонил)-L-аргинин

 $Fmoc-L-Lys-(англ. \ N_{\alpha}-Fmoc-N_{\omega}-(2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)-L-lysine)$

 N_{α} -Fmoc- N_{ω} – (2,2,4,6,7-пентаметилдигидробензофуран-5-сульфонил)-L-лизин

HCD – (англ. higher-energy collisional dissociation) диссоциация, вызванная столкновением электронов при высокой энергии

HCTU – (англ. O-(6-Chlorobenzotriazol-1-yl)-N,N,N`,N`- tetramethyluronium

hexafluorophosphate) О-(6-хлорбензотриазол-1-ил)-N,N,N`,N`-тетраметилурана гексафторофосфат

HECD – (англ. hot electron capture dissociation) диссоциация с захватом горячих электронов

IAA – (англ. iodoacetic acid) йодуксусная кислота

IAM – (англ. iodoacetamide) йодацетамид

MMTS – (англ. s-methyl methanethiosulfonate) метилметантиосульфонат

ppm – (англ. parts per million) частей на миллион

RPART – (англ. *recursive partitioning and regression trees*) рекурсивное разделение и деревья регрессии (библиотека для языка R)

SDC – (англ. sodium deoxycholate) натриевая соль дезоксихолиевой кислоты

SDS – (англ. sodium dodecyl sulphate) додецилсульфат натрия

TEABC – (англ. *triethylammonium bicarbonate buffer*) буферный раствор триэтиламмония бикарбоната

TFA – (англ. trifluoroacetic acid) трифторуксусная кислота

TIS – (англ. triisopropylsilane) триизопропилсилан

ТМР – (англ. 2,2,6,6-tetramethylpiperidine) 2,2,6,6-тетраметилпиперидин

введение

Актуальность и степень разработанности темы

В последние годы точность и чувствительность панорамного протеомного анализа сильно возросла. Основной причиной такого успеха служит изобретение масс-спектрометра типа Orbitrap, позволяющего за один стандартный эксперимент идентифицировать несколько тысяч белков в одном образце. Разумеется, увеличение глубины анализа поставило ребром вопрос изучения факторов, влияющих на его качество, таких, как химические модификации аминокислот в составе белков. С развитием методов протеогеномики, и вообще, с выходом на передний план так называемых мультиомик, проблема налаживания методов анализа данных встала особо остро. В современной науке ощущается тенденция перехода от обычной протеомики к комплексному исследованию генома, транскриптома и протеома вместе, и соотнесению одного другому.

Протеогеномный метод актуален, в первую очередь, для биомедицинских исследований. Исследуя белковые продукты мутантных генов, возможно обнаружить новые биомаркеры и так называемые неоантигены, специфичные участки белков, характерные для раковых клеток, которые представляют потенциальные мишени для раковых вакцин [1–3]. Поскольку в ходе протеогеномного анализа изучают одноаминокислотные замены, то модификации аминокислотных остатков, даже редкие, могут существенно влиять на результат. За пределами протеогеномики уже описано множество примеров химических модификаций, имитирующих важные, с биологической точки зрения, пост-трансляционные события. Так, Nielsen с соавт. [4] показали, что при пробоподготовке к протеомному анализу возможна артефактная модификация остатка лизина, молекулярная масса которой равна диглициновым меткам сайтов убиквитинирования. В других исследованиях было показано влияние химической обработки белков на результаты исследования сайтов фосфорилирования [5]. Иными словами, изучение влияния методов пробоподготовки на результаты протеомного анализа имеет важное значение для улучшения качества последнего. Особенно это важно в специализированных случаях, таких как исследование пост-трансляционных модификаций или протеогеномный анализ.

Исследования влияния алкилирующих агентов на результаты протеомного анализа ведутся давно, но количество таких работ невелико [6,7]. По большому счету, детальная оценка влияния различных алкилирующих агентов на результаты протеомного массспектрометрического анализа проведена только в одной работе [7]. Мы дополнили эти результаты анализом двух других алкилирующих агентов и отдельно внимательно изучили преобразование метионина в изотреонин, не затронутое в указанной статье.

Целью работы было исследование влияния разных способов искусственной модификации остатков цистеина при пробоподготовке к протеомному анализу для оптимизации метода протеогеномики.

Для достижения этой цели были поставлены следующие задачи:

- Оценка частоты преобразования метионина в изотреонин в протеомных данных и подбор оптимального для протеогеномики метода пробоподготовки, позволяющего избежать маскировки такой модификации под аминокислотную замену метионина на треонин.
- Проведение сравнительного анализа четырех часто используемых в протеомике алкилирующих агентов, а именно йодацетамида (IAM), 4-винилпиридина (4-VP), метилметантиосульфоната (MMTS) и хлорацетамида (CAM), и оценка влияния каждого из них на результат протеомного анализа.
- Сравнение спектров фрагментации пептидов, содержащих треонин и изотреонин для оценки возможности распознавания этих аминокислотных остатков в пептидах при анализе протеома.

Научная новизна работы и практическая ценность

В работе проведен анализ влияния методов искусственной обработки остатков цистеина на результаты протеомного масс-спектрометрического анализа. Помимо изученных ранее алкилирующих агентов, в анализ были включены два других, не проанализированных в сравнительных работах, но применяемых в протеомике, а именно 4-винилпиридин и метилметантиосульфонат. Впервые изучено преобразование метионина в изотреонин в процессе пробоподготовки. Исследована частота данного явления на нескольких наборах данных и выдвинуты предположения о его причинах. Проведен детальный анализ данных, в том числе методом «открытого» поиска, в ходе которого также впервые продемонстрированно наличие сдвига массы пептидов с метионином, свидетельствующее о его преобразовании в изотреонин.

Результаты работы и сделанные выводы помогут исследователям грамотно подбирать методы пробоподготовки к протеомному анализу, учитывая задачу исследования. Адекватно выбранный подход к обработке белков в ходе анализа способствует уменьшению вероятности ложных результатов и артефактов и в целом улучшает результат.

Методология и методы диссертационного исследования

В работе проведены как собственные эксперименты, так и обработка общедоступных данных других исследователей, что, в настоящий момент, является распространенной

практикой в протеомных исследованиях. Анализ данных проведен по современным методикам с использованием как готового программного обеспечения, так и самостоятельно написанных скриптов на двух языках программирования. Интерпретация массспектрометрических данных проведена в соответствии с рекомендациями глобального проекта «Протеом человека» [8]. Экспериментальные данные получены современными методами масс-спектрометрии, их описание приведено в соответствии с требованиями ведущих протеомных журналов к описанию масс-спектрометрических экспериментов. Данные, полученные в экспериментальной части работы, помещены в репозиторий ProteomeXchange и доступны другим исследователям для проверки и использования, что, также, соответствует требованию большинства научных изданий в области протеомики.

Положения, выносимые на защиту

1. Анализ преобразования метионина в изотреонин в процессе протеомного эксперимента показал, что метионин преобразуется в изотреонин, в среднем, не более, чем в 1% метионинсодержащих пептидов.

2. При использовании в качестве алкилирующих агентов йодсодержащих веществ, таких как йодацетамид и йодуксусная кислота, количество пептидов с преобразованием метионина в изотреонин может достигать 4% от всех пептидов, содержащих этот аминокислотный остаток. Эта реакция происходит чаще, если образцы проходят электрофоретическое разделение в полиакриламидном геле.

3. Путем оценки влияния соседствующих аминокислотных остатков на частоту преобразования метионина в изотреонин показано, что пролин, следующий за метионином в последовательности белка, увеличивает частоту этой реакции.

4. При сравнительном анализе четырех алкилирующих агентов, а именно йодацетамида, хлорацетамида, 4-винилпиридина и метилметантиосульфоната показано, что из исследованных соединений для протеомного анализа лучше всех подходит хлорацетамид. Он вызывает наименьшее количество побочных реакций и приводит к идентификации наибольшего количества пептидов.

5. Показано, что наиболее часто используемый алкилирующий агент йодацетамид может вызывать карбамидометилирование метионина и, тем самым, препятствовать идентификации пептидов, содержащих этот аминокислотный остаток. Продемонстрировано применение метода «открытого» поиска для учета подобных нежелательных модификаций аминокислотных остатков.

6. Путем сравнения хроматограмм и масс-спектров сходных синтетических пептидов, содержащих треонин или изотреонин, показано, что содержащие треонин и изотреонин

пептиды характеризуются разным временем выхода с хроматографической колонки, а также различаются по масс-спектрам фрагментации.

Личный вклад соискателя

Соискатель проводил поиск и анализ литературы, планирование эксперименов, пробоподготовку к протеомному анализу и анализ данных, интерпретацию результатов и подготовку публикаций. Работа была выполнена в лаборатории медицинской протеомики ИБМХ в период с 2015 по 2020 год.

Степень достоверности и апробация результатов

Достоверность результатов работы подтверждается согласованностью результатов современных экспериментальных и вычислительных методов протеомного анализа, использованных в работе. Обсуждение результатов проведено с учетом современных исследований, опубликованных в области белковой биохимии, протеомики и протеогеномики. Научные положения и выводы, изложенные в диссертации, обоснованы и подтверждены фактическим материалом.

Основные результаты диссертационной работы были представлены в виде стендовых докладов на конференциях таких, как конгресс Европейской Протеомной Ассоциации (EuPA 2016, Стамбул), конгресс международной организации «Протеом человека» (HUPO 2017, Дублин) и конгресс Европейской Протеомной Ассоциации (EuPA 2018, Сантьяго де Компостела) и других.

По теме диссертации опубликовано 18 работ, из которых 10 статей в рецензируемых научных журналах и 8 публикаций в трудах конференций.

ГЛАВА №1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Историческое развитие методов протеомного анализа

Понятие протеома как белкового эквивалента генома появилось не так давно, в 1994 году [9], в то время как работа с протеомами, а тем более с отдельными белками, была начата намного раньше. Еще в пятидесятые годы прошлого века появился метод определения аминокислотной последовательности пептидов, который в настоящее время принято называть методом деградации Эдмана [10]. Изучить и, в буквальном смысле, увидеть сам протеом исследователям удалось в начале семидесятых годов двадцатого века с появлением разделения белков при помощи электрофореза в полиакриламидном геле. Сначала белки разделяли по молекулярной массе в денатурирующих условиях в присутствии додецилсульфата натрия [11]. Практически сразу после этого появился метод так называемого двумерного электрофореза - разделения белков на геле в двух направлениях - по изоэлектрической точке и по молекулярной массе [12]. Уже в это время в рутинных протеомных методах закрепились такие этапы денатурации белков, как обработка образцов детергентами, например, додецилсульфатом натрия, а также нагревание вплоть до кипячения [11]. Двумерный электрофорез позволил не только увидеть протеом, но и сравнивать протеомы между собой, вырезать из геля и анализировать отдельные белки. Электрофорез как метод разделения белков используют в биомедицинскоих исследованиях до сих пор. Он позволяет визуализировать результат, а также получать очищенные белки или белковые комплексы.

Развитие методов масс-спектрометрии белков и пептидов, в том числе, появление мягкой ионизации соединений, например, ионизации электрораспылением [13], в сочетании с использованием для идентификации белков поиска по геномным последовательностям, дало толчок к развитию так называемой панорамной (shotgun) протеомики [14]. Смесь белков для этого анализа расщепляют на пептиды специфичной протеазой, например, трипсином, а потом получившиеся пептиды анализируют с использованием высокоэффективной жидкостной хроматографии, совмещенной с тандемной масс-спектрометрией (ВЭЖХ-МС/МС, или LC-MS/MS) [15]. В связи с этим, необходимость разделять белки в геле электрофорезом перед анализом отошла на второй план. В настоящее время стало возможным готовить лизаты из биологических образцов, проводить в них ферментативный гидролиз белков с образованием пептидов и сразу начинать хроматомасс-спектрометрический анализ.

Технический прогресс в протеомике повлек за собой необходимость развития методов анализа данных. Имея образец ткани или клеток, исследователь на выходе получал файлы с

масс-спектрами, которые представляют собой огромные списки масс пептидов и их фрагментов, зафиксированные прибором во время анализа. Сделать из этих списков масс понятные человеку списки пептидов и белков способны только сложные компьютерные программы, которые называются поисковыми машинами и работают по принципу сравнения имеющихся экспериментальных спектров с теоретическими спектрами, соответствующими известным пептидам из баз данных для конкретного организма.

Стартовавший в 2010 году проект "Протеом человека", являющийся логическим продолжением, но не смысловым аналогом, завершенного проекта "Геном человека", породил настоящую гонку за "глубокий протеом". Лаборатории в странах-участниках проекта начали анализировать протеомы и генерировать невероятное количество данных, публикуя статьи с все большим и большим количеством идентифицированных белков. Тут в протеомный анализ, в каком-то смысле, вернулся гель-электрофорез, поскольку разделять белки перед гидролизом и анализировать каждую фракцию отдельно, разделяя в ней пептиды, оказалось неплохой идеей для достижения все большего и большего разрешения метода. В конце концов скорострельный протеомный анализ достиг около 10 тысяч идентифицированных белков в одном образце [16].

На протяжении всего этого времени, протеомные методы развивались достаточно активно и, в большей степени это касалось масс-спектрометрии – наиболее технологически сложного этапа анализа. Отдельной областью для развития была биоинформатика, то есть обработка масс-спектрометрических данных, от которой тоже существенно зависит результат. Из одних и тех же спектров разными методами можно получить существенно отличающееся количество идентифицированных белков. В свою очередь, пробоподготовка, являясь немаловажным фактором получения хороших спектров, зачастую имела второстепенное значение. Занятые вопросами масс-спектрометрии и обработки результатов, исследователи не всегда уделяют должного внимания пробоподготовке и берут протоколы из похожих работ. Таким образом, одни и те же методы становятся традиционными и, не смотря на недостатки, копируются из работы в работу. Тем не менее, все же существует немало статей, посвященных анализу разных методов, сравнению их и оптимизации, о которых подробно будет сказано позже.

Разберемся какие методы протеомного анализа являются наиболее распространенными в настоящее время, какие они имеют преимущества и недостатки и какими ошибками в результатах чревата подготовка образцов по тому или иному методу.

1.2. Основные методы протеомного анализа

Метод панорамного протеомного анализа включает в себя много этапов (рис. 1), гомогенизацию и лизис образца [17], разделение белков электрофорезом [11], восстановление и алкилирование остатков цистеина [18], ферментативный гидролиз с образованием пептидов [10], а также хромато-масс-спектрометрический анализ этих пептидов [15]. В ходе пробоподготовки могут протекать побочные химические реакции, влияющие на результат протеомного анализа. Среди примеров, спонтанное окисление метионина [19] или неферментативное дезамидирование глутамина и аспарагина до, соответственно, глутаминовой и аспарагиновой кислот [20]. Одним из наиболее важных является этап восстановления дисульфидных связей и последующего алкилирования образованных свободных тиольных групп, поскольку на этом этапе достаточно часто происходит образование артефактов, влияющих на конечный результат [21].

Любая манипуляция с образцами влечет за собой потери материала и опасна загрязнением внешними белками такими, как человеческий кератин. По этой причине исследователи стремятся сократить количество этапов при пробоподготовке. С другой стороны, хорошая очистка смеси белков от других веществ и денатурация белков обеспечивает доступ ферменту к сайтам гидролиза. В качестве такого фермента чаще всего используют трипсин благодаря его высокой специфичности [22], однако, в особых случаях используют и другие ферменты [23]. Таким образом, при выборе стратегии пробоподготовки, необходимо соблюсти разумный компромисс между количеством этапов и качеством гидролиза. Далее рассмотрим три основных метода гидролиза, применяемых в настоящее время: гидролиз в геле, в растворе и на фильтрах (рис. 1). У всех трех методов есть недостатки и преимущества и во всех этих методах есть этапы, способные влиять на результат.



Рис. 1. Основные пути пробоподготовки к масс-спектрометрическому протеомному анализу: А – в геле; Б – в растворе; В – с использованием центрифужных фильтров. Основные этапы в разных вариантах: І – лизис образца; IIа – разделение белков методом электрофореза; IIб – перенос раствора белков в верхнюю камеру фильтра; III – восстановление и алкилирование остатков цистеина; IV – ферментативный гидролиз; Va – экстракция пептидов из геля; Vб – фильтрация пептидов сквозь мембрану фильтра, перемещение их в нижнюю камеру.

Гидролиз, как уже было сказано обычно это трипсинолиз, в геле проводят, в основном тогда, когда перед анализом белки разделяли по массе. Это необходимо для разделения смеси белков и достижения большей «глубины» протеомного анализа, то есть получения большего количества белковых идентификаций, или же, если исследователей интересуют белки из какого-то более узкого диапазона масс. В конце концов, до сих пор существуют задачи, для которых белки сначала разделяют двумерным электрофорезом, а потом анализируют их методом ВЖЭХ-МС [24]. Для всего этого был предложен метод проведения трипсинолиза белков, заключенных в гель [25]. В действительности идея хороша тем, что белки в геле денатурированы, и обеспечив доступ фермента в гель, о доступности сайтов гидролиза в самих белках можно волноваться уже меньше. Тем не менее, этот метод известен тем, что, в результате, пептидных идентификаций получается все же меньшее количество [26], поскольку, по-видимому, эффективность такого гидролиза ниже.

1.3. Биоинформатический анализ протеомных данных

1.3.1. Протеомный поиск

Полученные при помощи масс-спектрометра данные анализируют с использованием поисковых протеомных машин [27]. Поисковые программы сравнивают экспериментально полученные спектры фрагментации с теоретически предсказанным набором фрагментов для пептидов из базы данных, при этом в разных алгоритмах может учитываться как совпадение масс, так и интенсивностей фрагментов [28,29]. Каждое статистически значимое совпадение называют «спектральной идентификацией» (peptide spectrum match, PSM). Спектральные идентификации являются подтверждением аминокислотной последовательности пептидов, при этом чем больше таких спектров, тем надежнее идентифицирован пептид. Далее полученные пептиды «собирают» в белки [30]. Такой подход называется «снизу вверх» (bottom-up). Поскольку те или иные модификации аминокислотных остатков всегда существуют и возникают как *in vivo*, так и во время пробоподготовки *in vitro*, наиболее часто встречающиеся модификации учитываются при протеомном поиске как постоянные (то есть предполагается, что все такие остатки несут данную модификацию) или изменяющиеся, «вариабельные» (когда поисковая программа учитывает варианты с и без модификации). В настоящий момент существуют и используются и другие методы работы с протеомами [31,32], однако, панорамный анализ «снизу вверх» наиболее широко используется в практике биологических исследований.

1.3.2. Протеогеномный подход

Впервые термин «протеогеномика» был предложен тогда, когда протеомные данные стали применять для уточнения аннотации геномных данных [33]. Более же широкое распространение данная технология получила при интеграции геномных и протеомных данных с целью поиска белковых продуктов генов, содержащих те или иные полиморфизмы [34]. В основе этого метода лежит следующий принцип. Для обычного протеомного поиска используют консенсусную базу данных, содержащую аминокислотные последовательности белков изучаемого организма. При протеогеномном поиске в эту базу добавляют последовательности, содержащие варианты (Рис. 2). Такие варианты могут быть предсказаны на основании геномных данных, например, о раковых мутациях [35]. Также такое дополнение к базе данных может быть получено на основании транскриптомных данных, например информации о редактировании РНК [36] или альтернативном сплайсинге [37]. Далее проводят протеомный поиск по новой дополненной базе данных и получают информацию о белковых продуктах интересующих вариантов (Рис. 2).



Рис. 2. Принцип метода протеогеномики. Сначала в консенсусную базу данных для изучаемого организма добавляют вариантные последовательности, полученные на основе геномных данных. Далее проводят протеомный поиск по этой базе при помощи поисковой машины и экстрагируют из результатов пептиды, содержащие одноаминокислотные замены.

1.3.3. Расчет уровня ложноположительных результатов

Одна из основных проблем в протеомном поиске заключается в отборе достоверно идентифицированных пептидов и белков. Для этого применяют так называемый таргетдекойный подход (target-decoy approach) [38]. В базу данных аминокислотных последовательностей добавляют «декойные», то есть заведомо несуществующие в данном организме, пептиды. При обработке результатов поиска рассчитывают уровень ложноположительных результатов или FDR (false discovery rate) [39]. По сути FDR равен содержащемуся в результатах анализа проценту декойных пептидов, идентифицированных поисковой машиной как достоверных. В конечном итоге результаты фильтруют до нужного процента FDR, например, до 1%. Такую фильтровку можно проводить как на уровне спектральных идентификаций (PSM), так и на уровне пептидов и белков. Многие поисковые программы и инструменты последующей обработки результатов проводят фильтровку до нужного FDR автоматически [40–42].

Отдельной задачей является достоверная идентификация вариантных пептидов в протеогеномном анализе, или же модифицированных пептидов при изучении модификаций

аминокислотных остатков. Дело в том, что вероятность появления пептидов, содержащих одноаминокислотные замены (варианты) или некоторые модификации может отличаться. Вариантные белки в силу естественных причин могут реже встречаться в организмах [43,44]. Исследователи обратили внимание на проблему расчета FDR в группах довольно давно [45]. Ранее было предложено два принципиально различающихся подхода к протеомному поиску и расчету FDR при анализе вариантных пептидов [46]. Расчет комбинированного FDR (combined FDR) [46], или подход «все вместе» ("all-together") [35], не предполагал никакого отдельного поиска и расчета FDR для пептидов содержащих варианты. Референсную протеомную базу данных для изучаемого организма комбинировали с базой данных, содержащей вариантные последовательности. Далее проводили протеомный поиск против этой базы данных, а результаты фильтровали на уровне PSM до нужного процента FDR [46]. Расчет отдельного FDR (separate FDR) [46], или подход «один за другим» ("one-by-one") [35], предполагал составление раздельных баз данных с белками «дикого» типа и вариантными белками. Поиск сначала проводили по базе данных с белками «дикого» типа, получившиеся PSM фильтровали до нужного уровня FDR. Далее, проводили поиск спектров не совпавших ни с чем в первой базе данных, по базе содержащей вариантные последовательности. Уже в такой группе рассчитывали свой отдельный FDR и проводили фильтровку [35,46]. Таким образом удавалось более достоверно отобрать пептиды, относящиеся к определенной группе, например, вариантные пептиды, содержащие одноаминокислотные замены. Позднее FDR в группах стали рассчитывать по уточненной формуле [47].

1.4. Основные этапы пробоподготовки в протеомном анализе

1.4.1. Лизис образца и выделение смеси белков

Первым этапом любого метода является лизис образца. В качестве образцов используют почти любой биологический материал и, в зависимости от его природы, лизис образца и выделение белков проводят немного по-разному. Если в качестве образца используют ткани, то их необходимо гомогенизировать механически. Если в качестве образца используют клеточные культуры, то такие клетки отмывают от среды изотоническим буфером, а затем центрифугируют и лизируют осадок. На данном этапе важно избавиться от остатка солей, поскольку соли, в некоторых условиях, препятствуют хроматографическому разделению и электроспрейной ионизации [48]. В зависимости от метода гидролиза используют разные способы лизиса клеток. Если белки потом наносят на одномерный электрофорез, то в качестве лизирующего буфера, обычно, используют стандартный буфер Лэммли [11], содержащий додецилсульфат натрия (SDS) и восстанавливающий агент,

например, дитиотреитол (DTT) или β-меркаптоэтанол (BME). Восстанавливающий агент разрывает S-S связи в белках, a SDS связывается с белками, обеспечивая их денатурацию [49] и подвижность на электрофорезе, обратно пропорциональную их молекулярной массе [50]. Подвижность белков в полиакриламидном геле зависит от концентрации акриламида, то есть от величины пор в полимерной сети. Чем больше молекулярная масса белка, тем медленнее он движется в геле под воздействием электрического тока [50]. Лизис образца с использованием SDS очень эффективен, однако, он не подходит для двумерного электрофореза, поскольку при изоэлектрической фокусировке белков недопустимо использование ионных детергентов [51]. Самый главный недостаток использования SDS для лизиса клеток заключается в том, что, оставшись в образце, он, в последствии, влияет на время выхода пептидов с колонки во время хроматографии [52]. Помимо этого, SDS, если он присутствует в растворе пептидов, который анализируют на масс-спектрометре, мешает получить хороший результат. Будучи поверхностно-активным веществом, он препятствует образованию капель раствора при электроспрейной ионизации [53]. Допустимая концентрация SDS в растворе варьирует от 0.001% до 0.1% по результатам разных исследований [51]. Тем не менее, возможность использования SDS является преимуществом методов пробоподготовки в полиакридамидном геле, поскольку, связанные с этим гелем белки перед гидролизом можно отмывать практически от любых детергентов. Так, в случае двумерного электрофореза, не имея возможности на первом этапе применять SDS, большинство протоколов предполагают лизис клеток в буферах, содержащих мочевину, тиомочевину, а также такие неионные или цвиттерионные детергенты, как CHAPS, TrironX100 и другие [12].

При проведении гидролиза в растворе использовать солевые буферы, SDS, а также большинство детергентов нельзя, поскольку они не совместимы с хроматографией и массспектрометрией. Поэтому дальнейшее развитие в этой технической области пошло в двух направлениях: поиск и разработка совместимых с масс-спектрометрией детергентов [54–57] или подбор методов предварительной очистки раствора [58–60].

Одним из наиболее часто используемых детергентов, совместимых с хроматографией и масс-спектрометрией является натриевая соль дезоксихолевой кислоты (SDC). Особенностью этого реагента является то, что он выпадает в осадок в кислых условиях. Поскольку подкисление буфера и так производится после трипсинолиза для остановки ферментативной реакции, на этом этапе осажденный детергент просто удаляют из раствора при помощи центрифугирования [54,55]. Позже были разработаны специальные соединения, названные компаниями-разработчиками также «усилителями трипсина», а именно, ProteaseMAXTM [56] и RapiGestTM [57] (Рис. 3). Принцип их применения похож на метод дезоксихолата натрия. Так, в кислых условиях, после трипсинолиза, эти реагенты распадаются

на два компонента, один из которых выпадает в осадок, а второй – не препятствует хроматографии и масс-спектрометрии [56,57]. ProteaseMAX нашел свое применение в одних из самых прогрессивных и успешных с технической точки зрения работах [61]. RapiGest же получил меньшую популярность, по данным некоторых авторов, он уступает даже традиционному дезоксихолату [52], хотя этот реагент применяют и в настоящее время [62]. Существуют сообщения о том, что, в некоторых случаях, ProteaseMAX может являться источником артефактов [63]. В частности, использование ProteaseMAX перед трипсинолизом приводило к пальмитированию (+238 Да) и гидроксифарнезилированию (+220 Да) цистеинов. Предполагается, что этот реагент не подходит для экспериментов, связанных с липидными модификациями белков, а также для количественного анализа модификаций цистеинов [63].



Рис. 3. Детергенты, используемые в протеомике и способные выпадать в осадок в кислой среде после трипсинолиза. А: дезоксихолиевая соль натрия; Б: RapiGest; В: ProteaseMAX.

Как отмечалось ранее, в качестве альтернативы использованию осаждаемых детергентов для лизиса в растворе перед протеомным анализом применяют дополнительную очистку раствора от солей и детергентов. Так, довольно давно, существуют методы переосаждения белков [58], очистки белков на смоле C18 в специальных наконечниках или на колонках [59]. Все эти методы применяются, но неизбежно, ведут к потерям материала. В последнее время широкое распространение получил метод трипсинолиза на центрифужных фильтрах содержащих мембрану, состоящую из ацетата целлюлозы [64], названный впоследствии FASP (filter-aided sample preparation) [60] (Рис. 1). Этот метод, наряду с пробоподготовкой в геле, имеет большое преимущество при выделении мембранной фракции белков [65,66]. Лизаты загружают на центрифужные фильтры с размером пор около 20-30 кДа, такие фильтры не пропускают белки, но пропускают полученные в результате трипсинолиза пептиды. Далее раствор белков можно промывать буфером, что позволяет освободить его от большинства примесей, совместимых с фильтрацией данного типа. Все реакции проводятся в верхней камере фильтров, а получившиеся пептиды, потом смываются, сквозь фильтры в пробирку. Это обеспечивает сразу этап очистки смеси пептидов от слишком длинных фрагментов, получившихся в результате пропуска ферментом сайтов расщепления. Метод пробоподготовки на фильтрах приобрел широкое распространение и претерпел множество модификаций [67]. С целью упростить и ускорить этап пробоподготовки, а также сократить количество этапов, таким образом, уменьшив вероятность потерь и загрязнения образца, был разработан метод гидролиза в специальных наконечниках [68]. В отличие от метода центрифужных фильтров, этот метод пока не получил столь широкого распространения.

1.4.2. Обработка дисульфидных связей остатков цистеина для протеомного анализа

Как известно, в формировании третичной структуры белков участвуют дисульфидные (S-S) связи, соединяющие остатки цистеинов между собой. Это затрудняет массспектрометрический анализ пептидов, содержащих этот аминокислотный остаток. Полученные в результате расщепления трипсином или другими протеазами пептиды с сохраненными S-S-связями имеют массу, которую сложно предсказать на основе геномных данных, поэтому при протеомном поиске такие пептиды остаются неидентифицированными [18]. Для решения этой проблемы практически любой метод пробоподготовки предполагает восстановление остатков цистеина, то есть разрыв S-S-связей. Последние в свободном состоянии остаются реакционноспособными, поэтому их защищают от взаимодействий путем присоединения химически стабильных радикалов, чаще всего от алкилирующих агентов [18,69]. Реакции восстановления и алкилирования остатков цистеина, как правило, проводят после лизиса клеток и перед трипсинолизом белков, добавляя в раствор последовательно сначала восстанавливающий, а затем алкилирующий агенты [70].

1.4.2.1. Восстановление дисульфидных связей в третичной структуре белка

Для разрыва дисульфидных связей между остатками цистеина наиболее часто применяют дитиотреитол (DTT) [71], β-меркаптоэтанол (BME) [11] и трис(2-карбоксиэтил)фосфин (TCEP) [72]. Типичными для реакции восстановления условиями является температура 57°С-60°С и время около 40-60 мин [70]. Кроме перечисленных

соединений, встречаются упоминания и о других фосфинах [73,74], но они редки. Существует несколько работ, анализирующих эффект разных восстанавливающих агентов на результаты протеомного анализа [7,74]. В работе [74] авторы сравнивали DTT, BME, TCEP и редко применяемый трис(2-гидроксипропил)фосфин (THPP) и не обнаружили в их эффекте существенных различий. В работе [7] были проведены более сложные эксперименты, включающие анализ сочетания восстанавливающих и алкилирующих агентов. Кроме того, авторы провели пробоподготовку как в геле, так и в растворе, а точнее был использован так называемый метод центрифужных фильтров [60,64] (рис. 1). В данной работе наибольшее количество спектральных идентификаций получили при использовании в качестве восстанавливающего агента DTT, в случае пробоподготовки в геле и BME – в случае пробоподготовки в растворе. В сочетании с обоими реагентами наилучшим алкилирующим агентом соединением акриламид. В ранее опубликованных статьях [7,74] не описано никаких обнаруженных нежелательных побочных эффектов, возникающих на этапе восстановления.

1.4.2.2. Алкилирование остатков цистеина

Этап алкилирования следует за этапом восстановления и привлекает к себе большее внимание исследователей. Именно на этом этапе возникает наибольшее количество модификаций аминокислотных остатков. Существует несколько работ, в которых сравнивали различные алкилирующие агенты с точки зрения их эффективности и влияния на результат [6,7,74]. Сделанные в этих работах выводы частично согласуются, но есть и проитворечивые результаты. Так, например, в работе [7] не рекомендуется использовать йодацетамид, а в работе [74], наоборот, авторы делают вывод, что это наиболее подходящий алкилирующий агент. Подробно результаты работ, сравнивающих алкилирующие агенты, будут описаны далее.

Панорамный протеомный анализ, главным образом, преследует цель идентификации в отдельно взятом образце наибольшего количества пептидов, а, следовательно, и белков. Именно число идентифицированных пептидов и число соответствующих им спектральных идентификаций являются основными критериями эффективности и чувствительности протеомного анализа. Помимо этого, важную роль играет то, чтобы в результатах были представлены пептиды независимо от их аминокислотного состава, что обеспечивает воспроизводимость результатов [7]. Если какие-то аминокислотные остатки претерпевают модификации, не учитываемые при протеомном поиске, то содержащие их пептиды не могут быть идентифицированы. Их молекулярная масса изменяется, а экспериментально измеренный масс-спектр не совпадает с теоретическим, сформированным при протеомном

поиске из пептидов базы данных [75]. Таким образом, неучтенные модификации аминокислот сильно снижают качество результатов анализа.

В качестве алкилирующих агентов в протеомном анализе применяют: (i) галогаензамещенные амиды и кислоты, например, йодацетамид (IAM) [28], хлорацетамид (CAM) [4] и йодуксусную кислоту (IAA) [6], (іі) пиридины, в том числе, 4-винилпиридин (4-VP) и 2-винилпиридин (2-VP) [76], а также (iii) акриламид (AA) [77], (iv) метилметантиосульфонат (MMTS) [78] и (v) N-этилмалеимид (NEM) [79] (рис. 4). Алкилирование обычно проводят в течение 10-30 минут при комнатной температуре в темноте [70]. Повышение температуры реакции алкилирования может привести к возрастанию количества побочных продуктов [80]; кроме того, алкилирующие агенты неустойчивы на свету [81]. Один из вариантов метода также предполагает дополнительный этап: после обработки алкилирующим агентом в образец добавляют восстанавливающий агент, например, DTT, чтобы заблокировать избыток первого соединения, не прореагировавшего с остатками цистеина [25]. Это должно препятствовать алкилированию других аминокислотных остатков, которое может возникнуть при использовании некоторых соединений [21]. Такой этап имеет смысл в случае пробоподготовки в растворе, поскольку в геле, в котором зафиксированы целевые белки, после каждой стадии происходит отмывка реагентов буферным раствором [25].



Рис. 4. Модификации остатков цистеина используемыми в протеомике алкилирующими агентами.

Анализ литературы показывает, что чаще всего для алкилирования остатков цистеина в протеомике используют IAM [7]. В частности, IAM был использован более чем в 80% работ, опубликованных в двух основных специализированных журналах, а именно, «Molecular and Cellular Proteomics» и «Journal of Proteome Research», в 2016-2017 годах. Однако, в последнее время, все больше исследователей выбирают альтернативные алкилирующие агенты. Накопились сведения о том, что IAM является причиной нежелательных химических модификаций аминокислотных остатков [4,7,21,82]. Наиболее распространенной побочной реакцией является избыточное алкилирование (overalkylation) IAA и IAM. Этот эффект заключается в том, что реагенты могут алкилировать не только целевые сульфгидрильные группы цистеина, но и другие аминокислотные остатки, например, тирозин, гистидин, лизин, метионин и другие.

1.5. Алкилирование остатков цистеина: побочные реакции

В экспериментах с использованием синтетических пептидов было показано, что кроме цистеинов, алкилированию подвергаются в первую очередь N-концевые аминогруппы [21]. Во вторую очередь, модификации подвергаются другие аминокислотные остатки, содержащие метилтио- или аминогруппу (то есть, метионин и лизин) а также гистидин, содержащий имидазольное кольцо [21,83]. Кроме них, алкилированию могут подвергаться тирозин и Сконцевая карбоксильная группа пептида [83]. В одной из работ [21] авторы пробовали избежать эффекта алкилирования IAM аминогрупп, проводя реакцию в бикарбонате аммония, но этот подход не дал желаемого результата. В итоге, они предложили использовать для удаления избытка ІАМ тиоэфиры. Эти и другие аналогичные исследования [21,83] были выполнены на синтетических пептидах, и в соответствующих публикациях подробно описаны химические процессы, лежащие в основе эффекта избыточного алкилирования белков ІАМ и IAA. Однако, в них не проанализирован весь процесс пробоподготовки в протеомном анализе, и, кроме того, не использовался восстанавливающий агент. Протеомный эксперимент обязательно требует проведения реакции восстановления S-S связей перед проведением алкилирования, при этом, большая часть алкилирующего агента должна связываться с освободившимися тиольными группами остатков цистеина, а не с N-концом белка. Ко всему прочему, на этом этапе белок еще не расщеплен протеазами на пептиды, а к моменту образования последних реакция алкилирования, скорее всего, уже прекращается, поскольку, как указано выше, применяемые алкилирующие агенты распадаются на свету. Несмотря на эти соображения, многие работы указывают на то, что IAM и IAA все же алкилирует Nконцевые аминогруппы многих протеолитических пептидов [7,74,83,84].

Помимо работ с синтетическими пептидами, процесс алкилирования аминокислотных остатков изучали и на выделенных нерасщепленных белках [82]. Было показано, что IAA и IAM, помимо цистеина, алкилируют метионин, гистидин и лизин. В этой работе авторы рекомендуют подбирать эквимолярное соотношение алкилирующего агента и тиольных групп цистеина для того, чтобы алкилирующий агент расходовался полностью. В эксперименте с целым протеомом такое, разумеется, невозможно, поскольку образец представляет собой неизвестную смесь огромного количества белков. Кроме того, показано, что 4-винилпиридин также может вызывать избыточное алкилирование при повышенных температурах, интенсивность которого ниже, чем у IAA и IAM [82]. В ряде других исследований утверждается, что и CAM способен вызывать такой артефакт, но происходит это также при повышенных температурах [80].

Самый подробный анализ избыточного алкилирования различными агентами был проведен в работе [7], авторы которой показали, что в результатах проведенного ими анализа протеомов присутствуют пептиды, алкилированные IAM, IAA, AA и CAM на N-конце. Также, помимо цистеина, алкилирование было обнаружено на остатках тирозина, серина, треонина, гистидина, лизина, а также глутаминовой и аспарагиновой кислот. Стоит отметить, что при пробоподготовке в растворе в этом исследовании после реакции алкилирования добавляли восстанавливающие агенты для гашения избытка алкилирующих агентов. Из всех исследованных алкилирующих агентов САМ, вызвал наименьшее число нежелательных модификаций, а IAA – наибольшее. В этом эксперименте авторы не обнаружили алкилирование остатков метионина, хотя ранее было показано, что оно может происходить довольно часто [5,85]. Более детальный анализ с использованием особого массспектрометрического метода «многоэтапной активации» для анализа нейтральных потерь [86], позволил им разобраться, что происходит с метионином, и почему при использовании наиболее «агрессивных» алкилирующих агентов (таких как IAM и IAA) из результатов поиска практически пропадают пептиды, содержащие метионин. Кроме этого, в работе [7] использовали метод поиска модификаций пептидов в широком диапазоне масс [87]. При идентификации прекурсорных ионов, выставляли широкое окно ошибки масс (в данном случае ±250 Да), анализируя частоту встречаемости пептидов с теми или иными сдвигами масс. Такой метод, позднее, получил название «открытого поиска» [88], и о нем будет рассказано далее.

1.6. Нежелательные модификации алкилирующими агентами остатков метионина

Следует отметить, что метионин заслуживает отдельного рассмотрения, поскольку с ним, помимо алкилирования, происходят и другие известные модификации, как природные, так и артефактные. Метионин в большей степени, чем другие аминокислотные остатки, склонен к окислению в условиях протеомного эксперимента [19] (рис. 5А). Именно поэтому окисление остатка метионина в качестве «вариабельной» модификации является часто используемым параметром в протеомных поисках [89]. Окисление метионина может происходить как в клетке естественным путем [90–92], так и в процессе пробоподготовки (наряду с триптофаном) [19]. Окисление остатков метионина приводит к сдвигу массы на +16 Да (рис. 3А), а триптофана – на +4 Да, +16 Да, +20 Да и +32 Да с образованием остатков кинуренина , гидрокситриптофана, 3-гидрокситриптофана и N-формилкинуренина соответственно [93] (рис. 5). Известно, что окисленный метионин может претерпевать отщепление боковой цепи («нейтральную потерю») во время так называемой диссоциации, индуцированной соударением (*анел*. Collision-induced dissociation – CID), в результате которой

образуются пептидные ионы с молекулярной массой, уменьшенной на 64 Да [94] (рис. 3А). Показано, что такие пики чаще появляются в масс-спектрах после обработки белков йодсодержащими агентами, а именно IAM и IAA [7]. В то же время, нейтральная потеря боковой цепи метионина, вызывающая сдвиг молекулярной массы на минус 48 Да, также характерна для алкилирования IAM и IAA [7]. В экспериментах с синтетическими пептидами показано [7], что такие потери боковой цепи метионина вызывает не окисление, а именно обработка йодсодержащими агентами. Это противоречит некоторым результатам, полученным ранее [94], но соответствует другим данным [85]. В действительности, метионин теряет не 48 Да, а 105 Да (в случае обработки IAM), поскольку от него отщепляется не интактная, а модифицированная алкилирующим агентом боковая цепь [5,85]. Указанный сдвиг молекулярной массы на 105 Да включает в себя массу карбамидометила, равную 57 Да, и массу боковой цепи метионина, равную 48 Да (рис. 5Б).



Рис. 5. Модификации остатков метионина и последующие возможные нейтральные потери: А – окисление, Б – алкилирование йодацетамидом. Адаптировано из [95].



Рис. 6. Продукты окисления остатков триптофана. Адаптировано из [93].

В своей работе Müller и Winter [7] после детального анализа рекомендуют вместо йодсодержащих соединений использовать для алкилирования остатков цистеина АА. Также они получили удовлетворительное количество идентифицированных пептидов при минимуме побочных реакций при использовании хлорацетамида (САМ). Однако в ряде работполучены другие результаты. Например, при сравнении несколько отличного набора алкилирующих агентов, а именно, NEM, AA, IAM и 4-VP, авторы показали, что NEM вызывает наибольшую интенсивность алкилирования по N-концу пептида, за ним идет AA, только потом IAM и наименьшую вызывает 4-VP [74]. Данная работа также была выполнена на пептилах, полученных из белков после расщепления трипсином. При этом, авторы все равно рекомендуют использовать ІАМ, поскольку по их данным 4-VP не полностью алкилирует остатки цистеина. Стоит отметить, что эта работа выполнена не так тщательно, как исследование Müller и Winter [7]. В ней авторы не сравнивали комбинации алкилирующих и восстанавливающих агентов, а исследовали их эффекты отдельно. Также в ней: отсутствует «открытый» поиск [87], не проведены контрольные эксперименты на синтетических пептидах, а также не была выполнена «многоэтапная активация» [86], то есть не были проанализированы нейтральные потери так, как это делали Мюллер и Винтер [7]. Массспектрометрические данные их работы, вопреки сложившейся практике, не опубликованы в открытом доступе и не могут быть проверены и использованы для дальнейшего анализа. По

САМ имеются единичные указания на то, что этот агент способствует окислению остатков метионина и триптофана [96].

1.7. Выбор алкилирующих агентов при исследовании посттрансляционных модификаций

Одним из следствий алкилирования нецелевых аминокислотных остатков является то, что содержащие их пептиды остаются неидентифицированными, уменьшая покрытие протеома. Кроме этого, есть данные, что артефакты алкилирования могут стать причиной ложноположительных результатов при исследовании посттрансляционных модификаций белков, а также единичных аминокислотных замен. Так Nielsen с соавторами [4] показали, что IAM алкилирует остаток лизина, с образованием двойных аддуктов, молекулярная масса которых равна диглициновым меткам сайтов убиквитинирования (рис. 7). Особенно выраженно этот эффект проявлялся при нагревании. Таким образом, при изучении убиквитинирования не рекомендуется использовать IAM. Его предлагают заменить на САМ, который, по данным авторов этой работы, не склонен к алкилированию никаких аминокислотных остатков, кроме цистеина. В более поздней работе отмечается, что при нагревании выше 21°С такой эффект может вызвать и САМ [80]. В той же работе сообщается, что в качестве алкилирующего агента при исследовании сайтов убиквитинирования обычно применяют NEM, хотя его эффективность ниже, чем у IAM и САМ [97]. Несмотря на это, во многих работах по убиквитинированию все же используют IAM [98–100].



Рис. 7. Диглициновые метки, образующиеся на месте сайтов убиквитинирования по массе неотличимы от двойных аддуктов IAM. Адаптировано из [99].

При масс-спектрометрическом исследовании сайтов фосфорилирования обычно используют изоляционное окно 3,5 *m/z* (отношение массы иона к его заряду). При этом для

ионов с зарядами +2, +3 и +4 массы с разницей в 7 Да, 10,5 Да и 14 Да соответственно будут попадать в одно окно изоляции [5]. Таким образом, пептид со сдвигом массы на -105 Да, появляющийся в результатах при учете алкилирования IAM как модификации метионинов с отщепленной боковой цепью, будет попадать в то же изоляционное окно, что и такой же пептид с потерей остатка фосфорной кислоты массой 98 Да, характерной для фосфотреонина и фосфосерина [5]. Такие пептиды будут иметь один и тот же спектр фрагментации и будут технически неотличимы в результатах. Аналогичным образом добавочный ион Na⁺, имеющий массу 22 Да, в сочетании с карбамидометилированием (57 Да) дает сдвиг массы на 79 Да близкий по массе к самому фосфорилированию остатков треонина и серина (80 Да) [5]. Такие пептиды также будут иметь общий спектр фрагментации и без дополнительного разделения химерных спектров [101] один из них, а в некоторых случаях и оба пептида не будут идентифицированы. Более существенной проблемой может являться то, что сама фосфитная группа, имеющая массу 79,9663 Да, практически неотличима от сульфитной группы с массой 79,9568 Да.

1.8. Обработка остатков цистеина в протеогеномных исследованиях

С развитием интеграции геномных и протеомных данных (протеогеномики) проблема химических модификаций аминокислот, совпадающих по массе с аминокислотными заменами, несколько усугубилась. Дело в том, что в протеогеномных экспериментах в базу данных вносят пептиды с вариантами аминокислотных остатков, соответствующими закодированным в геноме или транскриптоме полиморфизмам [102]. Таким образом, поисковой машине предстоит найти спектры, соответствующие по молекулярной массе как пептиду с консенсусной последовательностью для данного организма, так и полиморфному пептиду. В случае, если какая-либо химическая модификация аминокислоты совпадает по молекулярной массе с разницей масс каких-либо аминокислот, и эта модификация случайно совпадет с сайтом аминокислотной замены, то такая замена может быть идентифицирована ошибочно.

1.9. Обработка остатков цистеина в редокс-протеомике

Кроме подробно описанных выше алкилирующих агентов, для той же цели используются и такие соединения как MMTS [76] и NEM [79]. Эти реагенты пришли в панорамный протеомный анализ из смежной области, а именно из редокс-протеомики [103]. Как уже упоминалось выше, NEM иногда использовался и в исследованиях сайтов убиквитинирования белков [80], а также в панорамной протеомике [79]. Редокс-протеомика –

это достаточно большая область протеомики, которая изучает окислительновосстановительные процессы в белках [103]. Одна из задач редокс-протеомики – изучение естественного состояния тиольных групп цистеинов. Таким образом, реакция алкилирования находятся в центре внимания методов, применяемых в этой области [104,105]. Стоит заметить, что MMTS, в отличие от всех остальных алкилирующих агентов, связываясь с цистеином, образует свою S-S связь и, таким образом, связывается обратимо (рис. 4). Это его свойство широко применяется для определения того, какие конкретно цистеины в белке были свободными от S-S связей, а какие – связанными [104]. Несмотря на это, MMTS используется и в панорамной протеомике, причем в ряде публикаций он признается более подходящим реагентом, чем тот же IAM [57]. В единичных работах было показано, что алкилирование при помощи NEM более эффективно, чем при использовании IAM и IAA [79,80]. При этом показано, что NEM способен алкилировать остатки лизина и гистидина [107], а также N-конец пептидов [74].

1.10. Современные решения для поиска модификаций аминокислотных остатков. «Открытый» поиск

Очевидно, что разные методы пробоподготовки лучше или хуже подходят для разных задач, будь то сравнение протеомов, исследование посттрансляционных модификаций или поиск несинонимичных мутаций на белковом уровне (табл. 1). Данные о недостатках и преимуществах тех или иных реагентов нередко противоречат друг другу. Таким образом, имеет смысл подбирать способ восстановления и алкилирования остатков цистеина для конкретных задач. Помимо пробоподготовки, на результат сильно влияет метод анализа данных, и в этой области также существуют разнообразные подходы. Важным шагом в развитии биоинформатических методов анализа протеомных данных в последние годы стал метод «открытого» поиска (рис. 8) [87].



профиль модификаций

Рис. 8. Общий принцип метода "открытого" поиска [87]. Сначала поисковая машина сравнивает экспериментальную массу прекурсорного иона с теоретическими массами пептидов из базы данных с широким окном допустимой ошибки определения массы ±500 Да, получая таким образом список "кандидатов". Затем спектр фрагментных ионов сравнивается со всеми теоретическими спектрами "кандидатов" со стандартным узким окном допустимой ошибки массы, и выбирается лучшее совпадение. Из чего исследователь получает информацию о сдвиге масс идентифицированного пептида, далее поиск продолжается аналогичным образом для всех прекурсорных ионов, и в конечном итоге можно получить информацию о количестве тех или иных ионов с определенными сдвигами масс.

В целом, идея метода состоит в том, что при анализе задается широкое окно допустимой ошибки определения масс для прекурсорных ионов, при этом идентификация основывается на классическом сравнении экспериментального спектра фрагментации с узким окном допустимой ошибки. Таким образом, на один экспериментальный спектр фрагментации приходится гораздо больше пептидов «кандидатов» из базы данных по сравнению с классическими поисками, однако точная идентификация по спектрам фрагментов позволяет контролировать уровень ложноположительных идентификаций (обычно методами «декойных» последовательностей). В результате можно оценить, какие сдвиги молекулярной массы по сравнению с теоретической чаще всего встречались.

Внося в качестве постоянных модификаций те, в которых исследователь уверен (например, алкилирование цистеина, если оно проводилось специально), можно увидеть, какие еще сдвиги массы затронули большое количество спектральных идентификаций. Таким

образом, проведя на первом этапе открытый поиск, исследователь делает вывод о наиболее часто встречающихся модификациях в конкретных данных. Далее эти модификации можно задать в качестве вариабельного параметра при основном протеомном поиске и получить большее количество идентифицированных пептидов, то есть более качественно интерпретированные результаты. Применение открытого поиска было продемонстрированно на протеомных данных [108,109]. Было показано, что учет модификаций, обнаруженных при открытом поиске, позволяет увеличить количество идентифицированных пептидов при основном поиске [108]. Кроме того, профиль модификаций, полученный при помощи открытого поиска, можно использовать для сравнения протеомов разных состояний [109].

Для открытого поиска разработаны специальные инструменты, например, поисковая программа MSFragger [88]. Кроме того, недавно разработанный на его основе инструмент AA_stat [108] позволяет для каждого найденного сдвига масс рассчитывать соотношение содержания аминокислотных остатков в пептидах, претерпевших этот сдвиг. Эта информация помогает выдвинуть предположение о том, какая конкретно аминокислота была модифицирована.

	IAA	IAM	AA	САМ	4-VP	MMTS	NEM
панорамная протеомика	не рекомендован избыточное алкилирование по N- концевой аминогруппе, Met, Hys, Lys, Tyr, C- конец [7,21,74,80,83,84], Ser, Asp, Glu [7]	не рекомендован избыточное алкилирование по N- концевой аминогруппе, Met, Hys, Lys, Tyr, C-term возможно использование с гашением избытка после алкилирования	рекомендован [7]	Рекомендован [7] возможно избыточное алкилирование при повышенной температуре [80], но с меньшей интенсивностью, чем IAA, IAM и AA [7]. может вызывать усиленное окисление Met и Trp [96]	возможно избыточное алкилирование при повышенной температуре [82], недостаточная эффективность алкилирования Cys [74]	Рекомендован [106] обратимое связывание, исключено гашение избытка после алкилирования	рекомендован возможно избыточное алкилирование по N- концу и Lys и Hys [107]
протеогеномика – поиск одноаминокислотных вариантов	не рекомендован способствует преобразованию Met в isoThr [110]	не рекомендован способствует преобразованию Met в isoThr [111]					
исследования ПТМ:							
фосфорилирование		применение исключено алкилирует Ser и Thr [5]					
убиквитинирование		применение исключено алкилирует Lys с образованием двойных аддуктов, не отличимых от сайтов убиквитинирования[4]		рекомендован при проведении экспериментов при 21°С [80]			рекомендован [80]
редокс-протеомика		используется в качестве необратимо связывающегося агента в комбинации с MMTS [104]				используется в качестве обратимо связывающегося агента в комбинации с NEM [104]	используется в качестве необратимо связывающегося агента в комбинации с MMTS [104]

Таблица 1. Использование алкилирующих агентов при решении различных задач протеомного анализа

1.11. Определение изомеров аминокислотных остатков в пептидах при помощи массспектрометрии

Как известно, изомеры химических соединений имеют одинаковую молекулярную массу. В связи с этим, определение конкретного изомера при помощи масс-спектрометрии, на первый взгляд, выглядит невозможной задачей. Тем не менее, современные масс-спектрометры измеряют не только массу прекурсорных ионов, соответствующую массе пептида, но и массу фрагментов, в результате чего получается масс-спектр второго порядка, или МС/МС. Разные изомеры одного и того же аминокислотного остатка способны по-разному влиять на устойчивость пептидной связи. В результате этого получающиеся фрагменты могут иметь разную интенсивность, и, следовательно, спектр фрагментации будет выглядеть по-разному. Подобное явление показано на нескольких парах изомеров аминокислотных остатков. Так наиболее изученными парами изомеров являются лейцин и изолейцин, аспарагиновая и изолейцин оба присутствуют в белках, в то время как формы аспарагиновой и, реже, глутаминовой кислоты являются продуктами дезаминирования, то есть часто встречаемой посттрансляционной модификации [112].

Было показано, что изолейцин образует ион иммония с m/z = 86. Это происходит при ионизации по типу бомбардировки быстрыми атомами (fast atom bombardment – FAB) [113]. Позже было описано различие в спектрах фрагментации лейцина и изолейцина при так называемой диссоциации с захватом горячих электронов (hot electron-capture dissociation – HECD) [114]. Оно заключается в неравной продукции ионов боковой цепи, а также *z*-ионов. В результате низкоэнергетической диссоциации, индуцированной соударением (low-energy collision-induced dissociation – CID) в качестве диагностического иона для различия лейцина и изолейцина был предложен вторичный продукт образования ионов иммония, а именно ион с m/z = 69 [115]. В дополнение к этому был предложен метод, сочетающий два вида ионизации: диссоциацию при переносе электрона (electron-transfer dissociation – ETD) и диссоциации, вызванной столкновением электронов при высокой энергии (higher-energy collisional dissociation – HCD) [116]. В этом методе в качестве диагностических ионов использовали *z*- и *w*-ионы, полученные на уровне MC³.

Немного другие подходы были использованы для того, чтобы распознавать по спектрам фрагментации аспарагиновую и изоаспарагиновую кислоты. Было показано, что при FAB в спектрах пептидов с Асп соотношение b_{n+1}/a_{n+1} (где n – это позиция Асп) больше, чем в спектрах аналогичных пептидов, содержащих изоАсп. Кроме того, ион иммония с m/z 88 является диагностическим, указывая на наличие изоАсп [117]. Позже было показано, что соотношение *y*-ионов к *b*-ионам при низкоэнергетической CID больше в спектрах пептидов,

содержащих изоАсп [118]. В данной работе также отмечают, что ион иммония является диагностическим. При диссоциации при захвате электрона (electron-capture dissociation – ECD) или ETD диагностическими ионами для различения Асп и изоАсп были названы ионы c_n+58 и z_n-57 [119–122]. Кроме различий в самих фрагментах, было показано, что при CID интенсивность ионов [M+Na]⁺ различается в шестичленных пептидах, содержащих Асп и изоАсп [123]. Существует несколько работ [119,122,124–126], использующих количественные методы при сравнении спектров фрагментации, получающиеся при использовании методов, основанных на реакциях электронов и ионов, таких как ECD [127] и ETD [128].

Задача определения α Глу и γ Глу при помощи масс-спектрометрии на первый взгляд выглядит аналогичной определению Асп и изоАсп. Однако этот вопрос намного хуже изучен и, по видимому, является более сложным. Для различия этих аминокислотных остатков также используют *z*- и *с*-ионы. При использовании ЕТD ион *z*–72 является диагностическим при определении наличия и положения γ Глу в пептидах, а ион *z*–72 – α Глу. Кроме того, если в непосредственной близости к N-концу пептида имеется носитель заряда, то в таких пептидах могут при ЕTD образовываться ионы с+57 и с+59, что также указывает на наличие в пептиде γ Глу [129].

Протеомный анализ является сложной, многоэтапной процедурой. Для получения значимых результатов нужно продуманно выбирать метод пробоподготовки и анализа данных. Из всей пробоподготовки наиболее уязвимым к появлению артефактов является этап защиты остатков цистеина. Работ, непосредственно сравнивающих эффект различных алкилирующих агентов не так много, а результаты в них иногда противоречат друг другу. Тем не менее, из них можно сделать вывод, что привычный большинству исследователей алкилирующий агент, йодацетамид подходит далеко не для всех задач протеомного анализа. Таким образом, задача подбора оптимального алкилирующего агента требует более внимательного изучения. Рациональный выбор алкилирующего агента особенно важен при исследовании посттрансляционных модификаций, а также в протеогеномных экспериментах, поскольку химические модификации аминокислотных остатков нередко совпадают по массе с посттрансляционными модификациями или с аминокислотными заменами. При анализе данных для выявления основных модификаций аминокислотных остатков следует применять метод открытого поиска, по результатам которого затем можно выбрать параметры основного поиска. В нашей работе мы провели сравнительный анализ алкилирующих агентов, оценили частоту некоторых модификаций метионина, а также провели открытый поиск для выявления других модификаций аминокислотных остатков.

ГЛАВА №2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Культура клеток

Клетки культур HeLa и HepG2 были получены из коллекции Института биомедицинской химии им. В.Н. Ореховича. Криогенно-консервированные клетки размораживали и высевали в культуральной среде (DMEM/F12) с добавлением 10% (мас./об.) фетальной бычьей сыворотки (FBS) и 100 ед./мл пенициллина/стрептомицина (все реактивы Gibco, Thermo Fisher Scientific, Германия) в увлажненном CO₂-инкубаторе в стандартных условиях (5% CO₂, 37°C). Среду меняли каждые 2 дня. При приготовлении образцов клеток для выделения белка клетки отделяли с помощью 0,25% раствора трипсина-ЭДТА (ПанЭко, Москва, Россия), трижды промывали PBS и подсчитывали. Аликвоты полученной клеточной суспензии центрифугировали, надосадочную жидкость удаляли и клетки замораживали в жидком азоте. Клеточные осадки хранили замороженными в парах жидкого азота до использования.

2.2. Подготовка образцов к масс-спектрометрическому анализу

2.2.1. Подготовка образцов бычьего сывороточного альбумина

Образцы бычьего сывороточного альбумина (BSA, Sigma-Aldrich, CША) были подготовлены для сравнения трех различных схем восстановления и алкилирования. Каждый образец содержал 50 мкг альбумина, растворенного в 50 мкл раствора 50 мМ хлорида тетраэтиламмония (TEABC; Sigma-Aldrich, CША). Затем к каждому образцу добавляли 5,5 мкл 5% (мас./об.) дезоксихолата натрия (SDC) в 50 мМ Трис-HCl, pH 8,5. После перемешивания, к каждому образцу добавляли 1,2 мкл 0,5 М дитиотреитола (DTT; Acros Organics, США) в 50мМ ТЕАВС до 10 мМ DTT. Образец №1 затем инкубировали в течение 20 минут при 56°С, тогда как образцы №2 и №3 выдерживали в течение 10 минут при 95°С. После этого, к каждому образцу добавляли 5,5 мкл 0,5 М раствора йодацетамида (IAM; Sigma-Aldrich, США) в 50 мМ ТЕАВС до конечной концентрации 50 мМ ІАМ. Образцы №1 и №2 инкубировали в течение 30 мин при комнатной температуре в темноте. Образец №3 инкубировали в течение 10 мин при 95°С, также в темноте. Затем белок расщепляли трипсином (Trypsin Gold, Promega, США). Фермент добавляли в соотношении 1:50 (мас./мас.), а полученную смесь инкубировали в течение ночи при 37°С. Реакцию трипсинолиза останавливали добавлением уксусной кислоты (5% мас./об.). Образцы встряхивали в течение 30 мин при 25°С (500 об./мин) и центрифугировали при 9300 g в течение 15 минут при 20°С (центрифуга 5415R, Eppendorf, Германия). Надосадочную жидкость перемещали на центрифужные фильтры (размер пор 30 кДа, Millipore, США) и центрифугировали при 13400 g в течение 20 минут при 20°С. После

этого в верхнюю камеру каждого фильтра добавляли по 200 мкл 50% (об./об.) муравьиной кислоты и образцы центрифугировали при 13400 g в течение 20 мин при 20°C до полного прохождения жидкости в нижнюю камеру. Образцы сушили в вакуумной центрифуге (Eppendorf Concentrator Plus, Eppendorf, Германия) при 45°C.

2.2.2. Подготовка образцов культур клеток

Лизис клеток и ферментативное расщепление белка были выполнены на основе метода, описанного ранее, с некоторыми изменениями [61]. Всю процедуру повторяли три раза, чтобы получить три повтора эксперимента. Две серии образцов с клетками линии HeLa и одну – с клетками линии HepG2 обрабатывали одинаковым образом. Клеточные осадки, содержащие 3,5 млн клеток, ресуспендировали в 170 мкл 0,1% детергента ProteaseMAX (Promega, CШA) в 50 мМ растворе бикарбоната аммония, смешанном с ацетонитрилом (ACN) в объемном соотношении 9:1. Затем смесь подвергали ультразвуковой обработке при помощи гомогенизатора Bandelin Sonopuls HD2070 (Bandelin Electronic, Германия) с интенсивностью 45% от максимальном, импульсами по 15 с, в течение 7 минут. Затем образцы центрифугировали при 15700 g в течение 20 минут при 20°С (центрифуга 5415R; Eppendorf, Германия) и собирали надосадочную жидкость. Общую концентрацию белка измеряли с использованием метода бицинхониновой кислоты (BCA Kit; Sigma-Aldrich, CШA). После этого образец разбавляли 50 мМ бикарбонатом аммония до концентрации белка 1 мг/мл. Получали восемь отдельных образцов по 50 мкл каждый.

После стадии алкилирования к каждому образцу добавляли трипсин (Promega Gold, Promega, США) в пропорции 1:50 (мас./мас.) фермента к общей массе белка. Образцы инкубировали в течение ночи при 37°С. Реакцию останавливали добавлением уксусной кислоты до 5% (об./об.). Затем образцы встряхивали в течение 10 минут при комнатной температуре и в течение 30 минут при 45°С. Надосадочную жидкость собирали и наносили на центрифужные фильтры с размером пор 30 кДа (Microcon Merc inc., Германия). Затем фильтры центрифугировали при 13400 g в течение 30 минут при 20°С. После этого в верхние камеры фильтров добавляли по 50 мкл 50% (об./об.) муравьиной кислоты и снова центрифугировали в тех же условиях. Затем образцы переносили в стеклянные емкости и высушивали в вакуумной центрифуге.

2.2.3. Синтез пептидов

Пептиды синтезировали в количестве около 0,5 мкмоль каждого с использованием автоматизированной рабочей станции для обработки жидкостей Hamilton Microlab Star
(Hamilton, Германия) с использованием лантерн SynPhas Lanterns (Mimotopes, Австралия) с тритильным линкером и присоединенными аминокислотами Fmoc-L-лизином или Fmoc-L-аргинином. Станция оснащена полипропиленовыми 96-луночными планшетами Arctic White LLC (PE Frit 25 мкм). Снятие защиты Fmoc проводили с использованием 20% (об./об.) 4-метилпиперидина в диметилформамиде (DMF) в течение 2 циклов по 10 мин. Связывание осуществляли с использованием 40 мМ раствора аминокислоты с соотношением 1:1:2 аминокислота/HCTU/TMP в DMF в течение 2 циклов по 30 мин. Аминокислоту добавляли с десятикратным избытком. Отщепление пептида проводили с использованием 94:1:2,5:2,5 TFA/TIS/вода/анизол в течение 2 часов. Готовые пептиды высушивали в вакуумной центрифуге.

2.3. Хроматомасс-спектрометрия

Анализ всех образцов проводили хроматомасс-спектрометрическим методом с небольшими различиями. Один микрограмм пептидов, разведенный в объеме 1-4 мкл воды/ацетонитрила/муравьиной кислоты в соотношениях 96/4/0,1% (об./об.), разделяли при помощи хроматографической системы UltiMate 3000 RSLC nano-system (Dionex, CША). Разделение проводили при помощи высокоэффективной обращенно-фазовой жидкостной хроматографии (ВЖЭХ) в обратном градиенте концентрации ацетонитрила. Сначала пептиды концентрировали и очищали на предколонке, непосредственно соединенной с аналитической колонкой. После разделения, пептиды сразу попадали в масс-спектрометр, соединенный с хроматографической системой «онлайн». Масс-спектры всех образцов получали при помощи масс-спектрометров типа Orbitrap из серии Q Exactive (Thermo Fisher Scientific, США) с использованием метода диссоциации, вызванной столкновением электронов при высокой энергии (higher-energy collisional dissociation – HCD), режима положительно заряженных ионов и зависимого от данных режима (data dependent acquisition – DDA).

2.3.1. Хроматомасс-спектрометрия образцов бычьего сывороточного альбумина

Для концентрирования и очистки пептидов в образцах BSA использовали предколонку Acclaim PepMap µ-Precolumn (300 мкм × 5 мм, размер частиц 5 мкм, величина пор 100 A, Thermo Fisher Scientific, США). Скорость потока при нанесении образца была 5 мкл/мин. Состав раствора для образцов был следующим: 4% (об./об.) CAN, 0,08% (об./об.) муравьиная кислота (FA), 0,015% трифторуксусная кислота (TFA). В качестве аналитической колонки использовали Acclaim C18 PepMap RSLC (150 мм × 75 мкм, размер частиц 2 мкм, величина пор 100 A, Thermo Fisher Scientific, США). Для разделения пептидов использовали 65-минутный

хроматографический метод. Скорость потока при разделении пептидов была 0,3 мл/мин. В качестве подвижной фазы A использовали: 0,015% TFA(об./об.), 0,08% FA (об./об.) в воде. В качестве подвижной фазы B использовали: 0,015% TFA(об./об.), 0,08% FA (об./об.), 80% ACN. Условия градиента были следующими: 0–3 мин, 4% раствор B; 3–28 мин, линейное увеличение до 55% раствора B; 28–31 мин линейное увеличение до 98% раствора B, 31–37 мин промывка колонки 98% раствора B, 37–40 мин линейное уменьшение до 4% раствора B; 40–47 мин уравновешивание колонки 4% раствора B.

Для масс-спектрометрического анализа образцов BSA использовали инструмент Q Exactive HF (Thermo Fisher Scientific, США). Для MC использовали разрешение 60000, таргетное значение AGC (Automatic Gain Control) 5×10^5 ионов и максимальное время накопления заряда 50 мс. Сканирование проводили в диапазоне от 400 до 1500 *m/z*. Для MC/MC разрешение было 15000, таргетное значение AGC 1×10^5 и максимальное время накопления заряда 100 мс. Сканирование проводили в диапазоне от 200 до 2000 *m/z*. Окно изоляции было 2 *m/z*, а максимальное число выбираемых ионов – 20. Использовали ступенчатую энергию соударения 35, порог интенсивности – 2×10^4 , заряды выбираемых ионов от +2 до +5 включительно.

2.3.2. Хроматомасс-спектрометрия образцов культур клеток

Один микрограмм пептидов в объемах от 1 до 4 мкл загружали на предколонку Acclaim µ-Precolumn (0,5 мм × 3 мм, размер частиц 5 мкм, Thermo Fisher Scientific, США) при скорости потока 10 мкл/мин. в течение 4 мин в изократическом режиме подвижной фазы С (2% ацетонитрила, 0,1% муравьиной кислоты). Затем пептиды разделяли с помощью высокоэффективной жидкостной хроматографии (ВЭЖХ, Ultimate 3000 Nano LC System, Thermo Fisher Scientific, CША) на колонке С18 длиной 15 см (Acclaim PepMap RSLC, внутренний диаметр 75 мкм, Thermo Fisher Scientific, США). Пептиды элюировали градиентом буфера В (80% ацетонитрила, 0,1% муравьиной кислоты) при скорости потока 0,3 мкл/мин. Общее время хроматографии составляло 90 минут, что включало начальные 4 минуты уравновешивания колонки с буфером А (0,1% муравьиной кислоты), затем градиент от 5 до 35% буфера В в течение 65 минут, затем 6 минут до достижения 99% буфера В, 10-минутное промывание 99% буфера В и 5-минутное уравновешивание в буфере А.

Масс-спектры получали в диапазоне 300-1500 *m/z* при разрешении 120000 (для MC). Тандемные масс-спектры получали с разрешением 15000 (MC/MC) в диапазоне от 100 *m/z* до значения *m/z*, определяемого по массе прекурсора, но не более 2000 *m/z*.

Поскольку три повтора эксперимента были сделаны в разное время, были использованы несколько разные условия масс-спектрометрического анализа. Для повторов I и

II использовали масс-спектрометр Q Exactive HF. Параметры, применяемые при анализе повтора I были следующими. Максимальное время накопление заряда составляло 50 мс и 100 мс для прекурсорных и фрагментов соответственно. Таргетное значение AGC для прекурсоров и фрагментов было установлено на 1×10^6 и 1×10^5 соответственно. Для изоляции прекурсорных ионов был выбран порог интенсивности в 5×10^5 . Для фрагментации выбирали до 15 ионов. Использовали ступенчатую энергию соударения (26, 28, 31). Для фрагментации отбирались прекурсоры в зарядном состоянии от +2 до +6. Окно изоляции составляло 4 *m*/*z*, а офсет 0,5 *m*/*z*. Динамическое исключение устанавливали автоматически на основании 15-секундной ширины хроматографического пика.

Для повтора II максимальное время накопления заряда составляло 50 мс и 110 мс для прекурсорных и фрагментных ионов соответственно. Таргетное значение AGC для прекурсоров и фрагментов было установлено на 1×10^6 и 1×10^5 соответственно. Для изоляции прекурсорных ионов был выбран порог интенсивности в $2,7 \times 10^5$. Для фрагментации выбирали до 25 ионов. Энергия соударения составляла 30 NCE. Для фрагментации отбирались прекурсоры в зарядном состоянии от +2 до +6. Окно изоляции составляло 2 *m/z*, а офсет 0,5 *m/z*. Динамическое исключение составляло 70 сек.

Для анализа повтора III использовали масс-спектрометр Q Exactive HF-X. Максимальное время накопления заряда составляло 50 мс и 150 мс для прекурсорных и фрагментных ионов соответственно. Таргетное значение AGC для прекурсоров и фрагментов было установлено на 1×10^6 и 1×10^5 соответственно. Для изоляции прекурсорных ионов был выбран порог интенсивности в $5,5 \times 10^5$. Для фрагментации выбирали до 20 ионов. Энергия соударения составляла 33 NCE. Для фрагментации отбирались прекурсоры в зарядном состоянии от +2 до +5. Окно изоляции составляло 2 *m/z*, а офсет 0,5 *m/z*. Динамическое исключение составляло 20 сек.

Все полученные в настоящем разделе методов хроматомасс-спектры доступны для публичного скачивания и использования на pecypce ProteomeXchange (<u>www.proteomexchange.org</u>) [130] под инвентарным номером PXD015889.

2.3.3. Масс-спектрометрия синтетических пептидов

Пептиды ресуспендировали в 0,1% TFA в воде, полученное разведение составляло 1:1000 перед масс-спектрометрическим анализом. Для анализа спектров пептидов, содержащих треонин и изотреонин были подготовлены смеси этих пептидов (подробнее в главе Результаты и обсуждение). Такие образцы также анализировали с использованием обращенно-фазовой высокоэффективной жидкостной хроматографии, совмещенной с масс-спектромертией (ВЭЖХ-МС). В качестве предколонки использовали колонку Acclaim PepMap длиной 2 см (внутренний

диаметр 75 мкм, размер частиц 3 мкм, величина пор 100 А; Thermo Scientific, USA).Скорость потока при нанесении образца на предколонку составляла 2 мкл/мин. Через 5 минут предколонку соединяли с аналитической колонкой Zorbax 300SB-C18 длиной 15 см (внутренний диаметр 75 мкм, размер частиц 3,5 мкм; Agilent, США). Раствор подвижной фазы А представлял собой воду с 0,1% (об./об.) муравьиной кислотой, а раствор подвижной фазы В – 80% (об./об.) ацетонитрил/вода с 0,1% (об./об.) муравьиной кислотой. Разделение проводили с использованием линейного градиента раствора В от 5% до 40% при скорости потока 300 нл/мин в течение 50 минут с последующей стадией промывки (10 минут при 99% раствора В) и стадией уравновешивания (10 минут при 5% раствора В).

Масс-спектрометрический анализ проводили при помощи прибора Q Exactive HF (Thermo Fisher Scientific, США). Использовался источник ионизации Nanospray Flex с напряжением ионизации 1800 В и температурой капилляра 200°С. Для MC использовали разрешение 60000, таргетное значение AGC (Automatic Gain Control) 5×10^5 ионов и максимальное время подачи образца 50 мс. Сканирование проводили в диапазоне от 400 до 1500 *m/z*. Для MC/MC разрешение было 15000, таргетное значение AGC 1×10^5 и максимальное время накопления заряда 100 мс. Сканирование проводили в диапазоне от 200 до 2000 *m/z*. Окно изоляции было 2 *m/z*, а максимальное число выбираемых ионов – 20. Использовали энергию соударения 35, порог интенсивности – 2×10^4 , заряды выбираемых ионов от +2 до +5 включительно.

Для анализа масс-спектров клеточных линий панели NCI60, взятых из работы [131] на предмет наличия в них замены метионина на изотреонин были использованы синтетические аналоги пептида 57(NQVAMNPTNTVFDAK)71 в белке HSC70. Пептиды с последовательностями NQVAMNPTNTVFDAK и NQVAMNPisoTNTVFDAK были заказаны у JPT Peptides (Германия). Пептиды разводили до концентрации 1 мг/л в 4% растворе ACN в воде с добавлением 0,1% муравьиной кислоты. Полученные образцы анализировали по тому же методу, что и образцы BSA.

2.4. Анализ данных

2.4.1.Протеогеномный поиск при помощи программы X!Tandem

Данные протеомов клеточных линий панели NCI60 были загружены с http://wzw.tum.de/proteomics/nci60 [131]. Для протеогеномного поиска была создана база данных, содержащая вариантные белки с одноаминокислотными заменами, характерными для колоректального рака. Объединенная база данных включала полную базу данных протеома человека из UniProt (выпуск от января 2013 года; 87638 белков) и белковую базу данных колоректального рака, сгенерированную из доступных геномных данных [132] (127486 белков).

Протеогеномный поиск проводили при помощи поисковой программы X!Tandem [133] версии 2012.10.01.1 с последующей обработкой результатов при помощи инструмента MPScore [42]. Для поиска использовали стратегию «раздельного FDR» [46]. Допустимая ошибка массы для прекурсорных и фрагментных ионов была 15 ppm и 0,3 Да соответственно. В качестве постоянной модификации использовали карбамидометилирование цистеинов, а в качестве вариабельной – ацетилирование N-конца. В качестве гидролизующего фермента был указан трипсин, а количество допустимых пропусков сайтов гидролиза – 2. Результаты фильтровали до 1% FDR при помощи MPscore на уровне PSM.

2.4.2. Количественный анализ без использования метки при помощи программы MaxQuant

Для определения относительного содержания пептидов применяли количественную оценку без использования меток на основе интенсивности. Полученные при помощи массспектрометра файлы формата RAW анализировали программой MaxQuant версии 1.5.2.8 [134]. Анализ MaxQuant включал первоначальный поиск с допустимой ошибкой массы прекурсора 20 ppm. Результаты первого поиска были использованы для повторной калибровки массы. В основном поиске допустимые ошибки масс прекурсорных и фрагментных ионов были установлены равными 6 ppm и 20 ppm соответственно. Поиск включал вариабельные модификации, такие как окисление метионина, дезаминирование аспарагина и глутамина и Nконцевое ацетилирование, а также карбамидометилцистеин в качестве фиксированной модификации. Минимальная длина пептида была установлена на семь аминокислот, и было разрешено максимум два пропущенных сайта гидролиза. Результаты фильтровали до 1% FDR на пептидном и белковом уровне автоматически при помощи MaxQuant.

2.4.3. Протеомный поиск при помощи программы IdentiPy

Файлы формата RAW с масс-спектрами были преобразованы в формат MGF с помощью MSConvert 3.0.11781 из пакета ProteoWizard [135]. Далее производили поиск полученных MGF файлов против базы данных SwissProt *Homo sapiens* (версия ноябрь 2018, 20409 белков) для наборов данных протеомов человека и против базы данных SwissProt *Mus musculus* (версия от февраля 2020 г., 17040 белков) для набора данных протеомов мышей, загруженных с www.uniprot.org. Поиск проводился с использованием поисковой системы IdentiPy (коммит 600c2a776499) [136] и постобработки данных MPScore [42]. Используемые

параметры поиска были следующими: трипсин в качестве фермента гидролиза; допустимо до 1 пропущенного сайта расщепления; допустимая ошибка массы прекурсора 10 ppm; точность массы фрагментных ионов 0,01 Да и 0,3 Да для данных MC/MC, полученных с использованием масс-анализаторов Orbitrap и Ion Trap соответственно; диапазон длин пептидов от 5 до 50 остатков; в качестве фиксированной модификации – специфическая для образца модификация цистеина и окисление метионина в качестве вариабельной модификации. База данных была дополнена декойными последовательностями, сгенерированными по реверсивному принципу. Для анализа общего числа идентифицированных пептидов и PSM результаты фильтровали автоматически при помощи MPScore, одновременно с поиском до 1% FDR на соответствующем уровне. Фиксированные модификации цистеина были следующими: для образцов № 60, 68, 92: никакие; № 61, 62, 65, 66, 69, 70, 73, 74, 93, 94, 97, 98: карбамидометилирование (+57,021 Да); № 63, 64, 71, 72, 95, 96: пиридилэтилирование (+105,058 Да), № 67, 75, 99: сульфенилирование (+45,988 Да).

После протеомного поиска мы сравнивали общее число идентифицированных пептидов и спектральных идентификаций (PSM), а также абсолютное и относительное число пептидов и PSM с Цис и Мет между всеми образцами и тремя повторами эксперимента. Чтобы получить более воспроизводимые результаты между повторами, мы нормализовали значения в каждом повторе эксперимента (повторы I, II и III) по наивысшему значению среди всех образцов (т. е. методов алкилирования). После этого была проведена оценка критерия Стьюдента попарно среди всех комбинаций образцов. Помимо значений р критерия Стьюдента, мы вычислили кратное изменение всех нормализованных значений в парах сравнения. Здесь мы используем термин «кратное изменение» как отношение большего значения к меньшему значению среднего среди всех повторов в одном образце. Статистический анализ проводился с использованием скриптов, написанных на R (3.6.3).

2.4.4. Анализ частоты преобразования Мет в изоТре в панорамных протеомных данных

Для того, чтобы оценить, какое количество Мет превращается в изоТре в протеомных экспериментах, мы использовали общедоступные данные из трех различных работ [7,131,137], а так же наши собственные данные, полученные в этой работе.

Для поиска использовали файлы MGF, поиск проводили при помощи IdentiPy против базы данных SwissProt соответствующего организма. Помимо стандартных параметров поиска, в качестве вариабельной модификации была добавлена замена метионина на треонин, по сдвигу массы совпадающая с преобразованием метионина в изотреонин (-29,9928 Да). Затем пептиды с идентифицированной модификацией Мет->Тре были отфильтрованы до группоспецифичного FDR 1% согласно [47] с использованием скриптов, написанных на языке Python на базе

библиотеки Pyteomics (v. 4.1.2) [138,139]. Такую же фильтровку проводили для всех пептидов содержащих метионин. После этого, сравнивали долю пептидов с Мет, несущих модификацию, между образцами, используя статистический критерий X². Соответствующие скрипты для статистического анализа были написаны на языке R.

2.4.5. Оценка влияния соседствующих аминокислотных остатков на частоту преобразования Мет в изоТре

Воспользовавшись тем фактом, что все образцы в [131] были приготовлены одинаковым методом, мы объединили данные глубоких протеомов 9 клеточных линий из панели NCI60 и проанализировали влияние соседних аминокислотных остатков на частоту преобразования Мет-изоТре. Для этого мы провели поиск спектров в формате RAW при помощи MaxQuant (v. 1.6.1.0) [40]. Затем файлы «Met- ThrSites.txt» и «peptides.txt», полученные в ходе этих поисков, использовали для расчета частоты модификации в зависимости от предшествующего и последующего аминокислотных остатков в последовательности. Частоту модификации оценивали как отношение количества преобразований Мет-изоТре, предшествующих или следующих за каждой аминокислотой во всех идентифицированных пептидах, к количеству всех комбинаций метионина с соответствующим аминокислотным остатком в идентифицированных пептидах. Например, это соотношение для аланина, предшествующего метионину, рассчитывали как количество модифицированного метионина, следующего за аланином, деленное на количество всех комбинаций АлаМет в идентифицированных пептидах. Эти соотношения сравнивали с использованием критерия X² таким же образом, как и при общей оценке частоты преобразования Мет в изоТре.

2.4.6. Открытый поиск и анализ с использованием программы AA_stat

Для того, чтобы убедиться, что наиболее частые сдвиги масс соответствуют ожидаемым модификациям и проверить, не больше ли их, чем ожидалось, мы выполнили открытый поиск [87] с последующим анализом результатов с помощью инструмента AA_stat [108]. Открытый поиск проводили с помощью MSFragger (версия 20170103.0) [88] со следующими параметрами: допустимая ошибка масс прекурсоров ± 500 Да, допустимая ошибка масс фрагментов 0,03 Да. AA_stat применялся для анализа результатов поиска с параметрами по умолчанию. Мы провели два раунда открытого поиска и анализа AA_stat. Первый поиск был необходим для подтверждения наличия модификаций, вызванных алкилированием остатков цистеина. Для первого поиска не было задано никаких вариабельных или фиксированных

модификаций. Затем был проведен второй поиск с фиксированными модификациями, соответствующими используемому методу алкилирования, чтобы увидеть другие модификации других аминокислотных остатков. Открытый поиск был проведен для всех трех наборов данных, использованных в этой работе: наших собственных экспериментальных данных и общедоступных данных для девяти клеточных линий NCI60 [131] и трех образцов мозга мышей [137].

2.4.7. Анализ масс-спектров синтетических пептидов

Файлы формата RAW, полученные при масс-спектрометрическом анализе синтетических пептидов сначала просматривали вручную при помощи программы Xcalibur Qual Browser 2.2 SP1.48 (Thermo Fisher Scientific, CША). Затем данные обрабатывали с помощью MaxQuant 1.5.2.8. Поиск выполнялся против специально сконструированной базы данных, содержащей последовательности как из группы человеческих пептидов, так и из группы модельных пептидов, перечисленных в таблице 2 (подробнее о них будет рассказано далее).

	пептиды с Тре	пептиды с изоТре	число Тре или изоТре в
			пептиде
	DNIQGITKPAIR	DNIQGI(isoT)KPAIR	1
	TSYAQHQQVR	(isoT)SYAQHQQVR	1
	QDPSVLHTEEMR	QDPSVLH(isoT)EEMR	1
	NNRPSEGPLQTR	NNRPSEGPLQ(isoT)R	1
	KTVGVEPAADGK	K(isoT)VGVEPAADGK	1
ДЫ	MSSPEDDSDTKR	MSSPEDDSD(isoT)KR	1
ити	KGESQTDIETR	KGESQ(isoT)DIEI(isoT)R	2
еческие пе	TGHSLLHTLYGR	(isoT)GHSLLH(isoT)LYGR	2
	KPLLESGTLGTK	KPLLESG(isoT)LG(isoT)K	2
	GAGTDDHTLIR	GAG(isoT)DDH(isoT)LIR	2
ЭЛОЕ	DRVTDALNATR	DRV(isoT)DALNA(isoT)R	2
ЭҺ	TLHPDLGTDKDK	(isoT)LHPDLG(isoT)DKDK	2
	TGNFQVTELGR	(isoT)GNFQV(isoT)ELGR	2
	TQPMTAQAASYR	(isoT)QPM(isoT)AQAASYR	2
	AVANQTSATFLR	AVANQ(isoT)SA(isoT)FLR	2
	VVTDTDETELAR	VV(isoT)D(isoT)DE(isoT)ELAR	3
	HFVALSTNTTK	HFVALS(isoT)N(isoT)(isoT)K	3
	VVPTTDHIDTEK	VVP(isoT)(isoT)DHID(isoT)EK	3
	TMQNTSDLDTAR	(isoT)MQN(isoT)SDLD(isoT)AR	3

Таблица 2. Список последовательностей синтетических пептидов использованных для сравнения масс-спектров пептидов с треонином и изотреонином

	STSTPTSPGPR	S(isoT)S(isoT)P(isoT)SPGPR	3
	FIDTTLATSR	FID(isoT)(isoT)LAI(isoT)SR	3
	YIGTGHADTTK	YIG(isoT)GHAD(isoT)(isoT)K	3
	FFNVLTTNTDGK	FFNVL(isoT)(isoT)N(isoT)DGK	3
	ETVEEQASTTER	E(isoT)VEEQAS(isoT)(isoT)ER	3
	TPYTDVNIVTIR	(isoT)PY(isoT)DVNIV(isoT)IR	3
	TGYTLDVTTGQR	(isoT)GY(isoT)LDV(isoT)(isoT)GQR	4
	HADIVTTTTHK	HADIV(isoT)(isoT)(isoT)(isoT)HK	4
	TALTTTISSR	(isoT)AL(isoT)(isoT)(isoT)ISSR	4
	YRDPTTVTTLR	YRDP(isoT)(isoT)V(isoT)(isoT)LR	4
	TDTLEDLFPTTK	(isoT)D(isoT)LEDLFP(isoT)(isoT)K	4
	NTLTQTTENLR	N(isoT)L(isoT)Q(isoT)(isoT)ENLR	4
	ETTVGAVTVTHK	E(isoT)(isoT)VGAV(isoT)V(isoT)HK	4
	TTGIVETHFTFK	(isoT)(isoT)GIVE(isoT)HF(isoT)FK	4
	QLYETTDTTTR	QLYE(isoT)(isoT)D(isoT)(isoT)(isoT)R	5
	LWDLTTGTTTR	LWDL(isoT)(isoT)G(isoT)(isoT)(isoT)R	5
	TVEQTATTTNK	(isoT)VEQ(isoT)A(isoT)(isoT)NK	5
	PAATAAGSLGSLK	PAA(isoT)AAGSLGSLK	1
	PDDTDDGSLGSLK	PDD(isoT)DDGSLGSLK	1
	PEETEEGSLGSLK	PEE(isoT)EEGSLGSLK	1
	PFFTFFGSLGSLK	PFF(isoT)FFGSLGSLK	1
I	PGGTGGGSLGSLK	PGG(isoT)GGGSLGSLK	1
пдил	PMMTMMGSLGSLK	PMM(isoT)MMGSLGSLK	1
пепт	PQQTQQGSLGSLK	PQQ(isoT)QQGSLGSLK	1
bie 1	PVVTVVGSLGSLK	PVV(isoT)VVGSLGSLK	1
лын	PIITIIGSLGSLK	PII(isoT)IIGSLGSLK	1
оде	PLLTLLGSLGSLK	PLL(isoT)LLGSLGSLK	1
W	PNNTNNGSLGSLK	PNN(isoT)NNGSLGSLK	1
	PPPTPPGSLGSLK	PPP(isoT)PPGSLGSLK	1
	PSSTSSGSLGSLK	PSS(isoT)SSGSLGSLK	1
	PTTTTTGSLGSLK	PTT(isoT)TTGSLGSLK	1
	PYYTYYGSLGSLK	PYY(isoT)YYGSLGSLK	1

2.4.8. Машинное обучение

Мы провели автоматическое распознавание образов для того, чтобы определить, возможно ли однозначно отличить пептиды содержащие треонин от пептидов содержащих изотреонин используя относительную интенсивность фрагментных ионов, связанных с треонином, в спектрах MC/MC. В данный анализ включали только пептиды, содержащие один остаток Тре или изоТре соответственно. Таким образом, список пептидов, используемых в автоматической классификации, был сокращен и содержал 21 пептид, содержащий Тре и 21 пептид, содержащий изоТре. Точность распознавания пептидов с Тре и изоТре оценивали путем усреднения результатов 100 запусков 5-кратной процедуры перекрестной проверки с использованием классификатора Random Forest – библиотеки Recursive Partitioning and Regression Trees (RPART) для R версии 4.1-10. Мы рассчитали общую точность для каждой возможной комбинации шести признаков. Весь этот анализ проводили при помощи скриптов, написанных на языке R.

ГЛАВА №3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1. Замена метионина в 61 позиции в белке HSC70 произошла в 8 из 9 клеточных линиях панели NCI60

В работе [131] представлены данные для клеточных линий панели американского Национального института рака NCI60. Для 9 таких линий, а именно: MCF7 (рак молочной железы), SK-OV-3 (рак яичника), U251 (злокачественная глиобластома), RXF-393 (рак почки), COLO205 (рак прямой кишки), NCI-H460 (рак легких), M14 (злокачественная меланома), ССRF-СЕМ (острый лимфоцитарный лейкоз) и PC-3 (рак предстательной железы), авторы предоставили данные «глубокого» протеома, с идентификацией нескольких тысяч генных продуктов. Мы использовали эти данные в протеогеномном анализе. Протеогеномный поиск проводили против человеческой базы данных, дополненной последовательностями, содержащими одноаминокислотные замены, характерные для раковых клеток. Последовательности белков с заменами были транслированы из экзомов рака прямой кишки, полученных в опубликованных данных [132]. Поиск и расчет FDR проводили по принципу "раздельного FDR" или подхода "один за другим" (см. Обзор литературы) [35,46]. Важно отметить, что используемая база данных представляла собой протеомную базу, содержащую участки полиморфизма из злокачественных опухолей толстой и прямой кишки, однако, она не имела точного совпадения с транслированными экзомами какой-либо из исследованных раковых клеточных линий, даже COLO205, имевшей сходное происхождение. Широко известно, что несмотря на то, что точечные раковые мутации обычно аккумулируются в одних и тех же генах, называемых онкогенами и онкосупрессоры, такие мутации существенно различаются у разных пациентов [140].

Отметим, что данный протеогеномный поиск не имел клинического значения и представлял собой отладку протеогеномного вычислительного конвейера, использованного в дальнейших работах.

Тем не менее, поиск показал, что в пептиде 57(NQVATNPTNTVFDAK)71 из конститутивного белка теплового шока HSC70 метионин в положении 61 заменен на треонин.

Это произошло в 8 из 9 исследуемых клеточных линиях. Более того, в 6 линиях был обнаружен пептид 57(NQVATNPTNTVFDAKR)72 то есть соответствующий исходному аналог, содержащий пропущенный сайт трипсинолиза, что дополнительно подтверждало достоверность обнаруженной находки.

3.2. Замена метионина в положении 61 белка HSC70 не закодирована в геноме

Замена Мет в позиции 61 в HSC70 на Тре встречается крайне редко и была обнаружена лишь у одного пациента с раком прямой кишки из более двухсот проанализированных в статье [131]. Однако, замена Мет на Тре встречается и в других белках и может являться причиной серьезных заболеваний [141]. Вскоре после того, как мы обнаружили данную замену в протеомах клеточных линий NCI60, в открытом доступе появились данные экзомов этих же клеточных линий [140]. Мы воспользовались этими данными для уточнения присутствия интересующей нас замены в геномах конкретных линий в соответствующей позиции. Результат такой проверки оказался отрицательным: нам не удалось обнаружить замены кодона метионина на кодон треонина в этих данных. Более того, как показал анализ методом BLAST, такой пептид не был обнаружен ни в одном живом организме. Таким образом, мы исключили возможность контаминации образцов чужеродной клеточной линией в процессе проведения клеточных экспериментов. Помимо геномного происхождения одноаминокислотной замены, возможным объяснением данного феномена могло послужить редактирование РНК. Однако для получения замены Мет на Тре должно было бы происходить редактирование кодона АУГ на АЦГ. Известно, что фермент ADAR не может осуществлять редактирование У на Ц, а, напротив, производит дезаминирование, преобразуя Ц в У [142]. Также не описано других видов редактирования РНК, способных осуществить такую замену. Таким образом, был сделан вывод, что преобразование метионина в треонин не закодировано в нуклеиновых кислотах, а произошло либо на этапе посттрансляционных модификаций, либо во время пробоподготовки к протеомному анализу.

3.3. Гипотеза о преобразовании метионина в изотреонин в процессе пробоподготовки

Разбираясь в возможных вариантах происхождения обнаруженной модификации метионина на этапе пробоподготовки, мы пришли к выводу, что преобразование метионина могло произойти не в треонин, а в изотреонин, то есть гомосерин. Подобная реакция происходит при обработке белков бромцианом. Она приводит к образованию из метионинов гомосерина или лактона гомосерина в зависимости от pH среды [143]. В более кислой среде и при более высокой концентрации бромциана полипептидная цепь рвется и метионин

преобразуется в лактон гомосерина. В более мягких условиях он может преобразоваться в гомосерин (изотреонин) без разрыва полипептидной цепи [144]. В базе данных модификаций аминокислотных остатков Unimod (www.unimod.org) присутствуют обе модификации под номерами 10 для гомосерина и 11 – для лактона гомосерина.

Известно, что йодуксусная кислота, как и йодацетамид способны алкилировать метионин с образованием карбоксиметилметионина и карбамидометилметионина соответственно [145]. Оба соединения под действием высокой температуры (около 100°С) в нейтральной среде способны преобразовываться в гомосерин и в лактон гомосерина [110]. Поскольку при проведении гель-электрофореза перед трипсинолизом образцы часто нагревают, а именно такой метод был применен в рассматриваемой работе [131], мы предположили, что преобразование метионина с получением изотреонина в анализируемых нами данных вполне возможно.

3.4. Анализ масс-спектров синтетических стандартов пептида 57(NQVATNPTNTVFDAK)71 и его варианта с изотреонином

Для того, чтобы убедиться, что в анализируемых нами данных протеомов клеточных линий NCI60 в пептиде 57(NQVAMNPTNTVFDAK)71 белка HSC70 метионин в 61 позиции преобразован именно в изотреонин, а не заменен на треонин, мы провели анализ с использованием синтетических стандартов. Были получены масс-спектры пептидов с последовательностями NQVATNPTNTVFDAK и NQVA(isoT)NPTNTVFDAK. При визуальном сравнении этих спектров с экстрагированным спектром соответствующего пептида из данных по клеточным линиям, видно, что последние больше похожи на спектр пептида с изоТре (см. рис. 9). В обоих случаях фрагмент b_6 с m/z 628,305 имеет заметно бо́льшую интенсивность, чем в спектре пептида с Tpe. Фрагмент b_6 в исследуемом пептиде имеет последовательность NQVA(T/isoT)N⁺, то есть является пептидом b_{n+1} , где n – положение треонина или изотреонина. В спектре синтетического пептида с изоТре интенсивность фрагмента b_6 составляет в среднем 3%. В спектрах клеточных линий такая интенсивность составляет в среднем 18%.

А. Масс-спектр из клеточной линии CNS_U251







Рис. 9. Масс-спектры пептида NQVATNPTNTVFDAK из клеточной линии CNS_U251 (A) и его синтетических аналогов с треонином (Б) и изотреонином (В). Интенсивность иона b₆ в спектре из клеточной линии сравнима с таковой в спектре пептида с изотреонином.

3.5. Количественный анализ пептида 57(NQVAMNPTNTVFDAK)71 и его варианта с изотреонином в панорамных протеомных данных клеточных линий панели NCI60

Далее мы проанализировали наличие немодифицированного варианта заинтересовавшего нас пептида в данных клеточных линий NCI60, а именно, пептида 57(NQVAMNPTNTVFDAK)71 в белке HSC70 в тех восьми линиях, в которых был обнаружен пептид с заменой Мет на изоТре. В результате такой немодифицированный пептид был обнаружен во всех этих восьми клеточных линиях. Более того, во всех них был обнаружен его аналог, содержащий один пропущенный сайт трипсинолиза, а именно пептид 57(NQVAMNPTNTVFDAKR)72. Количественный анализ без использования метки (label free quantification, LFQ) показал, что относительное содержание модифицированных вариантов пептидов меньше. Средняя интенсивность (LFQ intensity) пика, рассчитанная программой MaxQuant на основании интенсивности площадей хроматографических пиков в MC¹ (XIC) [146], для модифицированного пептида составляет примерно 2,3% от интенсивности для не модифицированного пептида (стандартное отклонение 1,3%, доверительные интервалы 1,1-4,3%) (Рис. 10).





3.6. Моделирование протеомной пробоподготовки на бычьем сывороточном альбумине для предсказания преобразования Мет в изоТре

Для того, чтобы внести больше ясности в причины преобразования метионина в изотреонин и понять, на каком именно этапе пробоподготовки это происходит, мы провели эксперименты на бычьем сывороточном альбумине (BSA). Для воссоздания условий стандартного протеомного эксперимента, альбумин сначала обработали дезоксихолатом натрия. Далее были протестированы три схемы восстановления и алкилирования, отличавшиеся температурой инкубации. Целью такой работы была оценка возможной роли нагревания в преобразовании метионина в изотреонин. В качестве восстанавливающего агента применяли дитиотреитол (DTT), а в качестве алкилирующего – йодацетамид (IAM). Образец №1 был восстановлен и проалкилирован при комнатной температуре. Образец №2 был нагрет до 95°С перед алкилированием, а образец №3 был нагрет до 95°С непосредственно во время алкилирования.

Последовательность BSA содержит пять остатков метионина. Метионины в положениях 1 и 208 практически не имеют шансов быть детектированы в ходе протеомного анализа. Мет-1 отщепляется в ходе созревания белка. Кроме этого, за ним следует аргинин и, так или иначе, после трипсинолиза образовывался бы пептид, состоящий всего из двух аминокислотных остатков, который не детектируется масс-спектрометром при стандартном анализе. Мет-208 входит в состав триптического пептида, состоящего из пяти аминокслотных остатков, что так же ниже порога детектирования. Такими образом, для анализа были доступны оставшиеся три остатка метионина в положениях 111, 469 и 571. Ожидаемо, нам удалось обнаружить пептиды, содержащие преобразованные метионины во всех трех позициях. Примеры спектров фрагментации модифицированного и немодифицированного пептидов представлены на рис. 11. Результаты данного анализа представлены в таблице 3 в виде числа спектральных идентификаций (PSM) пептидов, содержащих метионин во всех трех позициях, пептидов, содержащих окисленный метионин и метионин, преобразованный в изотреонин. Без нагревания образца количество PSM с метионином, преобразованным в изотреонин составило 2-8%. Однако нагревание образца во время алкилирования увеличило содержание таких PSM до 37% и сделало его соизмеримым с количеством PSM, содержащими окисленный метионин. Интересно, что метионин в положении 571 подвержен модификации в изотреонин в меньшей степени, чем остальные два. То же наблюдается и с окислением этого же остатка. По всей видимости, конформация белка в этой области молекулы делает остаток метионина менее уязвимым к модификациям.





Рис. 10. Масс-спектры прекурсорных ионов пептидов бычьего сывороточного альбумина. А: ЕТҮGDMADCEEK содержащий метионин, Б: ЕТҮGD(isoT)ADCEEK содержащий метионин, преобразованный в изотреонин.

Таблица 3. Спектральные идентификации (PSM) метионинсодержащих пептидов бычьего сывороточного альбумина. Показано общее число PSM, абсолютное и относительное число PSM с окисленным метионином и абсолютное и относительное число PSM с метионином, преобразованным в изотреонин, в образцах, подвергавшихся алкилированию при разных условиях. Образец №1 был восстановлен и проалкилирован при комнатной температуре. Образец №2 был нагрет до 95°С перед алкилированием, а образец №3 был нагрет до 95°С непосредственно во время алкилирования.

позиция Met в BSA	последовательность пептида	общее количество PSM		PSM с окисленным Мет (% от общего)			PSM с Мет-изоТре (% от общего)			
		№ 1	№ 2	Nº 3	Nº 1	Nº 2	Nº 3	Nº 1	№ 2	№ 3
M111	ETYGDMADCCEK	50	46	53	8 (16%)	14 (30%)	11 (21%)	4 (8%)	4 (9%)	10 (19%)
M469	MPCTEDYLSLILNR	159	183	268	25 (16%)	34 (19%)	52 (19%)	6 (4%)	6 (3%)	98 (37%)
M571	TV M ENFVAFVDK	124	113	95	6 (5%)	3 (3%)	5 (5%)	3 (2%)	3 (3%)	5 (5%)

Для всех трех пептидов обнаружены также их варианты с пропущенными сайтами гидролиза.

В целом, результаты эксперимента с BSA подтверждают предположение о том, что метионин преобразуется в изотреонин, и происходит это на этапе алкилирования в процессе пробоподготовки к протеомному анализу.

3.7. Анализ масс-спектров синтетических пептидов, содержащих Тре и изоТре

Предполагая, что метионин во время пробоподготовки модифицируется именно в изотреонин, а не заменяется на треонин, как в случае генетически закодированного полиморфизма, мы задались вопросом, можно ли отличить эти два аминокислотных остатка по масс-спектрам. Дело в том, что треонин и изотреонин имеют одинаковую молекулярную массу и, казалось бы, при помощи масс-спектрометрии должны быть неразличимы. Однако распознавание изомеров аминокислот по спектрам фрагментации проводили на других изомерных аминокислотах [112,129,147], о чем было подробно рассказано выше.

Для решения нашей задачи были синтезированы пептиды, по последовательности аминокислотных остатков соответствующие человеческим пептидам, образующимся в результате трипсинолиза. Выбранные пептиды содержали один, два, три, четыре или пять остатков треонина. Помимо этого, были синтезированы аналоги этих же пептидов, но содержащие изоТре на месте Тре. Всего таких пар было 36 (см. таблицу 2). В дополнение к человеческим пептидам были синтезированы 15 пар «модельных» пептидов. Такие пептиды имели схожие последовательности, но с различными аминокислотными остатками, окружающими треонин (и изотреонин соответственно). Каждый пептид содержал только один остаток Тре (или изоТре), окруженный с двух сторон двумя одинаковыми остатками другой аминокислоты (см. таблицу 2). Далее были детально изучены хроматограммы и спектры фрагментации таких пептидов при помощи HCD и масс-анализатора типа Orbitrap. С этой целью были приготовлены следующие образцы, содержащие смеси пептидов. Для человеческих пептидов каждый образец содержал 9 пептидов с Тре и изоТре по отдельности и 18 пептидов в смешанных образцах (всего 12 образцов). Для модельных пептидов образцы содержали все 15 пептидов с Тре и изоТре по отдельности и 30 пептидов в смеси (всего 3 образца). Все 15 образцов были проанализированы в 3 технических повторах.

3.7.1. Синтетические аналоги с треонином и изотреонином различаются по времени выхода с хроматографической колонки

На первом этапе нам удалось разделить смесь, содержащую аналогичные пептиды с Тре и изоТре при помощи жидкостной хроматографии. При сравнении хроматограмм смеси

аналогичных пептидов с хроматограммами образцов, содержащих эти пептиды по отдельности, отчетливо видно, что время выхода пептидов с изоТре меньше, чем пептидов с Тре (рис. 11). Средняя разница во времени выхода пептидов с Тре и изоТре составила 1.26 минуты, что в данных условиях градиента соответствует разнице концентраций ацетонитрила (ACN) в подвижной фазе, равной 7,37%. Кроме того, мы заметили корреляцию между количеством остатков треонина в пептиде и разницей во времени выхода такого пептида с его аналогом, содержащим изоТре. Иными словами, чем больше остатков треонина или изотреонина в пептиде, тем лучше такие пептиды разделяются при помощи хроматографии. Поскольку выборка в нашем эксперименте сравнительно невелика, для оценки статистической значимости этого наблюдения, мы разделили пептиды только на две группы. В первую группу вошли пептиды, содержащие один остаток треонина, во вторую – все остальные пептиды, то есть содержащие два и более остатков. Ожидаемо, разница времен выхода пептидов с Тре и изоТре в первой группе была ощутимо меньше, чем во второй. Для первой группы эти значения составили в среднем 0,77 минут, что соответствует 6,92% ACN, а для второй – 1,58 минут, что соответствует 7,36% АСЛ. В действительности, разделение пептидов, содержащих изомерные аминокислотные остатки, было описано ранее [122,148,149]. Кроме того, было показано, что влияние наличия лейцина или изолейцина в пептиде может быть предсказано с использованием феноменологической теории разделения пептидов [150].



Рис.11. Хроматограммы образцов, содержащих пептиды с Тре и изоТре по отдельности и смесь пептидов с Тре и изоТре. Пептиды с изоТре имеют меньшее время выхода с хроматографической колонки.

Для исследуемых пептидов мы рассчитали предсказанные времена выхода с колонки при помощи модели, описанной в [151]. Результаты показаны в таблице 4. Из этих результатов видно, что выход пептидов с изоТре с колонки происходит всегда раньше, чем пептидов с Тре. Предсказание времени выхода недостаточно точно для того, чтобы нормализовать эти данные и подвергнуть их автоматической классификации. Разница между реальными и предсказанными временами выхода одних и тех же пептидов больше, чем разница времен выхода аналогов, содержащих Тре и изоТре. Таким образом, при помощи хроматографии остатки Тре в пептидах возможно различить с остатками изоТре только в том случае, если изучать аналогичные пары пептидов.

последовательность	ВВ экспериментальное для Тре	ВВ экспериментальное для изоТре	ВВ предсказанное для Тре
MSSPEDDSDTKR	13,42	13,16	15,16
TDTLEDLFPTTK	38,7	37,82	48,48
VVTDTDETELAR	22,08	20,44	40,65
GAGTDDHTLIR	17,25	16,88	25,09
LWDLTTGTTTR	32,1	29,78	47,98
AVANQTSATFLR	25,83	23,96	44,77
DNIQGITKPAIR	23,57	21,97	32,62
ETTVGAVTVTHK	17,87	16,84	25,37
ETVEEQASTTER	18,51	16,19	24,01
FFNVLTTNTDGK	33,62	32,77	50,61
KGESQTDIEITR	19,64	18,73	23,01
KPLLESGTLGTK	22,38	21,78	34,06
KTVGVEPAADGK	15,89	14,93	22,16
NNRPSEGPLQTR	16,7	16,66	17,2
QDPSVLHTEEMR	22,52	21,63	28,35
TGHSLLHTLYGR	24,07	23,01	28,79
TGYTLDVTTGQR	26,41	22,36	34,79
TLHPDLGTDKDK	16,69	16,59	20,73
TMQNTSDLDTAR	20,07	19,55	33,88
TPYTDVNIVTIR	33	31,5	48,13

Таблица 4. Экспериментальные и предсказанные времена выхода с хроматографической колонки (ВВ) для исследуемых пептидов

TQPMTAQAASYR	21,37	19,98	31,59
TTGIVETHFTFK	29,26	26,39	41,49
VVPTTDHIDTEK	18,12	17,06	28,5
DRVTDALNATR	24,67	20,71	28,13
FIDTTLAITSR	32,16	29,99	52,78
HADIVTTTTHK	14,38	12,79	14,06
HFVALSTNTTK	22,45	19,88	32,15
NTLTQTTENLR	32,53	22,98	34,11
QLYETTDTTTR	23,19	17,76	31,63
STSTPTSPGPR	14,75	13,26	13,77
TGNFQVTELGR	41,69	41,01	37,56
TVEQTATTTNK	13,86	12,89	22,85
YIGTGHADTTK	14,09	13,64	19,75
YRDPTTVTTLR	29,15	23,12	30,57
TALTTTISSR	26,59	20,66	31,96
TSYAQHQQVR	13,85	13,46	12,18
PAATAAGSLGSLK	22,04	21,68	38,9
PNNTNNGSLGSLK	18,47	18,28	30,26
PDDTDDGSLGSLK	21,19	20,89	34,27
PEETEEGSLGSLK	21,73	21,42	32,81
PQQTQQGSLGSLK	18,9	18,75	28,21
PGGTGGGSLGSLK	19,25	18,99	27,62
PIITIIGSLGSLK	39,72	38,49	59,82
PLLTLLGSLGSLK	37,13	35,23	66,04
PMMTMMGSLGSLK	32,81	32,08	54,93
PFFTFFGSLGSLK	40,96	39,67	71,06
PPPTPPGSLGSLK	23,96	23,4	34,47
PSSTSSGSLGSLK	18,89	18,65	26,94
PTTTTTGSLGSLK	20,84	20,77	33,8
PYYTYYGSLGSLK	30,23	29,19	49,31
PVVTVVGSLGSLK	30,31	28,93	50,43

Возможность разделять пептиды в смеси позволила нам использовать спектры пептидов с Тре и изоТре также из образца, где присутствовали оба варианта пептидов. Таким образом число технических повторов масс-спектрометрического эксперимента возросло до шести.

3.7.2. Анализ спектров фрагментации пептидов с треонином и изотреонином

Главным интересом для решения задачи оптимизации метода протеомного анализа было сравнить масс-спектры фрагментации пептидов с Тре и изоТре, полученные в условиях типичного протеомного эксперимента. Для этого мы получили MC/MC спектры аналогичных пептидов при помощи ионизации по типу HCD и масс-анализатора типа Orbitrap. При сравнении мы обращали наибольшее внимание на *b*- и *у*-ионы, аннотированные при помощи программы MaxQuant. Визуально спектры аналогичных пептидов отличались (рис. 12). Данные относительной интенсивности пиков были получены из результатов MaxQuant и использованы для автоматической классификации спектров пептидов.



Рис. 12. Масс-спектры модельной пары пептидов PEETEEGSLGSLK (a) и PEE (isoT) EEGSLGSLK (b), показывающие разную относительную интенсивность характеристических ионов: b4 (b_n), b₂ (b_{n-1}), y₁₀ (y_n), y₉ (y_{n-1}), y₁₁ (y_{n+1}), y₈ (y_{n-2}).

Окончательный набор данных, использованный для автоматической классификации, состоял из 20 пептидов каждого класса. Для описания свойств спектра были выбраны шесть признаков, представляющих собой относительные интенсивности наиболее часто представленных ионов. Эти шесть ионов имели ненулевую интенсивность в более 80% пептидов. Среди ионов, выбранных в качестве признаков были: b_n, b_{n-2}, y_n, y_{n-1}, y_{n-2} и y_{n+1}, где n – это положение Тре или изоТре в пептиде. Заметим, что пептид b_{n+1} не вошел в число диагностических ионов по результатам анализа на выборке из пептидов. Во время предварительного визуального сравнения спектров фрагментации пептида 57(NQVA(T/isoT)NPTNTVFDAK)71, в котором впервые нами было обнаружено преобразование Мет в изоТре, именно ион b₆,то есть b_{n+1} оказался диагностическим для различения спектров. Однако при исследовании группы пептидов данный признак не показал диагностической значимости.

Целью анализа с применением автоматической классификации было найти комбинацию признаков, которые вносят наибольший вклад в идентификацию изомера треонина в пептиде. При анализе средней точности 1000 запусков 5-кратной перекрестной валидации для каждой комбинации признаков было обнаружено, что признак b_n сам по себе дает точность классификации 84 ± 2,4%. Добавление к нему других признаков не увеличивает точность ни отдельно, ни в комбинации.

Оцененные по шести техническим повторам средние значения относительных интенсивностей шести рассматриваемых ионов представлены в таблице 1 Приложения. Диаграммы, демонстрирующие сравнение групп пептидов с Тре и изоТре по шести исследуемым признакам в пептидах представлены на рис. 13. Из диаграмм видно, что по признаку b_n классы различаются достоверно, в то время как по остальным пяти признакам – не достоверно.



Рис. 13. Диаграммы, показывающие достоверность классификации с использованием относительной интенсивности шести ионов в качестве признаков. Диагностические ионы: b_n, b_{n-2}, y_n, y_{n-1}, y_{n-2} и y_{n+1}, где n – это положение Тре или изоТре в пептиде. Единственным признаком позволяющим достоверно классифицировать пептиды с Тре и изоТре оказался ион b_n.

Таким образом, MS/MC спектры пептидов, содержащих Тре и изоТре, полученные при HCD, отличаются по относительной интенсивности фрагмента b_n, где n – это положение треонина или изотреонина, соответственно. Этот признак не позволяет определять наличие изотреонина в панорамных протеомных данных без использования контроля. Однако это возможно с применением синтетических стандартов.

3.8. Сравнение методов алкилирования остатков цистеина при пробоподготовке к протеомному анализу

Одной из задач нашей работы было проведение сравнительного анализа четырех часто используемых в протеомике алкилирующих агентов, а именно йодацетамида (IAM), 4винилпиридина (4-VP), метилметантиосульфоната (MMTS) и хлорацетамида (CAM) и оценка влияния каждого из них на результат протеомного анализа. Для этого мы приготовили и проанализировали образцы из культуры человеческих клеток, различающиеся по методу алкилирования. Поскольку в разных работах, описывающих методы восстановления и алкилирования, применяются немного разные обозначения реагентов, в нашей работе мы выбрали обозначения одной из первых таких работ [6]. Соответствие наших обозначений обозначениям, применяемым в самой подробной сравнительной работе [7] представлено в таблице 5.

Таблица 5. Обозначения, используемые для алкилирующих агентов в нашей работе и их соответствие обозначениям условий восстановления и алкилирования в работе Müller и Winter [7]

Восстанавливающий агент	Алкилирующий агент (данная работа)	Условия по [7]	
	ІАМ (йодацетамид)	DI	
DTT (дитиотреитол)	IAA (йодуксусная кислота)	DIA	
	АА (акриламид)	DA	
	САМ (хлорацетамид)	DC	
	IAM	TI	
ТСЕР (трис(2-	IAA	TIA	
карбоксиметил)фосфин)	AA	ТА	
	CAM	TC	
	IAM	MI	
DME (0	IAA	MIA	
ыме (р-меркаптоэтанол)	AA	MA	
	CAM	MC	
-	-	CON	

Всю процедуру пробоподготовки проводили в 3 повторах, в результате чего получилось три группы по 8 образцов. В каждом повторе один образец был контрольным, то есть не подвергался ни восстановлению ни алкилированию. Остальные образцы были

восстановлены при помощи 10 мМ DTT в течение 30 минут при 57°С. Алкилирование проводили по следующим схемам:

№60, 68, 92: контроль, без восстановления, без алкилирования;

№61, 69, 93: йодацетамид (IAM) 20 мМ комнатная температура, 10 минут в темноте;

№62, 70, 94: IAM 20 мМ комнатная температура, 10 минут в темноте, после чего добавление DTT до 30 мМ;

№63, 71, 95: 4-винилпиридин (4-VP) 20 мМ комнатная температура, 10 минут в темноте;

№64, 72, 96: 4-VP 20 мМ комнатная температура, 10 минут в темноте, после чего добавление DTT до 30 мМ;

№65, 73, 97: хлорацетамид (САМ) 20 мМ комнатная температура, 10 минут в темноте;

№66, 74, 98: САМ 20 мМ комнатная температура, 10 минут в темноте, после чего добавление DTT до 30 мМ;

№67, 75, 99: метилметантиосульфонат (MMTS) комнатная температура, 10 минут в темноте.

Образцы №№ 60-67 получали из клеток линии HeLa и обозначали как повтор пробоподготовки I. Образцы №№ 68-75 также получали из клеток линии HeLa и обозначали, как повтор II. Образцы №№92-99 получали из клеток линии HepG2, и обозначали как повтор III. На этапе масс-спектрометрии каждый образец анализировали еще в трех технических повторах.

Три образца, алкилированные IAM, 4-VP и CAM были подвергнуты повторной обработке DTT для проверки эффекта гашения избытка алкилирующего агента. Этим образцам присваивали обозначения IAMq, 4-VPq, CAMq, соответственно. Такой подход нередко применяют в протеомике [70], однако, многие исследователи игнорируют данный этап [61]. Кроме IAM, избыток других алкилирующих агентов также убирают при помощи добавления восстановителя, например, 4-VP [76]. Образец, алкилированный MMTS, не подвергали такой процедуре. Комплекс цистеина и MMTS имеет собственную S-S связь (рис. 4), поэтому такое алкилирование обратимо, и последующая обработка DTT не имеет смысла. Несмотря на это, существуют единичные работы, в которых проводится гашение избытка MMTS добавлением DTT [152].

Для оценки пригодности исследуемых алкилирующих агентов для протеомики мы анализировали общее количество идентифицированных пептидов и спектральных идентификаций (PSM), а также количество пептидов и PSM, содержащих остатки цистеина и метионина по отдельности. Эти значения были нормализованы на наибольшее значение в каждом повторе пробоподготовки, они представлены на рис. 14 и рис. 1 и 2 Приложения. На рис. 14 представлены наиболее наглядные значения общего количества пептидов и отношения количества пептидов с Мет и Цис к общему количеству пептидов. На рисунках Приложения

приведены значения PSM и пептидов с Цис и Мет, не поделенные на общее количество пептидов и PSM. Такая нормализация была необходима, поскольку повторы пробоподготовки проводили в разное время и на разных приборах, и абсолютные значения исследуемых параметров существенно различались между такими повторами в связи с разной работой хроматомасс-спектрометрии. При этом общие закономерности сохранялись. Значения парного критерия Стьюдента (paired t-test), значения уровня значимости (p-value) и значения кратного изменения (fold change) приведены в таблице 2 Приложения. Сами нормализованные значения исследуемых параметров приведены в таблице 3 Приложения.



Рис. 14. Диаграммы, показывающие медианы и межквартильные интервалы нормализованных значений числа пептидов, идентифицированных после разных методов алкилирования.

Сокращения расшифрованы в таблице 5; q - образцы, в которых избыток алкилирующего агента был погашен добавлением DTT после алкилирования. Значения были нормализованы по наибольшему внутри 3 повторов пробоподготовки, что дало 9 значений, полученных после объединения технических повторов и повторов образца пробоподготовки. А: общее число идентифицированных пептидов; Б: отношение числа идентифицированных пептидов с Цис к количеству всех идентифицированных пептидов с Мет к количеству всех идентифицированных пептидов.

По общему числу идентифицированных пептидов контрольный образец не отличался от остальных образцов (рис. 14А). Ожидаемо, в этом образце практически отсутствуют идентифицированные пептиды с цистеином (рис. 14Б). Из-за сохранившихся S-S связей, такие пептиды были непредсказуемо связаны друг с другом и не могли быть идентифицированы. Такой результат вполне согласуется с тем, что было получено ранее [7]. На первый взгляд сравнительно большое количество идентифицированных пептидов в контрольном образце кажется неожиданным. Однако такой эффект, по-видимому, связан с тем, что в образцах, подверженных восстановлению и алкилированию, накапливаются побочные реакции, затрудняющие в результате идентификацию части пептидов. Таким образом, почти полное отсутствие идентифицированных пептидов с Цис компенсируется идентификацией других пептидов, содержащих аминокислотные остатки, уязвимые к побочным реакциям при алкилировании. Также показано, что количество пептидов с пропущенными сайтами трипсинолиза в контрольном образце не превышает количество таких пептидов в образцах, подверженных алкилированию [7]. Это косвенно может свидетельствовать о том, что отсутствие алкилирования не влияет на доступность сайтов гидролиза для фермента. Тем не менее, в работе [7] утверждается, что контрольные образцы демонстрировали бо́льшую вариабельность между повторами, чем алкилирование образцы. Этот факт, по мнению авторов, говорит о том, что неалкилированные белки претерпевают большое количество случайных модификаций [7]. Наши результаты не подтверждают данное наблюдение. В наших данных разброс значений между повторами в контрольном образце получился наименьшим (рис. 14А). Существуют и совсем новые работы, также подтверждающие тот факт, что, в целом, восстановление и алкилирование не ведет к увеличению числа идентифицированных пептидов и «углублению» протеомного анализа [153].

Наименьшее количество идентифицированных пептидов наблюдается для образца, алкилированного йодацетимидом (IAM) (рис. 14А), оно в 1,4 раза ниже по сравнению с контрольным образцом (р <0,001). Частично это вызвано тем, что в этом образце очень мало пептидов с Мет (рис. 14В): почти в четыре раза меньше по сравнению с контролем (р <0,001). Подобное было описано в работе Müller и Winter [7], где авторы предлагают очень детальное исследование причин отсутствия среди идентифицированных пептидов пептидов с Мет.

Большая часть Мет в пептидах в их работе претерпела либо окисление, либо алкилирование, либо потерю боковой цепи SH-CH₃ после окисления [7].

В отличие от образца, алкилированного с использованием IAM, образец, в котором избыток IAM был погашен добавлением DTT после алкилирования (образец «IAMq»), демонстрирует другой результат. Общее количество идентифицированных пептидов для этого образца показывает небольшое (в 1,33 раза) увеличение по сравнению с образцом IAM (р <0,001). Относительное количество пептидов с Цис для этого образца в 1,3 раза меньше по сравнению с образцом IAM (р <0,001), однако относительное количество пептидов с Мет является высоким и примерно таким же, как в контроле (рис. 14В). Таким образом, йодацетамид не является предпочтительным алкилирующим агентом для протеомного анализа, однако, в случае невозможности использования альтернативных агентов, избыток IAM следует погасить добавлением восстановителя после реакции алкилирования.

Общее количество идентифицированных пептидов для образца, алкилированного 4винилпиридином (4-VP), в среднем, немного ниже по сравнению с контролем (в 1,1 раза). Тем не менее, отклонение этих значений среди повторов для этого образца очень велико (рис. 14А), особенно по параметру относительного количества пептидов с Цис (рис. 14Б). Возможным объяснением такого эффекта может являться тот факт, что 4-VP плохо растворяется в воде. Это затрудняет работу с данным реагентом. Для проведения реакции с 4-VP в водных растворах, для начала готовят раствор 4-винилпиридина в диметилформамиде (DMF), а только потом добавляют полученный реагент в буфер с белками. Таким образом, велика вероятность неполного растворения 4-VP и, как следствие, его неоднородного распределения в конечном растворе, что может привести к избыточному или недостаточному алкилированию белков в образцах.

Гашение избытка 4-VP после реакции (4-VPq) улучшает воспроизводимость с уменьшением общего количества идентифицированных пептидов в 1,32 раза по сравнению с образцом без гашения 4-VP (р <0,001). Тем не менее, использование 4-VP защищает остатки метионина и приводит к высокому относительному количеству метионинсодержащих пептидов (рис. 14В).

Алкилирование хлорацетамидом (САМ) дало наибольшее количество идентифицированных пептидов, и этот результат был в высокой степени воспроизводимым (рис. 14А). Соотношения цистеин- и метионинсодержащих пептидов относительно всех идентифицированных пептидов также сравнительно высоки (рис. 14Б, В). Гашение этого реагента с помощью DTT после обработки ухудшает результат с точки зрения общего количества идентифицированных пептидов и, особенно, воспроизводимости. Наши наблюдения согласуются с данными других работ [7], где использование САМ в сочетании с DTT, также,

дало большое количество идентифицированных пептидов и PSM в сравнении с IAM. Существуют единичные сообщения о том, что CAM способен вызывать побочные реакции при алкилировании и модифицировать другие аминокислотные остатки [80,96], однако в данных работах речь идет об использовании CAM при повышенных температурах, чего следует внимательно избегать при проведении реакции. Кроме того, авторы [7] сообщают об алкилировании N-конца пептида при использовании CAM, IAM, AA и IAA, но подчеркивают, что CAM вызывает этот эффект меньше, чем другие реагенты.

Общее количество пептидов в образце, алкилированном метилметантиосульфонатом (MMTS) в 1,23 раза выше (р 0,04) по сравнению с образцом, алкилированным с использованием ІАМ. Последний который дал самое низкое количество пептидов, не считая экспериментов с гашением избытка. С другой стороны, число пептидов для ММТЅ ниже, чем количество пептидов, идентифицированных для образца, алкилированного с использованием САМ, в 1,17 раза (р 0,002). Относительное количество пептидов с Цис, идентифицированных для образца, алкилированного с использованием MMTS, также в 1,5 раза ниже по сравнению с образцом, алкилированным с использованием САМ (р <0,001). Метилметантиосульфонат используется в редокс-протеомике при исследовании нативной структуры S-S связей [104]. Метод основан на обратимом связывании MMTS с остатками цистеина. Использование MMTS позволяет выявить какие именно остатки цистеина изначально были связаны S-S связями в белке, а какие – были свободными. Некоторые авторы рекомендуют использовать именно этот реагент в панорамной протеомике [106]. Одним из возможных объяснений малого количества идентифицированных пептидов и PSM с Цис (рис. 14Б, рис. 1Б и 2Б приложения) в наших данных является следующее. Существуют сообщения о том, что MMTS может вызывать артефакты при участии остатков цистеина [104]. Дело в том, что MMTS образует дитиометановые группы, которые, в последствии, могут образовывать изомеры. Как следствие, в белках, алкилированных MMTS, возможно образование внутримолекулярных и межмолекулярных связей [104]. Если такой эффект в наших образцах действительно возникает, он может быть причиной сниженного числа идентифицированных пептидов и PSM с Цис в результатах нашего анализа. Однако такая гипотеза требует отдельной проверки.

Наблюдаемая картина для спектральных идентификаций (PSM) сходна с таковой для идентифицированных пептидов (рис. 1, 2 и Таблица 2 приложения).

3.8. Анализ частоты преобразования метионина в изотреонин в протеомных данных

Для общей оценки частоты преобразования метионина в изотреонин в протеомных данных мы использовали общедоступные данные из трех работ [7,131,137], а также наши собственные данные, полученные с применением разных алкилирующих агентов. Стоит

обратить внимание, что в работе Müller и Winter [7] образцы получали как методом пробоподготовки в геле, так и с использованием центрифужных фильтров (FASP) [60]. Метод центрифужных фильтров в данном случае может быть рассмотрен как пробоподготовка в растворе, поскольку все реакции проводят в верхней камере фильтра. Данные протеомов клеточных линий панели NCI60 [131] были получены одинаковым методом - разделением белков электрофорезом с последующим алкилированием при помощи IAM и трипсинолизом в геле. Кроме этого, мы проанализировали данные протеомов нескольких отделов мозга мыши [137]. В данной работе все образцы были приготовлены в растворе с использованием IAM. Для нашей работы мы использовали данные девяти отделов мозга [137].

В данных, полученных в нашей работе, наибольшая доля пептидов, содержащих преобразование метионина в изотреонин наблюдается в образце, алкилированном IAM (рис. 15А). В этом образце содержание соответствующего сдвига массы составляет около 3%. Доля пептидов, несущих модификацию Мет-изоТре в образцах, алкилированных другими агентами, практически одинакова (рис. 15А) и не превышает 1%.



Рис. 15. Отношение числа пептидов с преобразованием Мет-изоТре к числу всех идентифицированных пептидов, содержащих метионин. Показаны 95% доверительные интервалы критерия X². А: Данные из нашей работы; Б: Данные из работы Müller и Winter [7], образцы, отмеченные синим цветом, были приготовлены в геле, сокращения расшифрованы в таблице 4; В: Данные из работ по клеточным линиям NCI60 [131] (отмечены синим, все образцы приготовлены в геле) и Sharma с соавт. [137].

Похожий результат мы получили по данным из работы Müller и Winter [7]. Наибольшая доля преобразований метионина в изотреонин наблюдается в образцах, алкилированных

йодацетамидом (IAM) и йодоуксусной кислотой (IAA) (рис. 15Б), все образцы, обозначение которых оканчивается на «I» и «IA») по сравнению с образцами, алкилированными другими агентами. Процент пептидов с преобразованием Мет в изоТре в образцах IAM и IAA также составлял от 1 до 3%. Как и в наших данных, этот результат может быть вызван меньшим количеством пептидов, содержащих Мет в этих образцах (таблица 4 Приложения). Процент пептидов с модифицированным Мет в других образцах находился в диапазоне от 0,1 до 0,4%.

Как уже было сказано, все образцы в работе [131] были приготовлены единообразно, с использованием IAM и пробоподготовки в геле. В этих данных, для всех клеточных линий, за исключением PXF-393, процент пептидов, содержащих преобразования метионина в изотреонин составил 1,5-3% (рис. 15В). Данные из работы [131] представляют собой более «глубокие» протеомы, позволяющие получить большее количество идентифицированных пептидов и белков, чем предыдущие два рассмотренных набора данных. Кроме того, абсолютное количество пептидов с метионином, как и количество пептидов с модификацией Мет-изоТре, в этих данных выше (таблица 4 Приложения), что делает результат более статистически достоверным. В среднем, в каждой клеточной линии было обнаружено около 300 пептидов с модифицированным в изотреонин метионином.

Подобно данным клеточных линий NCI60 [131], данные по мышиному мозгу [137] обеспечивают глубокое покрытие протеома с большим количеством идентификаций пептидов и белков (таблица 3 Приложения). Тем не менее, доля пептидов с преобразованием Мет в изоТре в них составляет всего 0,1-0,2% (рис. 15В, таблица 4 Приложения).

Таким образом, даже оптимизированный метод пробоподготовки не позволяет полностью избежать преобразований метионина в изотреонин. До сих пор нельзя с уверенностью сказать, что конкретно вызывает это преобразование. Результаты эксперимента с альбумином, а также описанный ранее механизм реакции йодуксусной кислоты с метионином [110], в которой, вместо кислоты, может также участвовать соответствующий амид, подтверждают предположение о том, что преобразование происходит в результате алкилирования белков йодацетамидом. Однако небольшая часть остатков метионина преобразуется в изотреонин и в образцах, алкилированных другими реагентами и даже в контрольных образцах, неалкилированных вовсе. Результаты анализа данных из работы [7] наводят на мысль о том, что метионин чаще преобразуется в изотреонин, если образцы готовили в геле. Для проверки этой гипотезы мы выбрали еще два набора данных высокого разрешения: в одной из них авторы проводили пробоподготовку в геле [131], а во второй – в растворе [110]. Полученный результат, в некоторой степени, подтвердил наше предположение. Действительно, в данных из [131] процент пептидов с преобразованным в изотреонин метионином существенно превышал такой процент в данных из [110] (рис. 15В).

Воспользовавшись тем фактом, что все образцы NCI60 в работе [131] были подготовлены одинаковым образом, и число преобразований, найденных для этих данных, было самым высоким, мы проанализировали возможное влияние соседствующих аминокислотных остатков на частоту преобразования метионина. Результаты представлены в таблице 5 Приложения. Отношения количества модифицированных остатков метионина к общему числу комбинаций каждого другого остатка с метиониниом, а также 95% доверительные интервалы критерия X^2 представлены на рис. 16. Среди аминокислотных остатков, предшествующих метионину не было такого, который бы достоверно влиял на частоту преобразования (рис. 16А). Однако остаток пролина, следующий после метионина, повидимому, увеличивает частоту модификации Мет в изоТре (рис. 16В). Описанный механизм реакции йодуксусной кислоты с метионином [110], справедливый также для йодацетамида, предполагает образование карбоксиметилсульфониевой соли метионина с последующим образованием гомосерина, то есть изотреонина. Пролин в белках не является реакционноспособным аминокислотным остатком и, с химической точки зрения, не может влиять на эту реакцию. По всей видимости, пролин в соседстве с метионином влияет на пространственную структуру участка белковой цепи и делает такой метионин более доступным для реакции преобразования в изотреонин.



Рис. 16. Отношения количества модифицированных остатков метионина к общему числу комбинаций каждого другого остатка с метиониниом. Показаны 95% доверительные интервалы критерия X². А: аминокислотные остатки предшествующие метионину; Б: аминокислотные остатки следующие за метионином.

3.9. «Открытый» поиск и статистический анализ сдвигов молекулярной массы

Для того, чтобы проверить данные на наличие других модификаций аминокислотных остатков, в том числе метионина, мы провели открытый поиск [88] с последующим анализом его результатов с использованием инструмента AA_stat [108]. В результате были выявлены наиболее представленные сдвиги молекулярных масс пептидов, и предсказана их принадлежность к определенному аминокислотному остатку. Данные Müller и Winter [7] в открытом поиске не рассматривали из-за низкого разрешения спектров MC/MC. Из данных протеомов разных отделов мозга мыши [137] мы взяли для открытого поиска три отдела. Поиск проводили в два раунда. Первый раунд не включал никаких постоянных и вариабельных модификаций и нужен был для того, чтобы подтвердить наличие сдвигов массы, относящихся к

остаткам цистеина, соответствующих использованным алкилирующим агентам. Во второй раунде эти сдвиги масс были внесены как постоянные модификации для более точного анализа остальных, менее представленных сдвигов масс.

Как и ожидалось, первый открытый поиск показал сдвиг массы на +57.02 Да в образцах, обработанных IAM и CAM, что соответствует карбамидометилированию остатков цистеина. В образцах, алкилированных 4-VP, значительный сдвиг массы составил +105,05 Да, что соответствует пиридинэтилированию остатков цистеина. Для образца, полученного после обработки MMTS, сдвиг массы на +45,98 Да был связан с сульфенилированием остатков цистеина. После второго раунда открытого поиска, в котором указанные ожидаемые модификации использовались в качестве фиксированных, мы обнаружили ряд дополнительных сдвигов масс для пептидов, которые прошли фильтровку до 2% FDR. В таблице 6 приведены наиболее представленные сдвиги масс, полученные во втором раунде открытого поиска. Три сдвига масс было обнаружено во всех анализируемых наборах данных (табл. 6). В наших данных сдвиг массы на -128,1 Да присутствует в некоторых образцах, затрагивает 1-2,5% и не зависит от метода обработки остатков цистеина. В данных NCI60 [131] он присутствует во всех клеточных линиях и затрагивает около 0,3% пептидов. По данным [137] он также присутствует во всех образцах и затрагивает 0,5-0,7% пептидов. Согласно Unimod (http://www.unimod.org/) [154], этот сдвиг массы наиболее вероятно соответствует потере С-концевого остатка лизина. По результатам анализа при помощи AA stat, он не обогащен пептидами, содержащими какуюлибо конкретную аминокислоту, и, таким образом, модификация, вероятно, действительно происходит на конце пептида. Стоит заметить, что потеря С-концевого остатка при ионизации в масс-спектрометрии описана в литературе [155].

Второй сдвиг массы, присутствующий во всех трех наборах данных, составлял +42,01 Да и также не был приписан AA_stat к какому-либо конкретному остатку (рис. 17А). Он затронул 1–1,5% пептидов в наших данных, 1,5–3,5% пептидов в данных NCI60 [131] и 0,2– 0,3% пептидов в данных Sharma и соавт. [137]. Этот сдвиг массы наиболее вероятно соответствует ацетилированию N-конца белка, что является обычной посттрансляционной модификацией [156].


Рис. 17. Примеры результатов AA_stat, показывающих распределение аминокислот в трех бинах сдвигов масс из данных [131]. А: Сдвига массы на 42,007 Да соответствует ацетилированию на N-конце пептида; Б: Сдвиг массы на 57,019 Да соответствует карбамидометилированию остатков метионина; Б: Сдвиг массы на -29,997 Да соответствует преобразованию метионина в изотреонин.

Наиболее представленный сдвиг молекулярной массы, который также присутствовал во всех трех наборах данных, составил +57,02 Да с явным обогащением пептидами, содержащими метионин (рис. 17Б). Стоит обратить внимание, что карбамидометилирование остатков цистеина (+57,02 Да) использовали во время поиска в качестве фиксированной модификации. В наших данных этот сдвиг массы наблюдается только для образцов, алкилированных йодацетамидом (IAM), и отсутствует в тех образцах, где избыток IAM был погашен добавлением восстановителя (IAMq). В подгруппе обработки IAM эта модификация затрагивает от 6 до 13% пептидов. В данных NCI60 эта доля варьирует от 0,2 до 3%. В данных по отделам мышиного мозга этот сдвиг массы был обнаружен для 2 из 3 образцов, проанализированных при помощи открытого поиска. В одном из них доля пептидов, содержащих этот сдвиг массы, составляла 0,24%, а в другом она достигала 11,4%. При этом все эти образцы прошли одну и ту же процедуру алкилирования и гидролиза. Такие различия, повидимому, представляют собой погрешности пробоподготовки. В образце с повышенным карбамидометилированием метионина, вероятно, оказалась более высокая концентрация IAM.

Мы также обнаружили менее обильный, но интересный для нашей работы, сдвиг массы на -30 Да в 3 образцах из данных по NCI60. Этот сдвиг массы также был обогащен пептидами с метионином (рис. 17В). Доля пептидов, несущих этот сдвиг массы, составляла от 0,15 до 0,17%. Вероятнее всего этот сдвиг массы соответствует замене метионина на изотреонин. Из-за алгоритмических особенностей открытого поиска его чувствительность ниже, чем у обычного протеомного поиска, настроенного на конкретную вариабельную модификацию. Это объясняет более низкий процент пептидов с модификацией, обнаруженной в одних и тех же данных разными способами. Обнаружение сдвига массы на -30 Да при помощи открытого поиска и ассоциация этого сдвига к остатками метионина при помощи AA_stat подтверждают наличие модификации Мет-изоТре, обнаруженной нами при обычном протеомном поиске.

По результатам открытого поиска мы обнаружили и другие любопытные модификации аминокислотных остатков. Например, сдвиг массы на +183,03 Да по информации из Unimod соответствует аминоэтилбензолсульфонилированию на конце пептида. Действительно, в работе, в данных из которой мы обнаружили этот сдвиг массы [131], авторы использовали смесь ингибиторов протеаз фирмы Sigma. По информации от производителя, эта смесь содержит аминоэтилбензолсульфонил, то есть вполне может вызывать указанную модификацию. Такие наблюдения подтверждают ценность метода открытого поиска для обнаружения нестандартных модификаций аминокислотных остатков при протеомном анализе. Таблица 6А. Наиболее представленные в данных сдвиги молекулярной массы пептидов, обнаруженные в результатах открытого поиска. Аминокислотные остатки, к которым относятся сдвиги масс предсказаны при помощи инструмента AA_stat и интерпретированы по базе данных Unimod. Для каждого сдвига масс представлена доля затронутых пептидов в процентах. Данные 8 образцов в 3-х повторах (обозначены I, II, III) из нашей работы. Представлены образцы, для которых обнаружены наиболее обильные сдвиги масс.

сдвиг	populating volutioning	предсказанный	№ в	клетки HeLa и HepG2 обработанные разными алкилирующими агентами в растворе, (наши данные)								
массы, Да	возможная модификация	AA_stat)	Unimod	control I	IAM I	IAM II	IAM III	IAMq III	4-VP II	4-VPq II	CAM II	CAMq III
15,99	окисление	Мет	35									
57,019	карбамидометилирование	Мет	4		12,47	7,16	5,64					
42,009	ацетилирование	конец	1					1,55				0,8
52,91	замена 3 протонов на Fe	конец	1870									
31,98	двойное окисление	Мет, Три	425									
-128,097	потеря С-концевого Лиз	конец	313	1,25	1,11				2,15	2,21	2,5	

Таблица 6Б. Наиболее представленные в данных сдвиги молекулярной массы пептидов, обнаруженные в результатах открытого поиска. Аминокислотные остатки, к которым относятся сдвиги масс предсказаны при помощи инструмента AA_stat и интерпретированы по базе данных Unimod. Для каждого сдвига масс представлена доля затронутых пептидов в процентах. Данные 9 клеточных линий из [131]

слвиг		прелсказанный		Девять клеточных линий панели NCI60, обработка в геле								
массы, возможная модификация Да		а.о. (данные AA_stat)	Nº B Unimod	MCF7	SK- OV- 3	U251	RXF- 393	COLO205	NCI- H460	M14	CCRF- CEM	PC-3
15,99	окисление	Мет	35	7,27	8,7	6,87	7,02	6,51	8,5	7,09	6,93	10,72
57,019	карбамидометилирование	Мет	4	2,75	0,2	2,24		1,69	1,87	2,91	3,06	1
42,009	ацетилирование	конец	1	3,41	3,05	2,17	1,78	2,21	2,45	2,32	2,08	1,48
52,91	замена 3 протонов на Fe	конец	1870	1,45	1,52	0,84	0,9	1,09	1,57	1,05	1,34	0,87

31,98	двойное окисление	Мет, Три	425	0,88	1,38	0,76	1,03	1	1,28	0,97	0,65	1,88
- 128,097	потеря С-концевого Лиз	конец	313	0,34	0,3	0,22	0,36	0,27	0,27	0,32	0,3	0,33
128,09	добавление Лиз в результате транспептидации	конец	1301	0,71	0,8	0,77	0,8	0,81	0,64	0,62	0,69	0,56
-17,03	потеря аммония	конец	385	0,33	0,48	0,43	0,43		0,45	0,43	0,39	0,85
183,03	аминоэтилбензолсульфонилирование	конец	276	0,18	0,83		0,28	0,68	0,3	0,58		
55,91	замена 2 протонов на Ni	конец	953		0,47				0,56	0,53	0,51	0,6
37,94	замена 2 протонов на Са	конец	951	0,22	0,29				0,27	0,24	0,2	0,24
-18,01	дегидратация	конец или разные	23		0,16		0,22		0,2	0,19	0,16	0,3
-48,009	потеря боковой цепи окисленного Мет	Мет	526	0,16			0,24		0,2	0,22	0,18	
-30	Мет->изоТре	Мет	610	0,15						0,17	0,15	
-156,1	потеря С-концевого аргинина	конец	1287	0,1	0,1		0,12			0,14	0,1	

Таблица 6В. Наиболее представленные в данных сдвиги молекулярной массы пептидов, обнаруженные в результатах открытого поиска. Аминокислотные остатки, к которым относятся сдвиги масс предсказаны при помощи инструмента AA_stat и интерпретированы по базе данных Unimod. Для каждого сдвига масс представлена доля затронутых пептидов в процентах. Данные 3 отделов мозга мыши из [137]

сдвиг		предсказанный	№ в	Образцы мозга мыши, обработка в растворе [23]			
массы, Да	возможная модификация	AA_stat)	Unimod	brain_frac- tioned	CG_neu- rons	adult_micro- glia	
15,99	окисление	Мет	35	5,61	2,31	5,21	

57,019	карбамидометилирование	Мет	4		11,39	0,24
42,009	ацетилирование	конец	1	0,19	0,17	0,27
52,91	замена 3 протонов на Fe	конец	1870	1,37	0,55	1,68
31,98	двойное окисление	Мет, Три	425	0,19	0,14	0,67
- 128,097	потеря С-концевого Лиз	конец	313	0,7	0,5	0,55
128,09	добавление Лиз в результате транспептидации	конец	1301		0,14	0,26
-17,03	потеря аммония	конец	385	0,97		0,73
183,03	аминоэтилбензолсульфонилирование	конец	276			
55,91	замена 2 протонов на Ni	конец	953		0,18	0,19
37,94	замена 2 протонов на Са	конец	951			
-18,01	дегидратация	конец или разные	23	0,24	0,16	0,34
-48,009	потеря боковой цепи окисленного Мет	Мет	526			
-30	Мет->изоТре	Мет	610			
-156,1	потеря С-концевого аргинина	конец	1287	0,1		0,08

3.10. Окисление и карбамидометилирование метионина

После обнаружения в наших данных хорошо представленного сдвига массы на +57 Да, ассоциированного с пептидами, содержащими метионин, в образце, алкилированном с помощью IAM, мы повторили обычный протеомный поиск при помощи IdentiPy с карбамидометилированием метионина в качестве вариабельной модификации. Известно, что IAM также может алкилировать метионин [7,21,83]. Образцы, алкилированные с использованием САМ, анализировали для сравнения. Кроме карбамидометилирования мы включили в список вариабельных модификаций окисление метионина. Ранее было показано, что САМ может увеличивать окисление Met [96], и мы ожидали увидеть этот эффект в наших данных и сравнить его масштаб для образцов, алкилированных САМ и IAM.

В образцах, алкилированных с использованием IAM, в зависимости от повтора, доля пептидов, содержащих карбамидометилированный Мет, находилась в диапазоне от 40 до 80% по отношению ко всем идентифицированным пептидам, содержащим этот остаток. Доля пептидов с окисленным метионином составила всего лишь от 1 до 5%. В образцах, алкилированных с использованием САМ, напротив, было менее 1% пептидов, содержащих карбамидометилированный Мет по отношению ко всем Мет-содержащим пептидам. Доля пептидов с окисленным Мет в образцах, алкилированных САМ, составляла от 1 до 6%, аналогично образцам, алкилированным IAM.

Как уже было сказано выше, достаточно детальный анализ причин сниженного числа пептидов с Мет в образцах, алкилированных IAM был проведен ранее [7]. В этой статье авторы при помощи метода, близкого к методу открытого поиска, выяснили, что метионин подвергается окислению, карбамидометилированию и нейтральной потере боковой цепи SH-CH₃ после окисления (сдвиг массы -48 Да, идентификатор модификации на Unimod №526) [94]. Интересно, что потеря боковой цепи в остатках метионина при масс-спектрометрическом анализе случается не только после окисления, но и после карбамидометилирования. В таком случае, сдвиг массы также составляет -48 Да [7,95] (рис. 5). В работе [7] авторы указывают на то, что большая часть остатков метионина в их образцах претерпела такую модификацию. Однако в наших данных присутствует именно сдвиг массы на +57 Да, что соответствует карбамидометилированию метионина без нейтральной потери.

ЗАКЛЮЧЕНИЕ

В протеомике, исследователи не всегда уделяют внимание оптимизации метода подготовки образцов, концентрируясь на технически более сложной хроматомасс-

78

спектрометрии. Тем не менее, именно во время пробоподготовки в протеомном анализе могут возникать артефакты, существенно влияющие на результат. Наиболее уязвим к таким артефактам этап алкилирования остатков цистеина в белках, предшествующий ферментативному гидролизу. В работе удалось изучить некоторые побочные реакции этого процесса. Проведено сравнение нескольких алкилирующих агентов и проанализировано их влияние на частоту преобразования метионина в изотреонин. Кроме того, среди изученных алкилирующих агентов мы показали наиболее оптимальный для протеомных исследований хлорацетамид. Также исследовано влияние соседствующих с метионином аминокислотных остатков на частоту преобразования такого метионина в изотреонин. Помимо этой реакции для метионина характерны и другие модификации, случающиеся в процессе пробоподготовки, такие как, например, окисление и карбамидометилирование. В работе проиллюстрировано, что эти и другие нежелательные модификации аминокислотных остатков можно обнаружить в данных, применяя метод открытого поиска. Наиболее представленные сдвиги масс из результатов открытого поиска возможно использовать в качестве вариативных модификаций в основном протеомном поиске, что способно увеличить качество анализа.

Изотреонин (гомосерин), появляющийся в белках под действием химической модификации метионина по молекулярной массе равен треонину. Это приводит к тому, что эти аминокислотные остатки трудно различить при помощи масс-спектрометрии, а преобразование метионина в изотреонин может легко имитировать замену метионина на треонин, закодированную в геноме. Мы проанализировали хроматограммы и масс-спектры синтетических пептидов, содержащих указанные изомерные аминокислотные остатки. В результате нам далось описать различия в хроматографических свойствах и масс-спектрах фрагментации таких пептидов.

выводы

1. Методами масс-спектрометрического протеомного анализа показано, что преобразование метионина в изотреонин может затронуть до 3% пептидов, содержащих метионин. Этот эффект имеет большое значение для протеогеномики, поскольку такая модификация может имитировать замену метионина на треонин, закодированную в геноме. Преобразование метионина в изотреонин происходит чаще, если образец подвергается алкилированию йодсодержащими реагентами, то есть йодуксусной кислотой или йодацетамидом, а подготовка образца осуществляется в полиакриламидном геле. Анализ влияния соседствующих аминокислотных остатков на частоту преобразования метионина в

79

изотреонин показал, что пролин, следующий за метионином в последовательности белка, увеличивает частоту модификации.

2. Из четырех проанализированных алкилирующих агентов, а именно, йодацетамида, хлорацетамида, 4-винилпиридина и метилметантиосульфоната, хлорацетамид лучше других подходит для подготовки образцов к протеомному анализу, поскольку вызывает меньше побочных реакций и позволяет идентифицировать больше пептидов. Наиболее часто используемый для этого йодацетамид способен вызывать карбамидометилирование метионина, что приводит к уменьшению числа идентифицированных пептидов, содержащих этот остаток.

3. Результаты исследования показали, что пары пептидов, содержащих треонин и изотреонин, могут быть разделены с помощью обращенно-фазовой высокоэффективной жидкостной хроматографии в условиях, обычно используемых при панорамном протеомном анализе. Пептиды, содержащие треонин и изотреонин различаются по времени удержания на хроматографической колонке. Другим обнаруженным различием являются разные спектры фрагментации пептидов, содержащих треонин и изотреонин, в которых присутствие остатка изотреонина приводит к более высокому пику b_n-иона, где п – положение треонина в последовательности. В частности, до 84% всех пептидов, содержащих треонин в своей последовательности, были правильно классифицированы методом машинного обучения с использованием относительной интенсивности характеризующего пика.

ФИНАНСИРОВАНИЕ

Работа поддержана грантами Российского научного фонда № 14-15-00395 (руководитель Мошковский С.А.) и Российского фонда фундаментальных исследований № 18-315-00226 (руководитель Кузнецова К.Г.).

БЛАГОДАРНОСТИ

Автор благодарит руководителя и коллег за помощь в проведении работы и подготовки диссертации. В частности, автор благодарен д.б.н. Згоде В.Г. за проведение массспектрометрического анализа, к.б.н. Мойсе А.А. за проведение синтеза пептидов, к.б.н. Пятницкому М.А., Труфанову П.А., к.б.н. Чернобровкину А.Л., Левицкому Л.И. и Бубис Ю.А. за консультацию и помощь при обработке данных. Кроме того, автор благодарит к.ф.-м.н. Горшкова М.В., Ильину И.Ю., Карпова Д.С., Карпову М.А., Иванова М.В., Соловьеву Е.М., Кузикова А.В. и других коллег, участвовавших в проекте или помогавших его проведению.

ЛИТЕРАТУРА

- 1. Gubin M.M. et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens // Nature. 2014. Vol. 515, № 7528. P. 577–581.
- Polyakova A., Kuznetsova K., Moshkovskii S. Proteogenomics meets cancer immunology: mass spectrometric discovery and analysis of neoantigens. // Expert Rev. Proteomics. 2015. Vol. 12, № 5. P. 533–541.
- 3. Yadav M. et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing // Nature. Nature Research, 2014. Vol. 515, № 7528. P. 572–576.
- 4. Nielsen M.L. et al. Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry // Nat. Methods. 2008. Vol. 5, № 6. P. 459–460.
- Krüger R. et al. Iodoacetamide-alkylated methionine can mimic neutral loss of phosphoric acid from phosphopeptides as exemplified by nano-electrospray ionization quadrupole time-of-flight parent ion scanning // Rapid Commun. Mass Spectrom. 2005. Vol. 19, № 12. P. 1709–1716.
- Jiang X. et al. The effect of various S-alkylating agents on the chromatographic behavior of cysteine-containing peptides in reversed-phase chromatography // J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 2013. Vol. 915–916. P. 57–63.
- Müller T., Winter D. Systematic Evaluation of Protein Reduction and Alkylation Reveals Massive Unspecific Side Effects by Iodine-containing Reagents // Mol. Cell. Proteomics. 2017. Vol. 16, № 7. P. 1173–1187.
- Deutsch E.W. et al. Human Proteome Project Mass Spectrometry Data Interpretation Guidelines
 2.1 [Electronic resource] // Journal of Proteome Research. American Chemical Society, 2016.
 Vol. 15, № 11. P. 3961–3970. URL: /pmc/articles/PMC5096969/?report=abstract (accessed: 07.08.2020).
- Wilkins M.R. et al. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it // Biotechnol. Genet. Eng. Rev. 1996. Vol. 13, № 1. P. 19–50.
- Edman P. et al. Method for Determination of the Amino Acid Sequence in Peptides. // Acta Chem. Scand. 1950. Vol. 4. P. 283–293.
- Laemli U.K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4 // Nature. 1970. Vol. 227. P. 680–685.
- O'Farrell P.H. High resolution two-dimensional electrophoresis of proteins. // J. Biol. Chem. 1975. Vol. 250, № 10. P. 4007–4021.
- Fenn J.B. Electrospray wings for molecular elephants (Nobel lecture) // Angew. Chemie Int.
 Ed. 2003. Vol. 42, № 33. P. 3871–3894.
- 14. Washburn M.P., Wolters D., Yates J.R. Large-scale analysis of the yeast proteome by

multidimensional protein identification technology // Nat. Biotechnol. 2001. Vol. 19, № 3. P. 242–247.

- Pitt J.J. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. // Clin. Biochem. Rev. 2009. Vol. 30, № 1. P. 19–34.
- Wilhelm M. et al. Mass-spectrometry-based draft of the human proteome // Nature. 2014. Vol. 509, № 7502. P. 582–587.
- Moore S.M., Hess S.M., Jorgenson J.W. Extraction, Enrichment, Solubilization, and Digestion Techniques for Membrane Proteomics // J. Proteome Res. American Chemical Society, 2016.
 Vol. 15, № 4. P. 1243–1252.
- 18. Sechi S., Chait B.T. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification // Anal. Chem. 1998. Vol. 70, № 24. P. 5150–5158.
- 19. Crestfield A.M., Moore S., Stein W.H. The preparation and enzymatic hydrolysis of reduced and S-carboxymethylated proteins. // J. Biological Chem. 1963. Vol. 238, № 2. P. 622–627.
- Hao P. et al. Detection, Evaluation and Minimization of Nonenzymatic Deamidation in Proteomic Sample Preparation // Mol. Cell. Proteomics. 2011. Vol. 10, № 10. P. O111.009381.
- Boja E.S., Fales H.M. Overalkylation of a protein digest with iodoacetamide. // Anal. Chem.
 2001. Vol. 73, № 15. P. 3576–3582.
- 22. Leiros H.-K.S. et al. Trypsin specificity as elucidated by LIE calculations, X-ray structures, and association constant measurements. // Protein Sci. Wiley, 2004. Vol. 13, № 4. P. 1056–1070.
- 23. Trevisan-Silva D. et al. A multi-protease, multi-dissociation, bottom-up-to-top-down proteomic view of the Loxosceles intermedia venom // Sci. Data. Nature Publishing Groups, 2017. Vol. 4, № 1. P. 1–7.
- 24. Naryzhny S.N. et al. Virtual-Experimental 2DE Approach in Chromosome-Centric Human Proteome Project // J. Proteome Res. 2016. Vol. 15, № 2. P. 525–530.
- 25. Shevchenko A. et al. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels // Anal. Chem. 1996. Vol. 68, № 5. P. 850–858.
- Klont F. et al. Assessment of Sample Preparation Bias in Mass Spectrometry-Based Proteomics
 // Anal. Chem. 2018. Vol. 90, № 8. P. 5405–5413.
- Verheggen K. et al. Database search engines: Paradigms, challenges and solutions // Advances in Experimental Medicine and Biology. 2016. Vol. 919. P. 147–156.
- Silva A.S.C. et al. Accurate peptide fragmentation predictions allow data driven approaches to replace and improve upon proteomics search engine scoring functions // Bioinformatics / ed. Wren J. Oxford University Press (OUP), 2019. Vol. 35, № 24. P. 5243–5248.
- 29. Tiwary S. et al. High-quality MS/MS spectrum prediction for data-dependent and dataindependent acquisition data analysis // Nat. Methods. Nature Publishing Group, 2019. Vol. 16,

№ 6. P. 519–525.

- 30. Aebersold R., Mann M. Mass-spectrometric exploration of proteome structure and function // Nature. Nature Publishing Group, 2016. Vol. 537, № 7620. P. 347–355.
- Borràs E., Sabidó E. What is targeted proteomics? A concise revision of targeted acquisition and targeted data analysis in mass spectrometry // Proteomics. 2017. Vol. 17, № 17–18. P. 1700180.
- 32. Kelleher N.L. Top-down proteomics // Anal. Chem. 2004. Vol. 76, № 11. P. 197A-203A.
- 33. Ansong C. et al. Proteogenomics: needs and roles to be filled by proteomics in genome annotation // Briefings Funct. Genomics Proteomics. 2008. Vol. 7, № 1. P. 50–62.
- 34. Nesvizhskii A.I. Proteogenomics: concepts, applications and computational strategies // Nat.
 Methods. 2014. Vol. 11, № 11. P. 1114–1125.
- 35. Karpova M.A. et al. Exome-driven characterization of the cancer cell lines at the proteome level: the NCI-60 case study. // J. Proteome Res. 2014. Vol. 13, № 12. P. 5551–5560.
- 36. Levitsky L.I. et al. Adenosine- to- Inosine RNA Editing in Mouse and Human Brain Proteomes
 // Proteomics. 2019. Vol. 19, № 23. P. 1900195.
- 37. Zhu F.Y. et al. Proteogenomic analysis reveals alternative splicing and translation as part of the abscisic acid response in Arabidopsis seedlings // Plant J. Blackwell Publishing Ltd, 2017. Vol. 91, № 3. P. 518–533.
- 38. Elias J.E., Gygi S.P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry // Nat. Methods. Nat Methods, 2007. Vol. 4, № 3. P. 207–214.
- Benjamini Y., Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing // Journal of the Royal Statistical Society. Series B (Methodological). WileyRoyal Statistical Society, 1995. Vol. 57. P. 289–300.
- 40. Tyanova S., Temu T., Cox J. The MaxQuant computational platform for mass spectrometrybased shotgun proteomics. // Nat. Protoc. Nature Publishing Group, 2016. Vol. 11, № 12. P. 2301–2319.
- 41. Vaudel M. et al. PeptideShaker enables reanalysis of MS-derived proteomics data sets: To the editor // Nature Biotechnology. Nature Publishing Group, 2015. Vol. 33, № 1. P. 22–24.
- 42. Ivanov M. V. et al. Empirical multidimensional space for scoring peptide spectrum matches in shotgun proteomics // J. Proteome Res. 2014. Vol. 13, № 4. P. 1911–1920.
- Kryukov G. V., Pennacchio L.A., Sunyaev S.R. Most rare missense alleles are deleterious in humans: Implications for complex disease and association studies // Am. J. Hum. Genet. University of Chicago Press, 2007. Vol. 80, № 4. P. 727–739.
- 44. Ruggles K. V. et al. An Analysis of the Sensitivity of Proteogenomic Mapping of Somatic

Mutations and Novel Splicing Events in Cancer // Mol. Cell. Proteomics. 2016. Vol. 15, № 3. P. 1060–1071.

- Hu J.X., Zhao H., Zhou H.H. False discovery rate control with groups // J. Am. Stat. Assoc.
 NIH Public Access, 2010. Vol. 105, № 491. P. 1215–1227.
- 46. Woo S. et al. Proteogenomic strategies for identification of aberrant cancer peptides using largescale next-generation sequencing data // Proteomics. 2014. Vol. 14, № 23–24. P. 2719–2730.
- 47. Levitsky L.I. et al. Unbiased False Discovery Rate Estimation for Shotgun Proteomics Based on the Target-Decoy Approach // J. Proteome Res. 2017. Vol. 16, № 2. P. 393–397.
- 48. Ikonoraou M.G., Blades A.T., Kebarle P. Investigations of the Electrospray Interface for Liquid Chromatography/Mass Spectrometry // Anal. Chem. American Chemical Society, 1990. Vol. 62, № 9. P. 957–967.
- 49. Jones M.N. Surfactant interactions with biomembranes and proteins // Chem. Soc. Rev. The Royal Society of Chemistry, 1992. Vol. 21, № 2. P. 127.
- Shapiro A.L., Viñuela E., V. Maizel J. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels // Biochem. Biophys. Res. Commun. 1967. Vol. 28, № 5. P. 815–820.
- Kachuk C., Doucette A.A. The benefits (and misfortunes) of SDS in top-down proteomics // J. Proteomics. 2018. Vol. 175. P. 75–86.
- León I.R. et al. Quantitative Assessment of In-solution Digestion Efficiency Identifies Optimal Protocols for Unbiased Protein Analysis // Mol. Cell. Proteomics. 2013. Vol. 12, № 10. P. 2992–3005.
- 53. Loo R.R.O., Dales N., Andrews P.C. Surfactant effects on protein structure examined by electrospray ionization mass spectrometry // Protein Sci. Wiley-Blackwell, 1994. Vol. 3, № 11. P. 1975–1983.
- 54. Lin Y. et al. Sodium-deoxycholate-assisted tryptic digestion and identification of proteolytically resistant proteins // Anal. Biochem. Academic Press Inc., 2008. Vol. 377, № 2. P. 259–266.
- 55. Zhou J. et al. Evaluation of the application of sodium deoxycholate to proteomic analysis of rat hippocampal plasma membrane // J. Proteome Res. 2006. Vol. 5, № 10. P. 2547–2553.
- 56. Saveliev S. V. et al. Mass spectrometry compatible surfactant for optimized in-gel protein digestion // Anal. Chem. American Chemical Society, 2013. Vol. 85, № 2. P. 907–914.
- 57. Yu Y.-Q. et al. Enzyme-Friendly, Mass Spectrometry-Compatible Surfactant for In-Solution Enzymatic Digestion of Proteins // Anal. Chem. 2003. Vol. 75, № 21. P. 6023–6028.
- 58. Cohn E.J. et al. Preparation and Properties of Serum and Plasma Proteins. IV. A System for the Separation into Fractions of the Protein and Lipoprotein Components of Biological Tissues and Fluids // J. Am. Chem. Soc. American Chemical Society, 1946. Vol. 68, № 3. P. 459–475.

- 59. Schmelter C. et al. Comparison of two solid-phase extraction (SPE) methods for the identification and quantification of porcine retinal protein markers by LC-MS/MS // Int. J. Mol. Sci. MDPI AG, 2018. Vol. 19, № 12.
- 60. Wiśniewski J.R. et al. Universal sample preparation method for proteome analysis // Nat. Methods. 2009. Vol. 6, № 5. P. 359–362.
- 61. Pirmoradian M. et al. Rapid and deep human proteome analysis by single-dimension shotgun proteomics // Mol. Cell. Proteomics. 2013. Vol. 12, № 11. P. 3330–3338.
- Kaleja P. et al. Evaluation and improvement of protein extraction methods for analysis of skin proteome by noninvasive tape stripping // J. Proteomics. Elsevier B.V., 2020. Vol. 217. P. 103678.
- 63. Ji Y. et al. Surfactant-Induced Artifacts during Proteomic Sample Preparation // Anal. Chem.
 2015. Vol. 87, № 11. P. 5500–5504.
- 64. Manza L.L. et al. Sample preparation and digestion for proteomic analyses using spin filters //
 Proteomics. 2005. Vol. 5, № 7. P. 1742–1745.
- Doucette A., Crowell A. Precipitation of Detergent-Containing Samples for Top-Down and Bottom-Up Proteomics // Proteomics Technologies and Applications. IntechOpen, 2019. Vol. i. P. 13.
- 66. Nagaraj N. et al. Detergent-Based but Gel-Free Method Allows Identification of Several Hundred Membrane Proteins in Single LC-MS Runs // J. Proteome Res. 2008. Vol. 7, № 11. P. 5028–5032.
- 67. Potriquet J. et al. A modified FASP protocol for high-throughput preparation of protein samples for mass spectrometry // PLoS One / ed. Jacobs J.M. Public Library of Science, 2017. Vol. 12, № 7. P. e0175967.
- 68. Kulak N.A. et al. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells // Nat. Methods. 2014. Vol. 11, № 3. P. 319–324.
- Ploug M., Stoffer B., Jensen A.L. In situ alkylation of cysteine residues in a hydrophobic membrane protein immobilized on polyvinylidene difluoride membranes by electroblotting prior to microsequence and amino acid analysis // Electrophoresis. 1992. Vol. 13, № 1. P. 148– 153.
- 70. Gundry R.L. et al. Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow // Current Protocols in Molecular Biology. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2009. Vol. 77, № 5. P. 342–355.
- Cleland W.W. Dithiothreitol, a New Protective Reagent for SH Groups * // Biochemistry.
 American Chemical Society, 1964. Vol. 3, № 4. P. 480–482.
- 72. Getz E.B. et al. A comparison between the sulfhydryl reductants tris(2- carboxyethyl)phosphine

and dithiothreitol for use in protein biochemistry // Anal. Biochem. 1999. Vol. 273, № 1. P. 73–80.

- Rüegg U.T., Rudinger J. Reductive cleavage of cystine disulfides with tributylphosphine // Methods in Enzymology. 1977. Vol. 47. P. 111–116.
- 74. Suttapitugsakul S. et al. Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS-based proteomics // Mol. Biosyst. 2017. Vol. 13, № 12. P. 2574–2582.
- 75. Pappin D.J.C., Hojrup P., Bleasby A.J. Rapid identification of proteins by peptide-mass fingerprinting // Curr. Biol. 1993. Vol. 3, № 6. P. 327–332.
- 76. Sebastiano R. et al. A new deuterated alkylating agent for quantitative proteomics // Rapid Commun. Mass Spectrom. 2003. Vol. 17, № 21. P. 2380–2386.
- Brune D.C. Alkylation of cysteine with acrylamide for protein sequence analysis // Anal.
 Biochem. 1992. Vol. 207, № 2. P. 285–290.
- 78. Karala A.-R., Ruddock L.W. Does S-methyl methanethiosulfonate trap the thiol-disulfide state of proteins? // Antioxidants Redox Signal. 2007. Vol. 9, № 4. P. 527–531.
- 79. Paulech J., Solis N., Cordwell S.J. Characterization of reaction conditions providing rapid and specific cysteine alkylation for peptide-based mass spectrometry. // Biochim. Biophys. Acta -Proteins Proteomics. 2013. Vol. 1834, № 1. P. 372–379.
- 80. Xu P. et al. Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation // Cell. Cell Press, 2009. Vol. 137, № 1. P. 133–145.
- Kim S. et al. PubChem 2019 update: improved access to chemical data. // Nucleic Acids Res. Oxford University Press, 2019. Vol. 47, № D1. P. D1102–D1109.
- 82. Lundell N., Schreitmüller T. Sample preparation for peptide mapping-A pharmaceutical qualitycontrol perspective. // Anal. Biochem. 1999. Vol. 266. P. 31–47.
- 83. Yang Z., Attygalle A.B. LC/MS characterization of undesired products formed during iodoacetamide derivatization of sulfhydryl groups of peptides // J. Mass Spectrom. 2007. Vol. 42, № 2. P. 233–243.
- Woods A.G., Sokolowska I., Darie C.C. Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment // Biochem. Biophys. Res. Commun. 2012. Vol. 419, № 2. P. 305–308.
- 85. Lapko V.N., Smith D.L., Smith J.B. Identification of an artifact in the mass spectrometry of proteins derivatized with iodoacetamide // J. Mass Spectrom. 2000. Vol. 35, № 4. P. 572–575.
- 86. Schroeder M.J. et al. A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry // Anal. Chem. 2004. Vol. 76, № 13. P. 3590–3598.

- 87. Chick J.M. et al. A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. // Nat. Biotechnol. 2015. Vol. 33, № 7. P. 743–749.
- 88. Kong A.T. et al. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics // Nat. Methods. 2017. Vol. 14, № 5. P. 513–520.
- Verheggen K. et al. Anatomy and evolution of database search engines-a central component of mass spectrometry based proteomic workflows // Mass Spectrom. Rev. John Wiley and Sons Inc., 2017. P. 1–15.
- 90. Brot N., Weissbach H. Biochemistry and physiological role of methionine sulfoxide residues in proteins // Arch. Biochem. Biophys. 1983. Vol. 223, № 1. P. 271–281.
- 91. Veredas F.J., Cantón F.R., Aledo J.C. Methionine residues around phosphorylation sites are preferentially oxidized in vivo under stress conditions // Sci. Rep. Nature Publishing Group, 2017. Vol. 7, № 1. P. 40403.
- Berlett B.S., Stadtman E.R. Protein oxidation in aging, disease, and oxidative stress // Journal of Biological Chemistry. 1997. Vol. 272, № 33. P. 20313–20316.
- Lam X.M. et al. Site-Specific Tryptophan Oxidation Induced by Autocatalytic Reaction of Polysorbate 20 in Protein Formulation // Pharm. Res. 2011. Vol. 28, № 10. P. 2543–2555.
- Lagerwerf F.M. et al. Identification of oxidized methionine in peptides // Rapid Commun. Mass Spectrom. 1996. Vol. 10, № 15. P. 1905–1910.
- 95. Pilo A.L., McLuckey S.A. Oxidation of Methionine residues in polypeptide ions via gas-phase ion/ion chemistry // J. Am. Soc. Mass Spectrom. Elsevier Inc., 2014. Vol. 25, № 6. P. 1049– 1057.
- 96. Hains P.G., Robinson P.J. The Impact of Commonly Used Alkylating Agents on Artifactual Peptide Modification // J. Proteome Res. 2017. Vol. 16, № 9. P. 3443–3447.
- 97. Paulo J.A. Sample preparation for proteomic analysis using a GeLC-MS/MS strategy // J. Biol. Methods. 2016. Vol. 3, № 3. P. 45.
- 98. Udeshi N.D. et al. Large-scale identification of ubiquitination sites by mass spectrometry // Nat.
 Protoc. 2013. Vol. 8, № 10. P. 1950–1960.
- Su G., Jaffrey S.R. Proteomic identification of protein ubiquitination events // Biotechnol.
 Genet. Eng. Rev. 2013. Vol. 29, № 1. P. 73–109.
- Stes E. et al. A COFRADIC protocol to study protein ubiquitination // J. Proteome Res. 2014.
 Vol. 13, № 6. P. 3107–3113.
- 101. Zhang B. et al. DeMix workflow for efficient identification of cofragmented peptides in high resolution data-dependent tandem mass spectrometry // Mol. Cell. Proteomics. American Society for Biochemistry and Molecular Biology Inc., 2014. Vol. 13, № 11. P. 3211–3223.

- Moshkovskii S.A. et al. Identification of Single Amino Acid Substitutions in Proteogenomics // Biochem. 2018. Vol. 83, № 3. P. 250–258.
- 103. Ghezzi P., Bonetto V. Redox proteomics: Identification of oxidatively modified proteins // Proteomics. John Wiley & Sons, Ltd, 2003. Vol. 3, № 7. P. 1145–1153.
- Wojdyla K., Rogowska-Wrzesinska A. Differential alkylation-based redox proteomics -Lessons learnt // Redox Biol. Elsevier, 2015. Vol. 6. P. 240–252.
- Hill B.G. et al. Methods for the determination and quantification of the reactive thiol proteome
 // Free Radical Biology and Medicine. 2009. Vol. 47, № 6. P. 675–683.
- 106. Scheerlinck E. et al. Minimizing technical variation during sample preparation prior to labelfree quantitative mass spectrometry // Anal. Biochem. 2015. Vol. 490. P. 14–19.
- 107. Ying J. et al. Thiol oxidation in signaling and response to stress: Detection and quantification of physiological and pathophysiological thiol modifications // Free Radical Biology and Medicine. 2007. Vol. 43, № 8. P. 1099–1108.
- 108. Bubis J.A. et al. Validation of Peptide Identification Results in Proteomics Using Amino Acid Counting // Proteomics. 2018. Vol. 18, № 23. P. 1800117.
- 109. Tarasova I.A. et al. Profiling modifications for glioblastoma proteome using ultra-tolerant database search: Are the peptide mass shifts biologically relevant or chemically induced? // J. Proteomics. Elsevier, 2019. Vol. 191. P. 16–21.
- Gundlach H.G., Moore S., Stein W.H. The reaction of iodoacetate with methionine. // J. Biol.
 Chem. 1959. Vol. 234, № 7. P. 1761–1764.
- Chernobrovkin A.L. et al. Methionine to isothreonine conversion as a source of false discovery identifications of genetically encoded variants in proteogenomics // J. Proteomics. Elsevier B.V., 2015. Vol. 120. P. 169–178.
- Schiffter H.A. Medical Biotechnology and Healthcare // Comprehensive Biotechnology. Second edi / ed. Moo-Young M. Pergamon, 2011. P. 5320.
- 113. Aubagnac JL, El Amrani B, Devienne FM C.R. Characterization of leucine and isoleucine by use of Fab ionization method in tandem mass spectrometry // Org. Mass Spectrom. 1985. Vol. 20, № 6. P. 428–429.
- 114. Kjeldsen F. et al. Distinguishing of Ile/Leu amino acid residues in the PP3 protein by (hot) electron capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. // Anal. Chem. 2003. Vol. 75, № 6. P. 1267–1274.
- 115. Armirotti A., Millo E., Damonte G. How to Discriminate Between Leucine and Isoleucine by Low Energy ESI-TRAP MSn // J. Am. Soc. Mass Spectrom. 2007. Vol. 18. P. 57–63.
- 116. Lebedev A.T. et al. Discrimination of leucine and isoleucine in peptides sequencing with orbitrap fusion mass spectrometer // Anal. Chem. 2014. Vol. 86. P. 7017–7022.

- 117. Lloyd JR, Cotter ML, Ohori D D.D. Distinction of alpha- aspartyl and beta-aspartyl and alphaglutamyl and gamma-glutamyl- transferase peptides by fast atom bombardment tandem mass spectrometry // Biomed. Environ. Mass Spectrom. 1988. Vol. 15. P. 399–402.
- 118. Lehmann W.D. et al. Analysis of isoaspartate in peptides by electrospray tandem mass spectrometry // Protein Sci. 2000. Vol. 9, № 11. P. 2260–2268.
- 119. Cournoyer J.J. et al. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. // Protein Sci. 2005. Vol. 14, № 2. P. 452–463.
- O'Connor P.B. et al. Differentiation of aspartic and isoaspartic acids using electron transfer dissociation // J. Am. Soc. Mass Spectrom. 2006. Vol. 17, № 1. P. 15–19.
- 121. Chan W.Y.K., Chan T.W.D., O'Connor P.B. Electron transfer dissociation with supplemental activation to differentiate aspartic and isoaspartic residues in doubly charged peptide cations // Journal of The American Society for Mass Spectrometry. Springer New York LLC, 2010.
- Sargaeva N.P. et al. Sequence-specific predictive chromatography to assist mass spectrometric analysis of asparagine deamidation and aspartate isomerization in peptides // Electrophoresis.
 2011. Vol. 32, № 15. P. 1962–1969.
- 123. Wang B. et al. Differentiation of α- or β-Aspartic Isomers in the Heptapeptides by the Fragments of [M + Na]+ Using Ion Trap Tandem Mass Spectrometry // J. Am. Soc. Mass Spectrom. 2011. Vol. 22, № 8. P. 1453–1462.
- 124. Adams C.M., Zubarev R.A. Distinguishing and Quantifying Peptides and Proteins Containing d-Amino Acids by Tandem Mass Spectrometry // Anal. Chem. 2005. Vol. 77, № 14. P. 4571– 4580.
- 125. Cournoyer J.J., Lin C., O'Connor P.B. Detecting Deamidation Products in Proteins by Electron Capture Dissociation // Anal. Chem. 2006. Vol. 78, № 4. P. 1264–1271.
- 126. Tsvetkov P.O. et al. Isomerization of the Asp7 residue results in zinc-induced oligomerization of Alzheimer's disease amyloid beta(1-16) peptide. // Chembiochem. 2008. Vol. 9, № 10. P. 1564–1567.
- Zubarev R.A., Kelleher N.L., McLafferty F.W. Electron capture dissociation of multiply charged protein cations. A nonergodic process // Journal of the American Chemical Society. 1998. Vol. 120, № 13. P. 3265–3266.
- 128. Syka J.E.P. et al. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. // Proc. Natl. Acad. Sci. U. S. A. 2004. Vol. 101, № 26. P. 9528–9533.
- 129. Li X., Lin C., O'Connor P.B. Glutamine deamidation: differentiation of glutamic acid and gamma-glutamic acid in peptides by electron capture dissociation. // Anal. Chem. American Chemical Society, 2010. Vol. 82, № 9. P. 3606–3615.
- 130. Vizcaíno J. et al. ProteomeXchange provides globally coordinated proteomics data submission

and dissemination // Nat Biotech. 2014. Vol. 32, № 3. P. 223–226.

- Moghaddas Gholami A. et al. Global proteome analysis of the NCI-60 cell line panel. // Cell Rep. The Authors, 2013. Vol. 4, № 3. P. 609–620.
- 132. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer // Nature. Nature Publishing Group, 2012. Vol. 487, № 7407. P. 330–337.
- 133. Craig R., Beavis R.C. TANDEM: Matching proteins with tandem mass spectra // Bioinformatics. 2004. Vol. 20, № 9. P. 1466–1467.
- 134. Neuhauser N. et al. High performance computational analysis of large-scale proteome data sets to assess incremental contribution to coverage of the human genome // J. Proteome Res. 2013. Vol. 12, № 6. P. 2858–2868.
- 135. Kessner D. et al. ProteoWizard: open source software for rapid proteomics tools development // Bioinformatics. 2008. Vol. 24, № 21. P. 2534–2536.
- 136. Levitsky L.I. et al. IdentiPy: An Extensible Search Engine for Protein Identification in Shotgun Proteomics: research-article // J. Proteome Res. American Chemical Society, 2018. Vol. 17, № 7. P. 2249–2255.
- 137. Sharma K. et al. Cell type– and brain region–resolved mouse brain proteome // Nat. Neurosci.
 2015. Vol. 18, № 12. P. 1819–1831.
- 138. Goloborodko A.A. et al. Pyteomics A python framework for exploratory data analysis and rapid software prototyping in proteomics // J. Am. Soc. Mass Spectrom. Springer-Verlag, 2013. Vol. 24, № 2. P. 301–304.
- Levitsky L.I. et al. Pyteomics 4.0: Five Years of Development of a Python Proteomics Framework // J. Proteome Res. 2019. Vol. 18, № 2. P. 709–714.
- 140. Abaan O.D. et al. The exomes of the NCI-60 panel: A genomic resource for cancer biology and systems pharmacology // Cancer Res. American Association for Cancer Research, 2013. Vol. 73, № 14. P. 4372–4382.
- 141. Hidaka Y. et al. Human adenine phosphoribosyltransferase deficiency. Demonstration of a single mutant allele common to the Japanese // J. Clin. Invest. 1988. Vol. 81, № 3. P. 945–950.
- 142. Blanc V., Davidson N.O. C-to-U RNA editing: Mechanisms leading to genetic diversity // Journal of Biological Chemistry. J Biol Chem, 2003. Vol. 278, № 3. P. 1395–1398.
- 143. Donald V., Judith G.V. Biochemistry. 2nd ed. John Wiley & Sons., 1995. 115 p.
- 144. Schroeder W.A., Shelton J.B., Shelton J.R. An examination of conditions for the cleavage of polypeptide chains with cyanogen bromide: Application to catalase // Arch. Biochem. Biophys. Academic Press, 1969. Vol. 130, № C. P. 551–555.
- Lundblad R.L. Techniques in protein modification. London: CRC Press/Taylor & Francis Publishing Group, 1994. 243 p.

- 146. Cox J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ // Mol. Cell. Proteomics. American Society for Biochemistry and Molecular Biology Inc., 2014. Vol. 13, № 9. P. 2513–2526.
- 147. Biemann K., Martin S.A. Mass spectrometric determination of the amino acid sequence of peptides and proteins // Mass Spectrom. Rev. 1987. Vol. 6, № 1. P. 1–75.
- 148. Chelius D., Rehder D.S., Bondarenko P. V. Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. // Anal. Chem. 2005. Vol. 77, № 18. P. 6004–6011.
- 149. Winter D., Pipkorn R., Lehmann W.D. Separation of peptide isomers and conformers by ultra performance liquid chromatography // J. Sep. Sci. 2009. Vol. 32, № 8. P. 1111–1119.
- 150. Gorshkov A. V et al. Liquid chromatography at critical conditions: comprehensive approach to sequence-dependent retention time prediction. // Anal. Chem. American Chemical Society, 2006. Vol. 78, № 22. P. 7770–7777.
- 151. Gilar M. et al. Peptide retention prediction applied to proteomic data analysis. // Rapid Commun. Mass Spectrom. 2007. Vol. 21, № 17. P. 2813–2821.
- Zhou F. et al. Genome-scale proteome quantification by DEEP SEQ mass spectrometry. // Nat. Commun. 2013. Vol. 4. P. 2171.
- Wiśniewski J.R. et al. 'Shotgun' proteomic analyses without alkylation of cysteine // Anal. Chim. Acta. Elsevier B.V., 2020. Vol. 1100. P. 131–137.
- 154. Creasy D.M., Cottrell J.S. Unimod: Protein modifications for mass spectrometry // Proteomics.
 2004. Vol. 4, № 6. P. 1534–1536.
- 155. Wang Y. et al. Metastable decomposition at the peptide C-terminus: Possible use in protein identification // Rapid Commun. Mass Spectrom. John Wiley and Sons Ltd, 2020. Vol. 34, № 9.
- 156. Ree R., Varland S., Arnesen T. Spotlight on protein N-terminal acetylation // Experimental and Molecular Medicine. Nature Publishing Group, 2018. Vol. 50, № 7. P. 1–13.

ПРИЛОЖЕНИЕ

Таблица 1. Средние значения относительных интенсивностей (по шести техническим повторам) шести рассматриваемых ионов. Значения получены при помощи программы MaxQuant и использованы для автоматической классификации синтетических пептидов, содержащих треонин и изотреонин.

	1	•	
		относительная	относительная
пептил	ион р.	интенсивность	интенсивность
пентид	non on	для пептида с	для пептида с
		Tpe, %	изоТре, %
NNRPSEGPLQTR	b ₁₁	0,41	25,49
PAATAAGSLGSLK	b ₄	77,27	222,76
PNNTNNGSLGSLK	b ₄	25,98	118,45
PDDTDDGSLGSLK	b ₄	72,94	147,93
PEETEEGSLGSLK	b ₄	35,12	135,26
PQQTQQGSLGSLK	b ₄	73,71	156,8
PGGTGGGSLGSLK	b ₄	38,72	288,12
PIITIIGSLGSLK	b ₄	79,06	116,39
PLLTLLGSLGSLK	b_4	64,48	143,92
PMMTMMGSLGSLK	b ₄	43,86	127,71
PFFTFFGSLGSLK	b_4	19	79,79
PPPTPPGSLGSLK	b ₄	58,74	173,64
PSSTSSGSLGSLK	b ₄	21,62	137,37
PTTTTTGSLGSLK	b ₄	22,76	108,82
PYYTYYGSLGSLK	b ₄	11,66	64,3
PVVTVVGSLGSLK	b ₄	58,4	139,1

пептид	ион у _п	относительная интенсивность для пептида с Тре, %	относительная интенсивность для пептида с изоТре, %
MSSPEDDSDTKR	y ₃	418,51	729,36
DNIQGITKPAIR	y ₆	283,07	205,97
KTVGVEPAADGK	y ₁₁	10,73	0
NNRPSEGPLQTR	y ₂	115	90,91
QDPSVLHTEEMR	y 5	437,94	465,87
PAATAAGSLGSLK	y 10	20,71	14,87
PNNTNNGSLGSLK	y 10	21,33	22,01
PDDTDDGSLGSLK	y ₁₀	23,53	23,66
PEETEEGSLGSLK	y ₁₀	38,32	32,11
PQQTQQGSLGSLK	y 10	63,97	53,22
PGGTGGGSLGSLK	y ₁₀	11,63	10,2
PIITIIGSLGSLK	y ₁₀	25,71	13,8
PLLTLLGSLGSLK	y 10	25,02	13,58
PMMTMMGSLGSLK	y ₁₀	23,7	11,71
PFFTFFGSLGSLK	y ₁₀	18,46	5,31
PPPTPPGSLGSLK	y ₁₀	55,01	8,1
PSSTSSGSLGSLK	y ₁₀	24,45	16,53
PTTTTTGSLGSLK	y ₁₀	35,1	23,82
PYYTYYGSLGSLK	y ₁₀	9,74	4,23

PVVTVVGSLGSLK	y ₁₀	26,41	14,47
---------------	------------------------	-------	-------

пептид	ион b _{n-2}	относительная интенсивность для пептида с Тре, %	относительная интенсивность для пептида с изоТре, %
DNIQGITKPAIR	b ₅	1,75	6,72
NNRPSEGPLQTR	b 9	50,27	45,8
PAATAAGSLGSLK	b ₂	995,57	953
PNNTNNGSLGSLK	b ₂	574,25	560,81
PDDTDDGSLGSLK	b ₂	304,01	287,63
PEETEEGSLGSLK	b ₂	567,27	516,85
PQQTQQGSLGSLK	b ₂	648,53	605,49
PGGTGGGSLGSLK	b ₂	602,78	532,89
PIITIIGSLGSLK	b ₂	604,57	501,55
PLLTLLGSLGSLK	b ₂	628,6	644,02
PMMTMMGSLGSLK	b ₂	603,77	594,27
PFFTFFGSLGSLK	b ₂	547,38	509,92
PPPTPPGSLGSLK	b ₂	203,41	215,57
PSSTSSGSLGSLK	b ₂	249,13	278,73
PTTTTTGSLGSLK	b ₂	236,03	252,3
PYYTYYGSLGSLK	b ₂	501,49	460,14
PVVTVVGSLGSLK	b ₂	659,98	678,98

		относительная	относительная
паптип		интенсивность	интенсивность
пентид	ион y _{n-1}	для пептида с	для пептида с
		Tpe, %	изоТре, %
MSSPEDDSDTKR	y 2	397,4	418,25
DNIQGITKPAIR	y 5	111,76	213,52
KTVGVEPAADGK	y ₁₀	7,76	0
NNRPSEGPLQTR	y 1	267,64	504,28
QDPSVLHTEEMR	y 4	109,25	228,93
TSYAQHQQVR	y 9	10,23	41,36
PAATAAGSLGSLK	y 9	58,58	95,31
PNNTNNGSLGSLK	y 9	54,54	103,46
PDDTDDGSLGSLK	y 9	41,06	59,61
PEETEEGSLGSLK	y 9	29,48	58,98
PQQTQQGSLGSLK	y 9	15,4	22,3
PGGTGGGSLGSLK	y 9	140,87	189,2
PIITIIGSLGSLK	y 9	24	30,24
PLLTLLGSLGSLK	y 9	26,85	46,74
PMMTMMGSLGSLK	y 9	31,73	47,64
PFFTFFGSLGSLK	y 9	32,91	47,11
PPPTPPGSLGSLK	y 9	953,97	849,78
PSSTSSGSLGSLK	y 9	111,42	150,54
PTTTTTGSLGSLK	y 9	97,82	118,82
PYYTYYGSLGSLK	y 9	23,45	45,56
PVVTVVGSLGSLK	y 9	21,48	37,41

пептид	ион у _{п-2}	относительная интенсивность для пептида с Тре, %	относительная интенсивность для пептида с изоТре, %
DNIQGITKPAIR	y ₄	637,94	764,04
QDPSVLHTEEMR	y ₂	391,12	408,53
TSYAQHQQVR	y 7	292,95	300,86
PAATAAGSLGSLK	y ₈	103,22	125,96
PNNTNNGSLGSLK	y ₈	93,33	98,47
PDDTDDGSLGSLK	y 8	72,94	88,04
PEETEEGSLGSLK	y ₈	73,1	91,47
PQQTQQGSLGSLK	y 8	49,16	57,35
PGGTGGGSLGSLK	y 8	68,26	54,93
PIITIIGSLGSLK	y ₈	82,82	77,73
PLLTLLGSLGSLK	y ₈	74,78	91,67
PMMTMMGSLGSLK	y 8	87,84	90,64
PFFTFFGSLGSLK	y ₈	85	89,13
PPPTPPGSLGSLK	y ₈	527,71	527,21
PSSTSSGSLGSLK	y 8	97,31	98,01
PTTTTTGSLGSLK	y ₈	135,97	138,76
PYYTYYGSLGSLK	y ₈	81,81	90,91
PVVTVVGSLGSLK	y 8	85,19	97,55

пептид	ион у _{n+1}	относительная интенсивность для пептида с Тре, %	относительная интенсивность для пептида с изоТре, %
MSSPEDDSDTKR	y 4	48,67	14,71
DNIQGITKPAIR	y ₇	50,74	33,1
NNRPSEGPLQTR	y ₃	256,03	221,67
QDPSVLHTEEMR	y 6	446,62	421,68
PAATAAGSLGSLK	y ₁₁	15,88	16,28
PNNTNNGSLGSLK	y ₁₁	7,69	11,67
PDDTDDGSLGSLK	y 11	12,49	13,85
PEETEEGSLGSLK	y ₁₁	16,2	12
PGGTGGGSLGSLK	y 11	20,32	18,96
PIITIIGSLGSLK	y 11	4	1,06
PLLTLLGSLGSLK	y ₁₁	7,88	3,68
PMMTMMGSLGSLK	y ₁₁	5,14	2,65
PFFTFFGSLGSLK	y 11	0,64	0,38
PPPTPPGSLGSLK	y ₁₁	26,83	13,9
PSSTSSGSLGSLK	y ₁₁	24,64	22,29
PTTTTTGSLGSLK	y 11	19	7,75
PYYTYYGSLGSLK	y ₁₁	2,93	0
PVVTVVGSLGSLK	y ₁₁	10,26	3,82













Рис. 2. Диаграммы, показывающие медианы и межквартильные интервалы нормализованных значений числа идентифицированных пептидов после разных методов алкилирования; q образцы, в которых избыток алкилирующего агента был погашен добавлением DTT после алкилирования. Значения были нормализованы по наибольшему внутри 3 повторов пробоподготовки, что дало 9 значений, полученных после объединения технических повторов и повторов образца пробоподготовки. А: число пептидов с Цис; Б: число пептидов с Мет.

Таблица 2. Значения р паркого критерия Стьюдента и кратное изменение нормированных значений количества идентифицированных пептидов. Значения нормализованы в каждом повторе пробоподготовки по наивысшему значению среди всех образцов (то есть методоы алкилирования). Термин «кратное изменение» используется как отношение большего значения к меньшему значению среднего значения всех повторов в одном образце.

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,000286	1,32	мет1>мет2
4-VP	CAM	0,00252	1,13	мет1<мет2
4-VP	CAMq	0,666	1,01	мет1<мет2
4-VP	control	0,126	1,1	мет1<мет2
4-VP	IAM	0,00519	1,26	мет1>мет2
4-VP	IAMq	0,223	1,05	мет1<мет2
4-VP	MMTS	0,564	1,03	мет1>мет2
4-VPq	CAM	0,00012	1,49	мет1<мет2
4-VPq	CAMq	0,00000944	1,33	мет1<мет2
4-VPq	control	0,00431	1,45	мет1<мет2
4-VPq	IAM	0,727	1,04	мет1<мет2
4-VPq	IAMq	0,000506	1,39	мет1<мет2
4-VPq	MMTS	0,0000473	1,28	мет1<мет2
CAM	CAMq	0,0102	1,13	мет1>мет2
CAM	control	0,349	1,03	мет1>мет2
CAM	IAM	0,00000535	1,43	мет1>мет2
CAM	IAMq	0,00516	1,08	мет1>мет2
CAM	MMTS	0,00241	1,17	мет1>мет2
CAMq	control	0,208	1,09	мет1<мет2
CAMq	IAM	0,0112	1,27	мет1>мет2
CAMq	IAMq	0,319	1,04	мет1<мет2
CAMq	MMTS	0,302	1,04	мет1>мет2
control	IAM	0,000000524	1,39	мет1>мет2
control	IAMq	0,281	1,05	мет1>мет2
control	MMTS	0,0788	1,13	мет1>мет2
IAM	IAMq	0,000321	1,33	мет1<мет2
IAM	MMTS	0,0392	1,23	мет1<мет2
IAMq	MMTS	0,129	1,08	мет1>мет2

2А. Общее число идентифицированных пептидов

2Б. Число идентифицированных пептидов с цистеином

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,799	1,023	мет1<мет2

4-VP	CAM	0,000277	1,651	мет1<мет2
4-VP	CAMq	0,329	1,156	мет1<мет2
4-VP	control	0,0000171	17,944	мет1>мет2
4-VP	IAM	0,168	1,211	мет1<мет2
4-VP	IAMq	0,0137	1,24	мет1<мет2
4-VP	MMTS	0,807	1,02	мет1>мет2
4-VPq	CAM	0,00000248	1,614	мет1<мет2
4-VPq	CAMq	0,0297	1,13	мет1<мет2
4-VPq	control	0,000000627	18,359	мет1>мет2
4-VPq	IAM	0,0812	1,184	мет1<мет2
4-VPq	IAMq	0,00728	1,212	мет1<мет2
4-VPq	MMTS	0,937	1,043	мет1>мет2
CAM	CAMq	0,000248	1,428	мет1>мет2
CAM	control	0,0000000000000527	29,625	мет1>мет2
CAM	IAM	0,00000648	1,363	мет1>мет2
CAM	IAMq	0,000464	1,331	мет1>мет2
CAM	MMTS	0,000683	1,683	мет1>мет2
CAMq	control	0,00000382	20,751	мет1>мет2
CAMq	IAM	0,595	1,047	мет1<мет2
CAMq	IAMq	0,341	1,072	мет1<мет2
CAMq	MMTS	0,0821	1,179	мет1>мет2
control	IAM	0,0000000798	21,733	мет1<мет2
control	IAMq	0,00000088	22,251	мет1<мет2
control	MMTS	0,0000729	17,601	мет1<мет2
IAM	IAMq	0,794	1,024	мет1<мет2
IAM	MMTS	0,154	1,235	мет1>мет2
IAMq	MMTS	0,014	1,264	мет1>мет2

2В. Число идентифицированных пептидов с метионином

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,00356	1,163	мет1>мет2
4-VP	CAM	0,0587	1,081	мет1<мет2
4-VP	CAMq	0,0652	1,041	мет1<мет2
4-VP	control	0,349	1,063	мет1<мет2
4-VP	IAM	0,00000949	5,495	мет1>мет2
4-VP	IAMq	0,249	1,058	мет1<мет2
4-VP	MMTS	0,0607	1,109	мет1>мет2
4-VPq	CAM	0,00363	1,257	мет1<мет2
4-VPq	CAMq	0,0000196	1,211	мет1<мет2
4-VPq	control	0,062	1,236	мет1<мет2
4-VPq	IAM	0,0000304	4,724	мет1>мет2

Î.	1	1		1
4-VPq	IAMq	0,012	1,23	мет1<мет2
4-VPq	MMTS	0,224	1,049	мет1<мет2
CAM	CAMq	0,304	1,038	мет1>мет2
CAM	control	0,662	1,017	мет1>мет2
CAM	IAM	0,0000000247	5,939	мет1>мет2
CAM	IAMq	0,305	1,022	мет1>мет2
CAM	MMTS	0,000047	1,198	мет1>мет2
CAMq	control	0,761	1,021	мет1<мет2
CAMq	IAM	0,00000337	5,72	мет1>мет2
CAMq	IAMq	0,711	1,016	мет1<мет2
CAMq	MMTS	0,00561	1,154	мет1>мет2
control	IAM	0,00000031	5,84	мет1>мет2
control	IAMq	0,92	1,005	мет1>мет2
control	MMTS	0,0174	1,178	мет1>мет2
IAM	IAMq	0,00000000161	5,812	мет1<мет2
IAM	MMTS	0,00000125	4,957	мет1<мет2
IAMq	MMTS	0,00114	1,173	мет1>мет2

2Г. Отношение числа пептидов с цистемном к общему числу идентифицированных пептидов

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,0207	1,373	мет1<мет2
4-VP	CAM	0,00517	1,424	мет1<мет2
4-VP	CAMq	0,418	1,121	мет1<мет2
4-VP	control	0,0000414	19,343	мет1>мет2
4-VP	IAM	0,00475	1,493	мет1<мет2
4-VP	IAMq	0,172	1,159	мет1<мет2
4-VP	MMTS	0,879	1,03	мет1>мет2
4-VPq	CAM	0,324	1,037	мет1<мет2
4-VPq	CAMq	0,0389	1,225	мет1>мет2
4-VPq	control	0,000000161	26,55	мет1>мет2
4-VPq	IAM	0,118	1,087	мет1<мет2
4-VPq	IAMq	0,129	1,185	мет1>мет2
4-VPq	MMTS	0,0419	1,414	мет1>мет2
CAM	CAMq	0,0000154	1,27	мет1>мет2
CAM	control	0,0000000002 65	27,538	мет1>мет2
CAM	IAM	0,00981	1,048	мет1<мет2
CAM	IAMq	0,000285	1,229	мет1>мет2
CAM	MMTS	0,00043	1,466	мет1>мет2
CAMq	control	0,0000000171	21,68	мет1>мет2
CAMq	IAM	0,000000289	1,332	мет1<мет2
CAMq	IAMq	0,494	1,034	мет1<мет2

CAMq	MMTS	0,0804	1,155	мет1>мет2
control	IAM	0,0000000028 3	28,87	мет1<мет2
control	IAMq	0,00000226	22,413	мет1<мет2
control	MMTS	0,0000321	18,778	мет1<мет2
IAM	IAMq	0,0000594	1,288	мет1>мет2
IAM	MMTS	0,000108	1,537	мет1>мет2
IAMq	MMTS	0,00204	1,194	мет1>мет2

2Д. Отношение числа пептидов с метионином к общему числу идентифицированных пептидов

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,0000128	1,141	мет1<мет2
4-VP	CAM	0,000605	1,048	мет1>мет2
4-VP	CAMq	0,00414	1,036	мет1<мет2
4-VP	control	0,00226	1,038	мет1>мет2
4-VP	IAM	0,000000132	4,089	мет1>мет2
4-VP	IAMq	0,35	1,008	мет1<мет2
4-VP	MMTS	0,0211	1,076	мет1>мет2
4-VPq	CAM	0,0000506	1,195	мет1>мет2
4-VPq	CAMq	0,00063	1,101	мет1>мет2
4-VPq	control	0,000000362	1,184	мет1>мет2
4-VPq	IAM	0,00000185	4,665	мет1>мет2
4-VPq	IAMq	0,000133	1,132	мет1>мет2
4-VPq	MMTS	0,000132	1,228	мет1>мет2
CAM	CAMq	0,00000164	1,085	мет1<мет2
CAM	control	0,515	1,01	мет1<мет2
CAM	IAM	0,000000569	3,903	мет1>мет2
CAM	IAMq	0,0000105	1,056	мет1<мет2
CAM	MMTS	0,359	1,028	мет1>мет2
CAMq	control	0,00011	1,075	мет1>мет2
CAMq	IAM	0,000000555	4,237	мет1>мет2
CAMq	IAMq	0,000407	1,028	мет1>мет2
CAMq	MMTS	0,000174	1,115	мет1>мет2
control	IAM	0,000000526	3,94	мет1>мет2
control	IAMq	0,00306	1,046	мет1<мет2
control	MMTS	0,147	1,037	мет1>мет2
IAM	IAMq	0,000000631	4,12	мет1<мет2
IAM	MMTS	0,00000725	3,798	мет1<мет2
IAMq	MMTS	0,00162	1,085	мет1>мет2

2Е. Общее число спектральных идентификаций (PSM)

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,000198	1,256	мет1>мет2
4-VP	CAM	0,0136	1,127	мет1<мет2
4-VP	CAMq	0,637	1,01	мет1>мет2
4-VP	control	0,106	1,136	мет1<мет2
4-VP	IAM	0,0768	1,16	мет1>мет2
4-VP	IAMq	0,496	1,038	мет1<мет2
4-VP	MMTS	0,0217	1,16	мет1>мет2
4-VPq	CAM	0,000428	1,415	мет1<мет2
4-VPq	CAMq	0,0000704	1,243	мет1<мет2
4-VPq	control	0,0137	1,426	мет1<мет2
4-VPq	IAM	0,507	1,083	мет1<мет2
4-VPq	IAMq	0,00393	1,304	мет1<мет2
4-VPq	MMTS	0,0393	1,082	мет1<мет2
CAM	CAMq	0,00678	1,139	мет1>мет2
CAM	control	0,886	1,008	мет1<мет2
CAM	IAM	0,00335	1,307	мет1>мет2
CAM	IAMq	0,0065	1,086	мет1>мет2
CAM	MMTS	0,0000493	1,308	мет1>мет2
CAMq	control	0,117	1,147	мет1<мет2
CAMq	IAM	0,152	1,148	мет1>мет2
CAMq	IAMq	0,307	1,049	мет1<мет2
CAMq	MMTS	0,00535	1,149	мет1>мет2
control	IAM	0,0000349	1,317	мет1>мет2
control	IAMq	0,233	1,094	мет1>мет2
control	MMTS	0,00966	1,318	мет1>мет2
IAM	IAMq	0,0665	1,204	мет1<мет2
IAM	MMTS	0,859	1	мет1>мет2
IAMq	MMTS	0,00274	1,205	мет1>мет2

2Ж. Число спектральных идентификаций (PSM) с цистеином

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,682	1,042	мет1>мет2
4-VP	CAM	0,00211	1,578	мет1<мет2
4-VP	CAMq	0,686	1,065	мет1<мет2
4-VP	control	0,0000904	24,245	мет1>мет2
4-VP	IAM	0,264	1,273	мет1<мет2
4-VP	IAMq	0,111	1,158	мет1<мет2
4-VP	MMTS	0,692	1,131	мет1>мет2
4-VPq	CAM	0,0000633	1,644	мет1<мет2
4-VPq	CAMq	0,161	1,11	мет1<мет2

4-VPq	control	0,00000417	23,27	мет1>мет2
4-VPq	IAM	0,1	1,326	мет1<мет2
4-VPq	IAMq	0,00373	1,206	мет1<мет2
4-VPq	MMTS	0,865	1,085	мет1>мет2
CAM	CAMq	0,00016	1,481	мет1>мет2
CAM	control	0,0000000000109	38,256	мет1>мет2
CAM	IAM	0,0194	1,239	мет1>мет2
CAM	IAMq	0,000282	1,363	мет1>мет2
CAM	MMTS	0,000107	1,784	мет1>мет2
CAMq	control	0,0000032	25,823	мет1>мет2
CAMq	IAM	0,0722	1,195	мет1<мет2
CAMq	IAMq	0,249	1,087	мет1<мет2
CAMq	MMTS	0,0162	1,205	мет1>мет2
control	IAM	0,00000145	30,867	мет1<мет2
control	IAMq	0,000000199	28,064	мет1<мет2
control	MMTS	0,00003	21,439	мет1<мет2
IAM	IAMq	0,466	1,1	мет1>мет2
IAM	MMTS	0,025	1,44	мет1>мет2
IAMq	MMTS	0,00362	1,309	мет1>мет2

23. Число спектральных идентификаций (PSM) с метионином

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,00379	1,132	мет1>мет2
4-VP	CAM	0,0634	1,094	мет1<мет2
4-VP	CAMq	0,325	1,027	мет1<мет2
4-VP	control	0,176	1,133	мет1<мет2
4-VP	IAM	0,000000184	5,43	мет1>мет2
4-VP	IAMq	0,328	1,059	мет1<мет2
4-VP	MMTS	0,0109	1,225	мет1>мет2
4-VPq	CAM	0,00729	1,239	мет1<мет2
4-VPq	CAMq	0,000564	1,163	мет1<мет2
4-VPq	control	0,0886	1,283	мет1<мет2
4-VPq	IAM	0,00000189	4,796	мет1>мет2
4-VPq	IAMq	0,0251	1,2	мет1<мет2
4-VPq	MMTS	0,529	1,082	мет1>мет2
CAM	CAMq	0,114	1,065	мет1>мет2
CAM	control	0,643	1,036	мет1<мет2
CAM	IAM	0,0000000000857	5,942	мет1>мет2
CAM	IAMq	0,152	1,033	мет1>мет2
CAM	MMTS	0,00000756	1,34	мет1>мет2
CAMq	control	0,329	1,104	мет1<мет2
CAMq	IAM	0,0000000318	5,577	мет1>мет2
CAMq	IAMq	0,476	1,031	мет1<мет2
CAMq	MMTS	0,000636	1,258	мет1>мет2
control	IAM	0,00000422	6,155	мет1>мет2
control	IAMq	0,465	1,07	мет1>мет2

control	MMTS	0,00663	1,388	мет1>мет2
IAM	IAMq	0,00000000434	5,753	мет1<мет2
IAM	MMTS	0,0000000293	4,433	мет1<мет2
IAMq	MMTS	0,0000198	1,298	мет1>мет2

2И.	Отношение	числа РЅМ с ц	истеином к общ	ему числу PSM
-----	-----------	---------------	----------------	---------------

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,00379	1,205	мет1<мет2
4-VP	CAM	0,0634	1,367	мет1<мет2
4-VP	CAMq	0,325	1,046	мет1<мет2
4-VP	control	0,176	27,443	мет1>мет2
4-VP	IAM	0,000000184	1,424	мет1<мет2
4-VP	IAMq	0,328	1,091	мет1<мет2
4-VP	MMTS	0,0109	1,018	мет1>мет2
4-VPq	CAM	0,00729	1,135	мет1<мет2
4-VPq	CAMq	0,000564	1,151	мет1>мет2
4-VPq	control	0,0886	33,061	мет1>мет2
4-VPq	IAM	0,00000189	1,182	мет1<мет2
4-VPq	IAMq	0,0251	1,104	мет1>мет2
4-VPq	MMTS	0,529	1,227	мет1>мет2
CAM	CAMq	0,114	1,307	мет1>мет2
CAM	control	0,643	37,524	мет1>мет2
CAM	IAM	0,0000000000857	1,041	мет1<мет2
CAM	IAMq	0,152	1,254	мет1>мет2
CAM	MMTS	0,00000756	1,392	мет1>мет2
CAMq	control	0,329	28,715	мет1>мет2
CAMq	IAM	0,0000000318	1,361	мет1<мет2
CAMq	IAMq	0,476	1,042	мет1<мет2
CAMq	MMTS	0,000636	1,065	мет1>мет2
control	IAM	0,00000422	39,068	мет1<мет2
control	IAMq	0,465	29,935	мет1<мет2
control	MMTS	0,00663	26,954	мет1<мет2
IAM	IAMq	0,00000000434	1,305	мет1>мет2
IAM	MMTS	0,0000000293	1,449	мет1>мет2
IAMq	MMTS	0,0000198	1,111	мет1>мет2

2К.	Отношение	числа PSM	[с метион	нином к	общему	числу PSM	[
						2	

метод_1	метод_2	значение р	кратное изменение	характер изменения
4-VP	4-VPq	0,0646	1,119	мет1<мет2
4-VP	CAM	0,0189	1,033	мет1>мет2
4-VP	CAMq	0,776	1,037	мет1<мет2
4-VP	control	0,0000949	1,005	мет1>мет2
4-VP	IAM	0,0195	4,285	мет1>мет2
4-VP	IAMq	0,48	1,017	мет1<мет2
4-VP	MMTS	0,854	1,056	мет1>мет2
4-VPq	CAM	0,0414	1,155	мет1>мет2
4-VPq	CAMq	0,232	1,079	мет1>мет2
4-VPq	control	0,00000967	1,125	мет1>мет2
4-VPq	IAM	0,0396	4,794	мет1>мет2
4-VPq	IAMq	0,369	1,1	мет1>мет2

4-VPq	MMTS	0,227	1,182	мет1>мет2
CAM	CAMq	0,0000549	1,071	мет1<мет2
CAM	control	0,000000000000625	1,027	мет1<мет2
CAM	IAM	0,0909	4,15	мет1>мет2
CAM	IAMq	0,000524	1,05	мет1<мет2
CAM	MMTS	0,000898	1,023	мет1>мет2
CAMq	control	0,0000000235	1,042	мет1>мет2
CAMq	IAM	0,0000000808	4,443	мет1>мет2
CAMq	IAMq	0,387	1,02	мет1>мет2
CAMq	MMTS	0,206	1,095	мет1>мет2
control	IAM	0,0000000000735	4,262	мет1>мет2
control	IAMq	0,000000433	1,022	мет1<мет2
control	MMTS	0,00000262	1,051	мет1>мет2
IAM	IAMq	0,0000444	4,357	мет1<мет2
IAM	MMTS	0,0000467	4,057	мет1<мет2
IAMq	MMTS	0,0426	1,074	мет1>мет2

Таблица 3. Абсолютные ненормализованные значения по всем образцам и повторам 3А. Общее число пептидов

образец	повтор I	пептиды I	повтор II	пептиды II	повтор III	пептиды III
control	60_1	5009	68_1	8862	92_1	3339
control	60_2	5258	68_2	9065	92_2	3432
control	60_3	5040	68_3	9115	92_3	3487
IAM	61_1	3761	69_1	6283	93_1	2630
IAM	61_2	3958	69_2	6110	93_2	2394
IAM	61_3	3740	69_3	6230	93_3	2344
IAMq	62_1	4836	70_1	10087	94_1	3028
IAMq	62_2	4949	70_2	9959	94_2	2930
IAMq	62_3	4484	70_3	9607	94_3	2737
4-VP	63_1	5100	71_1	9549	95_1	2499
4-VP	63_2	5240	71_2	8927	95_2	2410
4-VP	63_3	5196	71_3	8911	95_3	2406
4-VPq	64_1	3726	72_1	7458	96_1	1292
4-VPq	64_2	3666	72_2	8058	96_2	NA
4-VPq	64_3	3653	72_3	8043	96_3	1558
CAM	65_1	5294	73_1	9934	97_1	3145
CAM	65_2	5544	73_2	10048	97_2	3154
CAM	65_3	5352	73_3	10293	97_3	3158
CAMq	66_1	5291	74_1	9637	98_1	2330
CAMq	66_2	5253	74_2	9636	98_2	2361
CAMq	66_3	4939	74_3	9579	98_3	2362
MMTS	67_1	4653	75_1	NA	99_1	2374
MMTS	67_2	4806	75_2	9879	99_2	2361
MMTS	67_3	4951	75_3	9602	99_3	2355

образец	повтор I	пептиды I	повтор II	пептиды II	повтор III	пептиды III
control	60_1	21	68_1	68	92_1	6
control	60_2	29	68_2	62	92_2	6
control	60_3	22	68_3	67	92_3	4
IAM	61_1	471	69_1	1042	93_1	213
IAM	61_2	502	69_2	1008	93_2	190
IAM	61_3	453	69_3	1028	93_3	188
IAMq	62_1	433	70_1	1488	94_1	175
IAMq	62_2	443	70_2	1501	94_2	165
IAMq	62_3	411	70_3	1399	94_3	144
4-VP	63_1	228	71_1	1392	95_1	154
4-VP	63_2	234	71_2	1275	95_2	159
4-VP	63_3	241	71_3	1272	95_3	152
4-VPq	64_1	365	72_1	1016	96_1	116
4-VPq	64_2	385	72_2	1127	96_2	NA
4-VPq	64_3	380	72_3	1146	96_3	152
CAM	65_1	605	73_1	1571	97_1	253
CAM	65_2	622	73_2	1605	97_2	246
CAM	65_3	611	73_3	1638	97_3	254
CAMq	66_1	530	74_1	1196	98_1	129
CAMq	66_2	516	74_2	1202	98_2	125
CAMq	66_3	494	74_3	1182	98_3	138
MMTS	67_1	379	75_1	NA	99_1	92
MMTS	67_2	381	75_2	1378	99_2	99
MMTS	67_3	405	75_3	1379	99_3	94

3Б. Число идентифицированных пептидов с цистеином

3В. Число идентифицированных пептидов с метионином

образец	повтор I	пептиды I	повтор II	пептиды II	повтор III	пептиды III
control	60_1	1225	68_1	2450	92_1	738
control	60_2	1284	68_2	2539	92_2	762
control	60_3	1242	68_3	2536	92_3	775
IAM	61_1	87	69_1	818	93_1	121
IAM	61_2	88	69_2	784	93_2	103
IAM	61_3	85	69_3	790	93_3	105
IAMq	62_1	1194	70_1	3005	94_1	710
IAMq	62_2	1226	70_2	2996	94_2	681
IAMq	62_3	1101	70_3	2899	94_3	635
4-VP	63_1	1276	71_1	2820	95_1	578
4-VP	63_2	1318	71_2	2618	95_2	533
4-VP	63_3	1313	71_3	2671	95_3	547
4-VPq	64_1	1097	72_1	2429	96_1	344

4-VPq	64_2	1078	72_2	2597	96_2	NA
4-VPq	64_3	1079	72_3	2586	96_3	405
CAM	65_1	1244	73_1	2852	97_1	682
CAM	65_2	1277	73_2	2891	97_2	680
CAM	65_3	1247	73_3	3028	97_3	692
CAMq	66_1	1335	74_1	2948	98_1	552
CAMq	66_2	1318	74_2	2967	98_2	567
CAMq	66_3	1280	74_3	2954	98_3	575
MMTS	67_1	1013	75_1	NA	99_1	533
MMTS	67_2	1052	75_2	2712	99_2	546
MMTS	67_3	1099	75_3	2622	99_3	531

3Г. Отношение числа пептидов с цистемном к общему числу идентифицированных

пептидов

образец	повтор I	пептиды І	повтор II	пептиды II	повтор III	пептиды III
control	60_1	0,0042	68_1	0,0077	92_1	0,0018
control	60_2	0,0055	68_2	0,0068	92_2	0,0017
control	60_3	0,0044	68_3	0,0074	92_3	0,0011
IAM	61_1	0,1252	69_1	0,1658	93_1	0,081
IAM	61_2	0,1268	69_2	0,165	93_2	0,0794
IAM	61_3	0,1211	69_3	0,165	93_3	0,0802
IAMq	62_1	0,0895	70_1	0,1475	94_1	0,0578
IAMq	62_2	0,0895	70_2	0,1507	94_2	0,0563
IAMq	62_3	0,0917	70_3	0,1456	94_3	0,0526
4-VP	63_1	0,0447	71_1	0,1458	95_1	0,0616
4-VP	63_2	0,0447	71_2	0,1428	95_2	0,066
4-VP	63_3	0,0464	71_3	0,1427	95_3	0,0632
4-VPq	64_1	0,098	72_1	0,1362	96_1	0,0898
4-VPq	64_2	0,105	72_2	0,1399	96_2	NA
4-VPq	64_3	0,104	72_3	0,1425	96_3	0,0976
CAM	65_1	0,1143	73_1	0,1581	97_1	0,0804
CAM	65_2	0,1122	73_2	0,1597	97_2	0,078
CAM	65_3	0,1142	73_3	0,1591	97_3	0,0804
CAMq	66_1	0,1002	74_1	0,1241	98_1	0,0554
CAMq	66_2	0,0982	74_2	0,1247	98_2	0,0529
CAMq	66_3	0,1	74_3	0,1234	98_3	0,0584
MMTS	67_1	0,0815	75_1	NA	99_1	0,0388
MMTS	67_2	0,0793	75_2	0,1395	99_2	0,0419
MMTS	67_3	0,0818	75_3	0,1436	99_3	0,0399

3Д. Отношение числа пептидов с метионином к общему числу идентифицированных пептидов

		образец	повтор I	пептиды І	повтор II	пептиды II	повтор III	пептиды III
--	--	---------	----------	-----------	-----------	------------	------------	-------------

control	60_1	0,2446	68_1	0,2765	92_1	0,221
control	60_2	0,2442	68_2	0,2801	92_2	0,222
control	60_3	0,2464	68_3	0,2782	92_3	0,2223
IAM	61_1	0,0231	69_1	0,1302	93_1	0,046
IAM	61_2	0,0222	69_2	0,1283	93_2	0,043
IAM	61_3	0,0227	69_3	0,1268	93_3	0,0448
IAMq	62_1	0,2469	70_1	0,2979	94_1	0,2345
IAMq	62_2	0,2477	70_2	0,3008	94_2	0,2324
IAMq	62_3	0,2455	70_3	0,3018	94_3	0,232
4-VP	63_1	0,2502	71_1	0,2953	95_1	0,2313
4-VP	63_2	0,2515	71_2	0,2933	95_2	0,2212
4-VP	63_3	0,2527	71_3	0,2997	95_3	0,2273
4-VPq	64_1	0,2944	72_1	0,3257	96_1	0,2663
4-VPq	64_2	0,2941	72_2	0,3223	96_2	NA
4-VPq	64_3	0,2954	72_3	0,3215	96_3	0,2599
CAM	65_1	0,235	73_1	0,2871	97_1	0,2169
CAM	65_2	0,2303	73_2	0,2877	97_2	0,2156
CAM	65_3	0,233	73_3	0,2942	97_3	0,2191
CAMq	66_1	0,2523	74_1	0,3059	98_1	0,2369
CAMq	66_2	0,2509	74_2	0,3079	98_2	0,2402
CAMq	66_3	0,2592	74_3	0,3084	98_3	0,2434
MMTS	67_1	0,2177	75_1	NA	99_1	0,2245
MMTS	67_2	0,2189	75_2	0,2745	99_2	0,2313
MMTS	67_3	0,222	75_3	0,2731	99_3	0,2255

3Е. Общее число спектральных идентификаций (PSM)

образец	повтор I	PSM I	повтор II	PSM II	повтор III	PSM III
control	60_1	14058	68_1	14829	92_1	5608
control	60_2	14049	68_2	15018	92_2	6208
control	60_3	13485	68_3	15291	92_3	6073
IAM	61_1	11321	69_1	11095	93_1	4728
IAM	61_2	11566	69_2	10783	93_2	4008
IAM	61_3	11674	69_3	11194	93_3	3919
IAMq	62_1	10859	70_1	18062	94_1	5380
IAMq	62_2	10560	70_2	17718	94_2	5425
IAMq	62_3	9286	70_3	17383	94_3	4836
4-VP	63_1	11866	71_1	17242	95_1	4219
4-VP	63_2	11823	71_2	17038	95_2	4098
4-VP	63_3	11820	71_3	17060	95_3	4074
4-VPq	64_1	8562	72_1	14853	96_1	2492
4-VPq	64_2	8579	72_2	15294	96_2	NA
4-VPq	64_3	8410	72_3	15324	96_3	3072
CAM	65_1	12343	73_1	17879	97_1	5620
------	------	-------	------	-------	------	------
CAM	65_2	12697	73_2	18286	97_2	5465
CAM	65_3	11319	73_3	18719	97_3	5493
CAMq	66_1	12091	74_1	17749	98_1	4208
CAMq	66_2	10689	74_2	17518	98_2	4155
CAMq	66_3	10355	74_3	17701	98_3	4083
MMTS	67_1	9774	75_1	NA	99_1	3902
MMTS	67_2	9755	75_2	16801	99_2	3858
MMTS	67_3	7746	75_3	15820	99_3	3925

3Ж. Число спектральных идентификаций (PSM) с цистеином

образец	повтор I	PSM I	повтор II	PSM II	повтор III	PSM III
control	60_1	37	68_1	85	92_1	6
control	60_2	42	68_2	83	92_2	10
control	60_3	32	68_3	92	92_3	5
IAM	61_1	1217	69_1	1864	93_1	345
IAM	61_2	1240	69_2	1798	93_2	284
IAM	61_3	1243	69_3	1839	93_3	278
IAMq	62_1	827	70_1	2662	94_1	278
IAMq	62_2	801	70_2	2630	94_2	274
IAMq	62_3	711	70_3	2539	94_3	227
4-VP	63_1	408	71_1	2721	95_1	248
4-VP	63_2	385	71_2	2655	95_2	258
4-VP	63_3	442	71_3	2683	95_3	254
4-VPq	64_1	583	72_1	2086	96_1	188
4-VPq	64_2	631	72_2	2240	96_2	NA
4-VPq	64_3	590	72_3	2241	96_3	248
CAM	65_1	1163	73_1	2784	97_1	423
CAM	65_2	1223	73_2	2929	97_2	403
CAM	65_3	1113	73_3	3034	97_3	399
CAMq	66_1	1004	74_1	2153	98_1	207
CAMq	66_2	880	74_2	2151	98_2	199
CAMq	66_3	883	74_3	2167	98_3	204
MMTS	67_1	758	75_1	NA	99_1	153
MMTS	67_2	749	75_2	2260	99_2	159
MMTS	67_3	600	75_3	2207	99_3	159

33. Число спектральных идентификаций (PSM) с метионином

образец	повтор I	PSM I	повтор II	PSM II	повтор III	PSM III
control	60_1	3634	68_1	4259	92_1	1224
control	60_2	3635	68_2	4404	92_2	1349
control	60_3	3552	68_3	4391	92_3	1344
IAM	61_1	217	69_1	1566	93_1	197

IAM	61_2	205	69_2	1478	93_2	146
IAM	61_3	230	69_3	1488	93_3	142
IAMq	62_1	2696	70_1	5675	94_1	1192
IAMq	62_2	2646	70_2	5633	94_2	1207
IAMq	62_3	2296	70_3	5598	94_3	1085
4-VP	63_1	2921	71_1	5370	95_1	917
4-VP	63_2	2964	71_2	5242	95_2	884
4-VP	63_3	2919	71_3	5368	95_3	891
4-VPq	64_1	2364	72_1	5057	96_1	620
4-VPq	64_2	2344	72_2	5129	96_2	NA
4-VPq	64_3	2347	72_3	5183	96_3	782
CAM	65_1	2917	73_1	5462	97_1	1212
CAM	65_2	3002	73_2	5571	97_2	1146
CAM	65_3	2633	73_3	5725	97_3	1158
CAMq	66_1	3039	74_1	5788	98_1	959
CAMq	66_2	2681	74_2	5640	98_2	950
CAMq	66_3	2590	74_3	5724	98_3	955
MMTS	67_1	2183	75_1	NA	99_1	854
MMTS	67_2	2132	75_2	4775	99_2	866
MMTS	67_3	1823	75_3	4501	99_3	867

ЗИ. Отношение числа PSM с цистеином к общему числу PSM

образец	повтор I	PSM I	повтор II	PSM II	повтор III	PSM III
control	60_1	0,0026	68_1	0,0057	92_1	0,0011
control	60_2	0,003	68_2	0,0055	92_2	0,0016
control	60_3	0,0024	68_3	0,006	92_3	0,0008
IAM	61_1	0,1075	69_1	0,168	93_1	0,073
IAM	61_2	0,1072	69_2	0,1667	93_2	0,0709
IAM	61_3	0,1065	69_3	0,1643	93_3	0,0709
IAMq	62_1	0,0762	70_1	0,1474	94_1	0,0517
IAMq	62_2	0,0759	70_2	0,1484	94_2	0,0505
IAMq	62_3	0,0766	70_3	0,1461	94_3	0,0469
4-VP	63_1	0,0344	71_1	0,1578	95_1	0,0588
4-VP	63_2	0,0326	71_2	0,1558	95_2	0,063
4-VP	63_3	0,0374	71_3	0,1573	95_3	0,0623
4-VPq	64_1	0,0681	72_1	0,1404	96_1	0,0754
4-VPq	64_2	0,0736	72_2	0,1465	96_2	NA
4-VPq	64_3	0,0702	72_3	0,1462	96_3	0,0807
CAM	65_1	0,0942	73_1	0,1557	97_1	0,0753
CAM	65_2	0,0963	73_2	0,1602	97_2	0,0737
CAM	65_3	0,0983	73_3	0,1621	97_3	0,0726
CAMq	66_1	0,083	74_1	0,1213	98_1	0,0492

CAMq	66_2	0,0823	74_2	0,1228	98_2	0,0479
CAMq	66_3	0,0853	74_3	0,1224	98_3	0,05
MMTS	67_1	0,0776	75_1	NA	99_1	0,0392
MMTS	67_2	0,0768	75_2	0,1345	99_2	0,0412
MMTS	67_3	0,0775	75_3	0,1395	99_3	0,0405

3К. Отношение числа PSM с метионином к общему числу PSM

образец	повтор I	PSM I	повтор II	PSM II	повтор III	PSM III
control	60_1	0,2585	68_1	0,2872	92_1	0,2183
control	60_2	0,2587	68_2	0,2932	92_2	0,2173
control	60_3	0,2634	68_3	0,2872	92_3	0,2213
IAM	61_1	0,0192	69_1	0,1411	93_1	0,0417
IAM	61_2	0,0177	69_2	0,1371	93_2	0,0364
IAM	61_3	0,0197	69_3	0,1329	93_3	0,0362
IAMq	62_1	0,2483	70_1	0,3142	94_1	0,2216
IAMq	62_2	0,2506	70_2	0,3179	94_2	0,2225
IAMq	62_3	0,2473	70_3	0,322	94_3	0,2244
4-VP	63_1	0,2462	71_1	0,3114	95_1	0,2174
4-VP	63_2	0,2507	71_2	0,3077	95_2	0,2157
4-VP	63_3	0,247	71_3	0,3147	95_3	0,2187
4-VPq	64_1	0,2761	72_1	0,3405	96_1	0,2488
4-VPq	64_2	0,2732	72_2	0,3354	96_2	NA
4-VPq	64_3	0,2791	72_3	0,3382	96_3	0,2546
CAM	65_1	0,2363	73_1	0,3055	97_1	0,2157
CAM	65_2	0,2364	73_2	0,3047	97_2	0,2097
CAM	65_3	0,2326	73_3	0,3058	97_3	0,2108
CAMq	66_1	0,2513	74_1	0,3261	98_1	0,2279
CAMq	66_2	0,2508	74_2	0,322	98_2	0,2286
CAMq	66_3	0,2501	74_3	0,3234	98_3	0,2339
MMTS	67_1	0,2233	75_1	NA	99_1	0,2189
MMTS	67_2	0,2186	75_2	0,2842	99_2	0,2245
MMTS	67_3	0,2353	75_3	0,2845	99_3	0,2209

Таблица 4. Число пептидов с метионином, преобразованным в изотреонин и число общее число пептидов, содержащих остатки метионина. Данные группо-специфичную прошли фильтровку 4А. Данные нашей работы

образец	пептиды с Мет	пептиды с Мет- изоТре
cont	1465	8
IAM	591	19
IAMq	2244	10
4-VP	1836	10
4-VPq	1921	16
CAM	2190	10

CAMq	2126	8
MMTS	2148	11

образец	пептиды с Мет	пептиды с Мет- изоТре
InGel_CON	1085	6
InSolution_CON	2574	5
InGel_DA	1055	3
InGel_MA	11164	21
InGel_TA	1339	4
InSolution_DA	2237	7
InSolution_MA	2699	6
InSolution_TA	2423	5
InGel_DC	1182	3
InGel_MC	1260	8
InGel_TC	1121	4
InSolution_DC	2570	8
InSolution_MC	2508	4
InSolution_TC	2734	4
InGel_DI	239	7
InGel_MI	533	7
InGel_TI	537	5
InSolution_DI	1665	15
InSolution_MI	1747	17
InSolution_TI	1644	16
InGel_DIA	109	3
InGel_MIA	187	5
InGel_TIA	253	4
InSolution_DIA	109	3
InSolution_MIA	187	5
InSolution_TIA	253	4

4Б. Данные из работы [7]

4В. Данные по клеточным линиям панели NCI60 из [131]

клеточная линия	пептиды с Мет	пептиды с Мет- изоТре
COLO-205	12369	305
PXF-393	18426	69
PC-3	22023	369
NCI-H460	8034	243
U-251	12829	189
SKOV-3	11786	219
M14	12646	339
MCF-7	11560	335
CCRF-CEM	9582	225

отдел	пептиды с Мет	пептиды с Мет- изоТре
brain	15100	26
CG_neurons	13270	21
cort_neurons	13941	21
oligodendro	21222	27
adultmicroglia	18049	19
brain_fract	10501	24
prefrnt_cortex	19695	42
collab_neur	14741	11
thalamus	19971	25

4Γ.	Данные по	отделам	мозга	мыши	ИЗ	[137]
-----	-----------	---------	-------	------	----	-------

Таблица 5. Число метионинов, преобразованных в изотреонин предшествующих каждому аминокислотному остатку и следующему за ним, а также общее число пептидов, сожержащих комбинацию каждого остатка с остатком метионина

а.о. предшествующий метионину	число Мет-изоТре следующее за a.o.	комбинация	общее число пептидов с комбинацией
A	118	AM	7053
С	16	СМ	1082
D	107	DM	5043
Е	127	EM	8145
F	58	FM	2696
G	166	GM	5179
Н	23	HM	1597
Ι	57	IM	4375
K	27	KM	936
L	117	LM	8044
Ν	147	NM	3923
Р	83	PM	3198
Q	87	QM	4962
R	18	RM	301
S	154	SM	5015
Т	91	ТМ	4024
V	144	VM	5718
W	6	WM	518
Y	64	YM	2341
а.о. следующий	число Мет-изоТре	комбинация	общее число пептидов с
А	предшествующее а.о. 173	МА	7449
C C	22	MC	928
D	125	MD	5349
E	125	ME	8092
F	32	MF	2787
G	93	MG	5106
H	7	MH	1390

Ι	118	MI	4121
К	134	МК	6796
L	123	ML	8140
Ν	101	MN	4500
Р	218	MP	3879
Q	84	MQ	4527
R	119	MR	4089
S	61	MS	4703
Т	77	MT	4401
V	202	MV	5952
W	10	MW	510
Y	19	MY	1823