

ОБЗОРЫ

©Коллектив авторов

МОДИФИКАЦИЯ ОСТАТКОВ ЦИСТЕИНА ДЛЯ МАСС-СПЕКТРОМЕТРИЧЕСКОГО ПРОТЕОМНОГО АНАЛИЗА: ФАКТЫ И АРТЕФАКТЫ

К.Г. Кузнецова^{1}, Е.М. Соловьева^{2,3}, А.В. Кузиков^{1,4}, М.В. Горшков², С.А. Мошковский^{1,4}*

¹Научно-исследовательский институт биомедицинской химии им. В.Н. Ореховича, 119121, Москва, ул. Погодинская, 10; *эл. почта: kuznetsova.ks@gmail.com

²Институт энергетических проблем химической физики им. В.Л. Тальрозе, Федеральный исследовательский центр химической физики РАН им. Н.Н. Семенова, 119334, Москва, Ленинский пр., 38, корп. 2

³Московский физико-технический институт (государственный университет), 141700, Долгопрудный, Институтский пер., 9

⁴Российский национальный исследовательский медицинский университет им. Н.И. Пирогова, 117997, Москва, ул. Островитянова, 1

Масс-спектрометрический протеомный анализ на стадии пробоподготовки включает в себя искусственное восстановление дисульфидных связей, образованных остатками цистеина. Такие связи при их сохранении в нативном состоянии затрудняют последующий ферментативный гидролиз и интерпретацию результатов исследования. Для предотвращения повторного образования дисульфидных связей остатки цистеина защищают специальными группами, чаще всего, путём алкилирования. В данном обзоре рассмотрены используемые способы модификации остатков цистеина во время пробоподготовки, а также возможные артефакты этого этапа, в частности, побочные реакции используемых для этого алкилирующих агентов с другими аминокислотными остатками. Наиболее распространённым алкилирующим соединением для защиты остатков цистеина в масс-спектрометрическом протеомном анализе является йодацетамид. Однако данные литературы в этой области указывают на то, что данный реагент вызывает больше побочных реакций, чем другие используемые агенты, такие как хлорацетамид и акриламид. Последние могут быть рекомендованы для более широкого использования. В обзоре также обсуждаются особенности модификации остатков цистеина и их влияние на эффективность поиска посттрансляционных модификаций и белковых продуктов генов, содержащих одонуклеотидные замены.

Ключевые слова: протеомика; пробоподготовка; алкилирование; цистеин; масс-спектрометрия

DOI: 10.18097/PBMC20206601018

ВВЕДЕНИЕ

Понятие протеома как белкового эквивалента генома появилось не так давно, в 1994 году [1], в то время как работа с протеомами, а тем более с отдельными белками, была начата намного раньше. Ещё в пятидесятые годы прошлого века появился метод определения аминокислотной последовательности пептидов, который в настоящее время принято называть методом деградации Эдмана [2]. Изучить и, в буквальном смысле, увидеть сам протеом исследователям удалось в начале семидесятых годов двадцатого века с появлением метода разделения белков при помощи электрофореза в полиакриламидном геле. Сначала белки разделяли по молекулярной массе в денатурирующих условиях в присутствии додецилсульфата натрия [3]. Практически сразу после этого появился метод так называемого двумерного электрофореза – разделения белков на геле в двух направлениях: по изоэлектрической точке и по молекулярной массе [4]. 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 [3]. 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 биомедицинских 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 [5], в сочетании с использованием для идентификации белков поиска по геномным последовательностям, дало толчок к развитию так называемой панорамной (shotgun) протеомики [6]. Смесь белков для этого анализа расщепляют на пептиды специфичной протеазой, например, трипсином, а потом получившиеся пептиды анализируют с использованием высокоэффективной жидкостной хроматографии, совмещённой с тандемной масс-спектрометрией (ВЭЖХ-МС/МС, или LC-MS/MS) [7]. В связи с этим, необходимость разделения белков 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 хромато-масс-спектрометрический анализ. Полученные масс-спектры анализируют с использованием поисковых протеомных машин [8]. 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 [9, 10]. 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 надёжнее идентифицирован пептид. Далее полученные пептиды “собирают” в белки [11]. Такой подход называется “снизу вверх” (top-down). 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 [12, 13], однако панорамный анализ “снизу вверх” наиболее широко используется в практике биологических исследований.

Метод панорамного протеомного анализа состоит из большого числа этапов (рис. 1), включающих гомогенизацию и лизис образца [14], разделение белков электрофорезом [3], восстановление и алкилирование остатков цистеина [15], ферментативный гидролиз с образованием пептидов [2], а также хромото-масс-спектрометрический анализ этих пептидов [7]. В ходе пробоподготовки могут протекать побочные химические реакции, влияющие на результат протеомного анализа, такие как, например, спонтанное окисление метионина [16] или неферментативное дезамидирование глутамина и аспарагина до, соответственно, глутаминовой и аспарагиновой кислот [17]. 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 [18, 19].

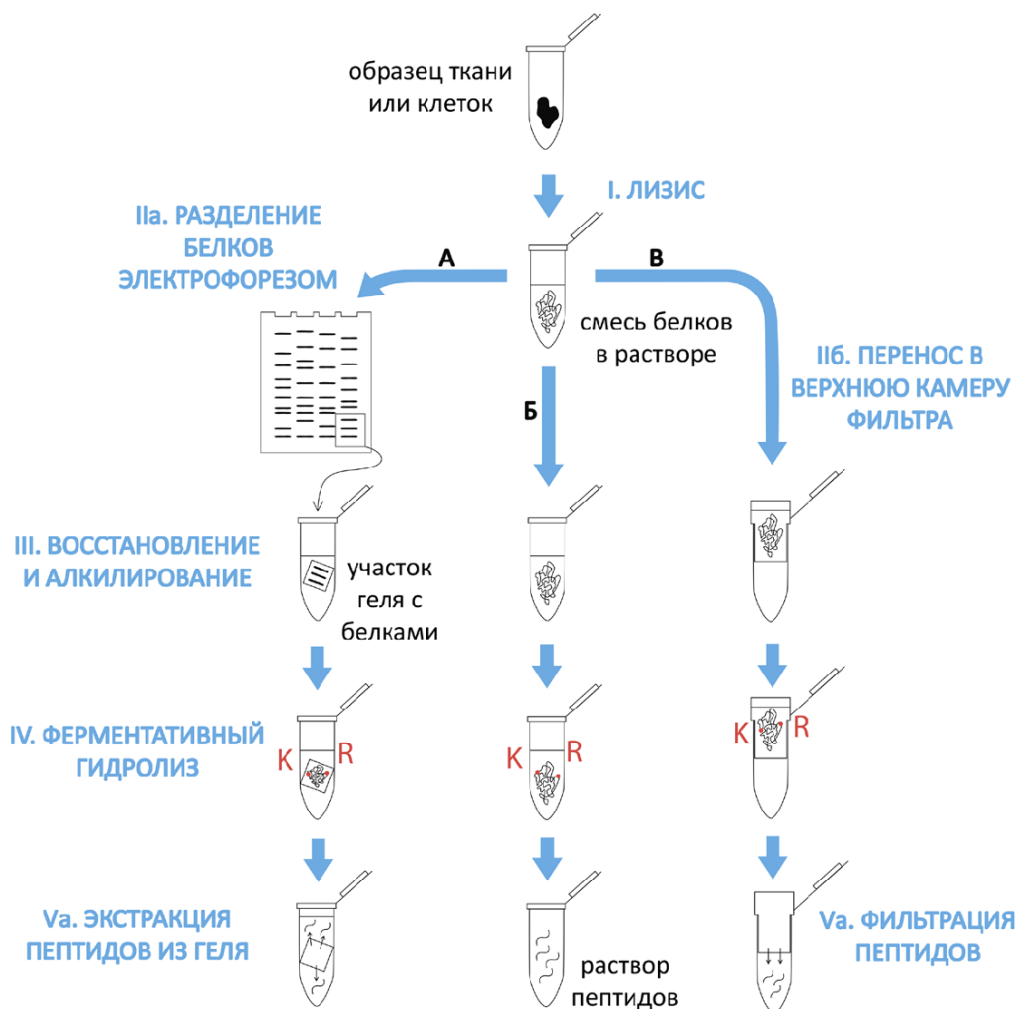


Рисунок 1. Основные пути пробоподготовки к масс-спектрометрическому протеомному анализу: А – в геле; Б – в растворе; В – с использованием центрифужных фильтров. Основные этапы в разных вариантах: I – лизис образца; IIa – разделение белков методом электрофореза; IIб – перенос раствора белков в верхнюю 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.

1. ОБРАБОТКА ДИСУЛЬФИДНЫХ СВЯЗЕЙ ОСТАТКОВ ЦИСТЕИНА ДЛЯ ПРОТЕОМНОГО АНАЛИЗА

Как известно, в формировании третичной структуры белков участвуют дисульфидные (S-S) связи, соединяющие остатки цистеина между собой. Это затрудняет масс-спектрометрический анализ пептидов, содержащих этот аминокислотный остаток. Полученные в результате расщепления трипсином или другими протеазами пептиды с сохранёнными S-S-связями имеют массу, которую сложно предсказать на основе геномных данных, поэтому при протеомном поиске такие пептиды остаются неидентифицированными [15]. 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 путём присоединения химически стабильных радикалов, чаще всего от алкилирующих агентов [15, 20]. 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 [21].

2. ВОССТАНОВЛЕНИЕ ДИСУЛЬФИДНЫХ СВЯЗЕЙ МЕЖДУ ОСТАТКАМИ ЦИСТЕИНА

Обычно восстановление занимает около 40-60 мин при температуре около 57-60°C [21]. Наиболее распространёнными восстанавливающими агентами для разрыва S-S-связей между остатками цистеина являются β-меркаптоэтанол (BME) [3], дитиотреитол (DTT) [22] и трис(2-карбокsetил)фосфин (TCPEP) [23]. В литературе есть данные об использовании и других фосфинов [24, 25], но в протеомике их применяют редко. Известно несколько работ, в которых сопоставлены результаты протеомного анализа, проведённого с использованием разных восстанавливающих агентов [25, 26]. В одной из них [25] авторы не нашли существенных отличий в эффективности восстановления при помощи DTT, BME, TCPEP и редко применяемого трис(2-гидроксипропил)фосфина (THPP). В другой работе [26] авторы проводили более сложные эксперименты, сравнивая сочетания восстанавливающих и алкилирующих агентов, а также методов пробоподготовки в геле и в растворе. В последнем случае был использован так называемый метод центрифужных фильтров [27, 28] (рис. 1). Все реакции, в том числе восстановление, алкилирование и трипсинолиз, проводили в верхней камере фильтров, а полученные в итоге пептиды смывали сквозь фильтры в нижнюю камеру, обеспечивая, таким образом, очистку получившихся образцов. В результате, в случае пробоподготовки в геле, наибольшее количество спектральных идентификаций получено при восстановлении дисульфидных связей при помощи DTT, в то время как в растворе его превзошел метод с использованием BME.

Оба реагента показали наилучший результат в сочетании с акриламидом в качестве алкилирующего соединения. Стоит отметить, что ни одна из перечисленных работ не указывает на появление артефактов в результатах анализа под воздействием тех или иных реагентов на этапе восстановления.

3. ЗАЩИТА ТИОЛЬНЫХ ГРУПП ЦИСТЕИНА: АЛКИЛИРУЮЩИЕ АГЕНТЫ

В отличие от восстановления, этап алкилирования вызывает наибольший интерес среди исследователей, изучающих влияние пробоподготовки на результаты протеомного анализа. 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 [25, 26, 29]. Сделанные в этих работах выводы частично согласуются, но есть и противоречивые результаты. Так, например, в работе [26] не рекомендуется использовать йодацетамид, а в работе [25], 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 [26]. 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 [30]. Таким образом, неучтённые модификации аминокислот сильно снижают качество результатов анализа.

В качестве алкилирующих агентов в протеомном анализе применяют: (i) галогензамещённые амиды и кислоты, например йодацетамид (IAM) [28], хлорацетамид (CAM) [31] и йодуксусную кислоту (IAA) [29], (ii) пиридины, в том числе 4-винилпиридин (4-VP) и 2-винилпиридин (2-VP) [32], а также (iii) акриламид (AA) [33], (iv) метилметантисульфат (MMTS) [34] и (v) N-этилмалеимид (NEM) [35] (рис. 2). Алкилирование обычно проводят в течение 10-30 мин при комнатной температуре в темноте [21]. Повышение температуры реакции алкилирования может привести к возрастанию количества побочных продуктов [36]; кроме того, алкилирующие агенты неустойчивы на свету [37]. Один из вариантов метода

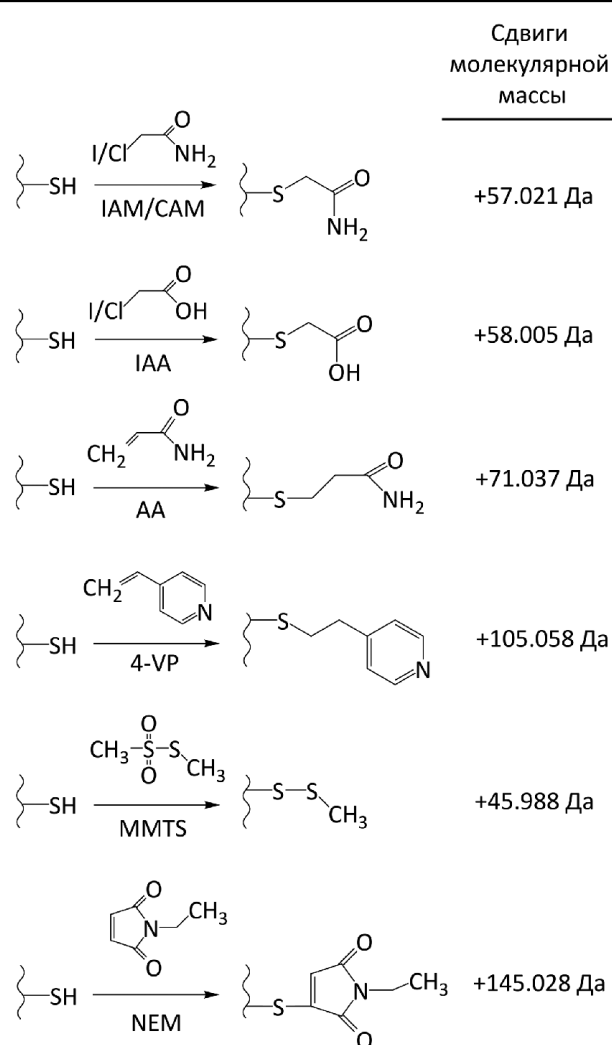


Рисунок 2. Модификации остатков цистеина используемыми в протеомике алкилирующими агентами.

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 [38]. Это должно препятствовать алкилированию других аминокислотных остатков, которое может возникнуть при использовании некоторых соединений [18]. 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 [38].

Анализ литературы показывает, что чаще всего для алкилирования остатков цистеина в протеомике используют IAM [26]. IAM был использован в более чем 80% работ, опубликованных в двух основных протеомных журналах, а именно в журналах *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

[18, 19, 25, 26, 31, 39, 40]. Наиболее распространённой побочной реакцией является избыточное алкилирование (overalkylation) IAM и IAA. 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.

4. АЛКИЛИРОВАНИЕ ОСТАТКОВ ЦИСТЕИНА: ПОБОЧНЫЕ РЕАКЦИИ

В экспериментах с использованием синтетических пептидов было показано, что кроме цистеинов алкилированию подвергаются, в первую очередь, N-концевые аминокислоты [18]. Во вторую очередь модификации подвергаются другие аминокислотные остатки, содержащие метилтио- или аминокислотную группу (то есть метионин и лизин) а также гистидин, содержащий имидазольное кольцо [18, 41]. Кроме них алкилированию могут подвергаться тирозин и C-концевая карбоксильная группа пептида [41]. В одной из работ [18] авторы пробовали избежать эффекта алкилирования IAM аминокислот, проводя реакцию в бикарбонате аммония, но этот подход не дал желаемого результата. В итоге они предложили использовать для удаления избытка IAM тиоэфир. Эти и другие аналогичные исследования [18, 41] были выполнены на синтетических пептидах, и в соответствующих публикациях подробно описаны химические процессы, лежащие в основе эффекта избыточного алкилирования белков IAM 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 ещё не расщеплён 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 всё же алкилируют N-концевые аминокислоты многих протеолитических пептидов [25, 26, 41, 42].

Помимо работ с синтетическими пептидами, процесс алкилирования аминокислотных остатков изучали и на выделенных расщеплённых белках [40]. Было показано, что IAA и IAM, помимо цистеина, алкилируют метионин, гистидин и лизин. В этой работе авторы рекомендуют подбирать эквимольное 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 [40]. В ряде других исследований утверждается, что и САМ способен вызывать такой артефакт, но происходит это также при повышенных температурах [36].

Самый подробный анализ избыточного алкилирования различными агентами был проведён в работе [26], авторы которой показали, что в результатах проведенного ими анализа протеомов присутствуют пептиды, алкилированные IAM, IAA, AA и САМ 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 [43, 44]. Более детальный анализ с использованием особого масс-спектрометрического метода “многоэтапной активации” для анализа нейтральных потерь [45] 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 [26] использовали метод поиска модификаций пептидов в широком диапазоне масс [46]. 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” [47], и о нём будет рассказано далее.

5. 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 [16] (рис. 3А). Именно поэтому окисление остатков метионина в качестве “вариабельной” модификации является часто используемым параметром в протеомных поисках [48]. Окисление метионина может происходить как в клетке естественным путём [49-51], так и в процессе пробоподготовки (наряду с триптофаном) [16]. 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 [52] (рис. 4). Известно, что окисленный метионин может претерпевать отщепление боковой цепи (“нейтральную потерю”) во время так называемой диссоциации, индуцированной столкновениями (*англ.* collision-induced dissociation – CID), в результате которой образуются пептидные ионы с молекулярной массой, уменьшенной на 64 Да [53] (рис. 3А). Показано, что такие пики чаще появляются в масс-спектрах после обработки белков йодсодержащими агентами, а именно IAM и IAA [26]. В то же время, нейтральная потеря боковой цепи метионина, вызывающая сдвиг молекулярной массы на -48 Да, также характерна для алкилирования IAM и IAA [26]. В экспериментах с синтетическими пептидами показано [26], что такие потери боковой цепи метионина вызывает не окисление, а именно обработка йодсодержащими агентами. Это противоречит некоторым результатам, полученным ранее [53], но соответствует другим данным [44]. В действительности, метионин теряет

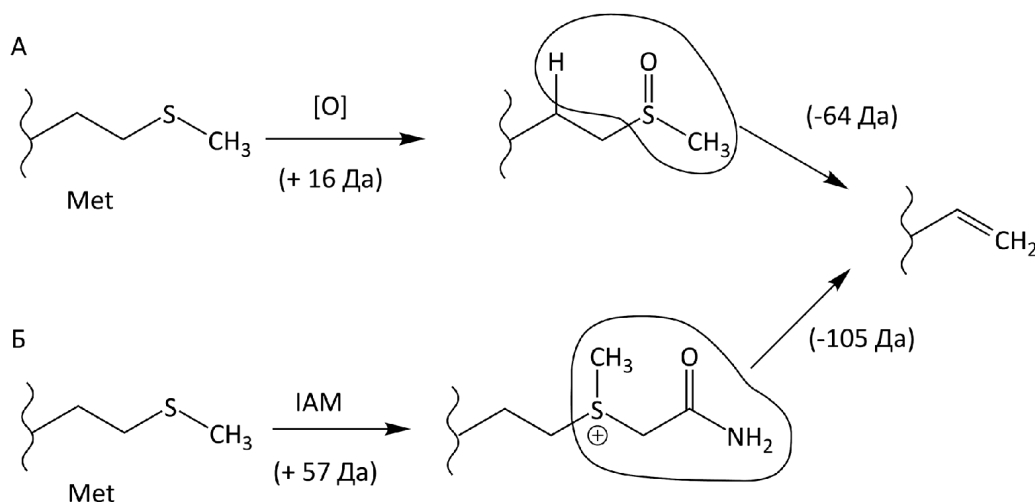


Рисунок 3. Модификации остатков метионина и последующие возможные нейтральные потери: А – окисление, Б – алкилирование йодацетамидом. Адаптировано из [70].

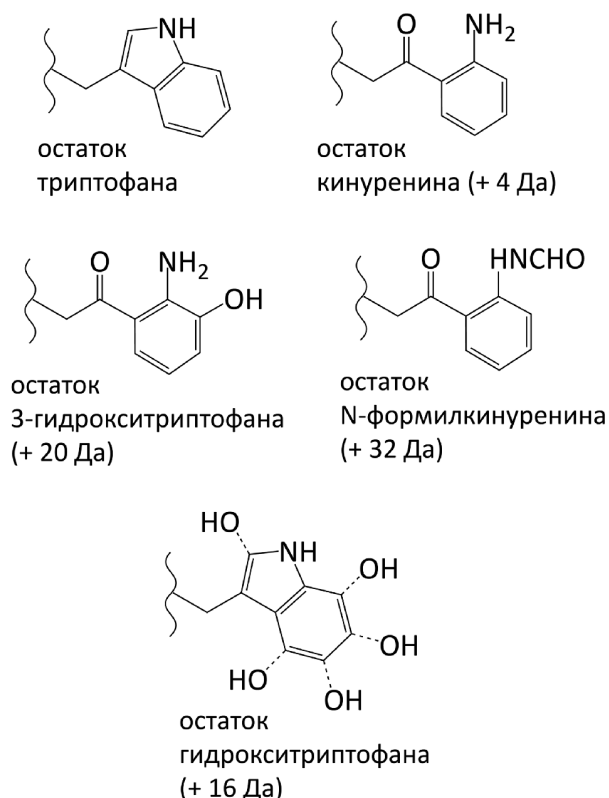


Рисунок 4. Продукты окисления остатков триптофана. Адаптировано из [52].

не 48 Да, а 105 Да (в случае обработки IAM), поскольку от него отщепляется не интактная, а модифицированная алкилирующим агентом боковая цепь [43, 44]. Указанный сдвиг молекулярной массы на 105 Да включает в себя массу карбамидометила, равную 57 Да, и массу боковой цепи метионина, равную 48 Да (рис. 3Б).

В своей работе Müller и Winter [26] после детального анализа рекомендуют вместо йодсодержащих соединений использовать для алкилирования остатков цистеина AA. Также они получили удовлетворительное количество идентифицированных пептидов при минимуме побочных реакций при использовании хлорацетамида (SAM). Однако в ряде работ получены 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 идёт AA, только потом IAM и наименьшую вызывает 4-VP [25]. Данная работа также была выполнена на пептидах, полученных из белков после расщепления трипсином. При этом, авторы все равно рекомендуют использовать IAM, 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 [26]. В ней авторы не сравнивали комбинации алкилирующих и восстанавливающих агентов, а исследовали их эффекты отдельно. Также в ней: отсутствует “открытый” поиск [46], не проведены контрольные

эксперименты на синтетических пептидах, а также не была выполнена “многоэтапная активация” [45], то есть не были проанализированы нейтральные потери так, как это делали Müller и Winter [26]. Масс-спектрометрические данные их работы, вопреки сложившейся практике, не опубликованы в открытом доступе и не могут быть проверены и использованы для дальнейшего анализа. По SAM имеются единичные указания на то, что этот агент способствует окислению остатков метионина и триптофана [54].

6. 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 соавт. [31] показали, что IAM алкилирует остаток лизина с образованием двойных аддуктов, молекулярная масса которых равна диглициновым меткам сайтов убиквитинирования (рис. 5). Особенно выражено этот эффект проявлялся при нагревании. Таким образом, при изучении убиквитинирования не рекомендуется использовать IAM. Его предлагают заменить на SAM, 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°C такой эффект может вызвать и SAM [36]. 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 SAM [55]. Несмотря на это, во многих работах по убиквитинированию всё же используют IAM [56-58].

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 [43]. Таким образом, пептид со сдвигом массы на -105 Да, появляющийся в результатах при учёте 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 [43]. 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) даёт сдвиг массы на 79 Да, близкий по массе

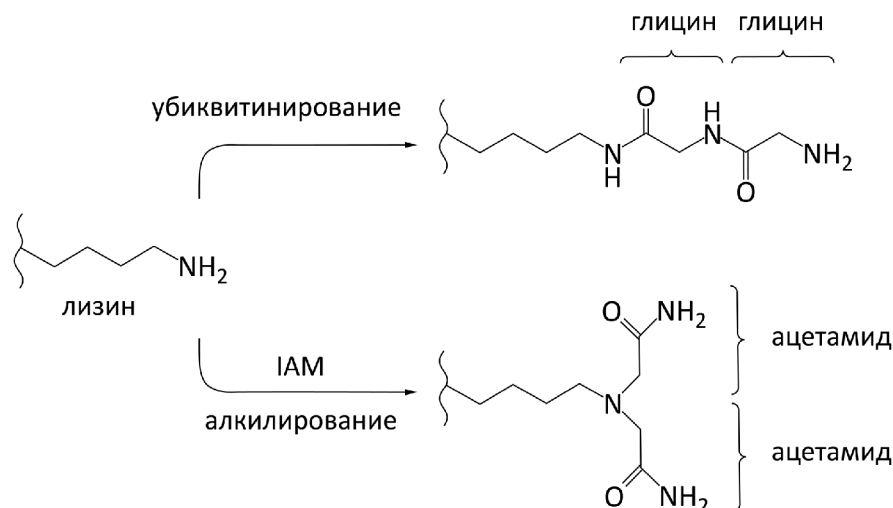


Рисунок 5. Диглициновые метки, образующиеся на месте сайтов убиквитинирования по массе неотличимы от двойных аддуктов IAM. Адаптировано из [57].

к самому фосфорилированию остатков треонина и серина (80 Да) [43]. Такие пептиды также будут иметь общий спектр фрагментации, и без дополнительного разделения химерных спектров [59] 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.

7. 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 транскриптом полиморфизмам [60]. 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. Подобное совпадение было обнаружено нашей группой при протеогеномном исследовании раковых мутаций [19]. Тогда в консенсусную базу данных человеческих белков были добавлены варианты пептиды с учётом этих мутаций. Экспериментальные масс-спектры, полученные из образцов нескольких раковых клеточных линий панели NCI60 [61], были проанализированы поисковой программой с использованием этой базы данных. Оказалось, что в новой дополненной базе данных в определённом месте присутствовала замена метионина на треонин,

она же нашлась в протеомах 8 из 9 клеточных линий. Однако при проверке выяснилось, что в геноме конкретных клеточных линий такой замены нет. Дело в том, что при алкилировании белков IAM часть остатков метионина была преобразована в остатки гомосерина (изотреонина) с молекулярной массой, равной массе остатка треонина. Один из таких метионинов совпал с заменой в геноме триплета метионина на триплет треонина в результате раковой мутации, и был обнаружен нами в протеоме. Преобразование метионина в гомосерин было описано давно [62], но в составе белков эта реакция до нас обнаружена не была. Действительно, доля остатков метионина, подверженных такой модификации, невелика и в среднем не превышает 1% в случае стандартной пробоподготовки и панорамного протеомного анализа [Кузнецова и соавт., статья готовится к печати]. Реагируя с IAA, метионин образует гомосеринлактон [62]. Описанный механизм этой реакции позволяет предположить, что IAM также способен вызывать это преобразование.

8. ОБРАБОТКА ОСТАТКОВ ЦИСТЕИНА В РЕДОКС-ПРОТЕОМИКЕ

Кроме подробно описанных выше алкилирующих агентов, для той же цели используются и такие соединения как MMTS [32] и NEM [35]. Эти реагенты пришли в панорамный протеомный анализ из смежной области, а именно из редокс-протеомики [63]. Как уже упоминалось выше, NEM иногда использовался и в исследованиях сайтов убиквитинирования белков [36], а также в панорамной протеомике [35]. Редокс-протеомика – это достаточно большая область протеомики, которая изучает окислительно-восстановительные процессы в белках [63]. Одна из задач редокс-протеомики – изучение естественного состояния тиольных групп цистеинов. Таким образом, реакция алкилирования находится в центре внимания методов, применяемых в этой области [64, 65]. 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. 2). Это его свойство широко применяется для определения того, какие конкретно цистеины в белке были свободными от S-S-связей, а какие – связанными [64]. Несмотря на это, MMTS используется и в панорамной протеомике, причём в ряде публикаций он признаётся 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 [35, 36]. При этом показано, что NEM способен алкилировать остатки лизина и гистидина [67], а также N-конец пептидов [25].

9. БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ МОДИФИКАЦИЙ АМИНОКИСЛОТНЫХ ОСТАТКОВ МЕТОДОМ “ОТКРЫТОГО” ПОИСКА

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 (таблица). 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. 6) [46].

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 identifications (обычно методами “декойных” последовательностей). В результате можно оценить, какие сдвиги молекулярной массы, по сравнению с теоретической, чаще всего встречались.

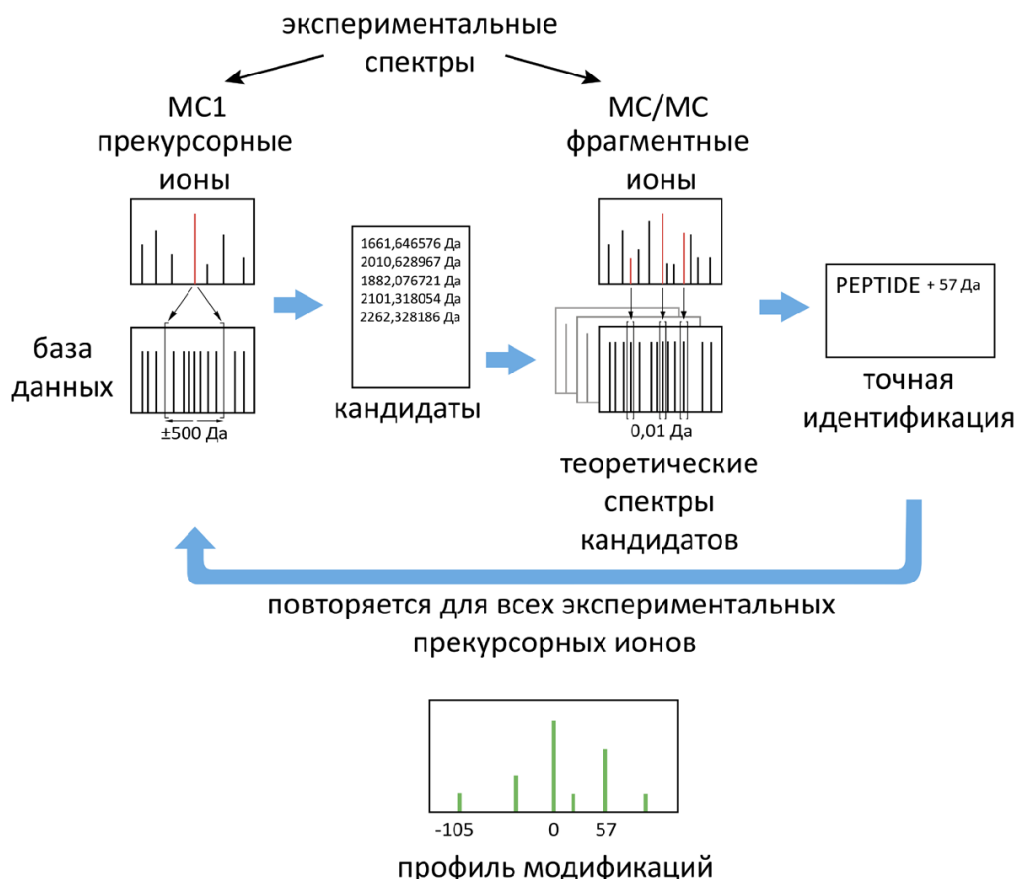


Рисунок 6. Общий принцип метода “открытого” поиска [43]. Сначала поисковая машина сравнивает экспериментальную массу прекурсорного иона с теоретическими массами пептидов из базы данных с широким окном допустимой ошибки определения массы ± 500 Да, получая таким образом список “кандидатов”. Затем спектр 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 определёнными сдвигами масс.

МОДИФИКАЦИЯ ОСТАТКОВ ЦИСТЕИНА В ПРОТЕОМИКЕ

Таблица. Использование алкилирующих агентов при решении различных задач протеомного анализа

	IAA	IAM	AA	SAM	4-VP	MMTS	NEM
панорамная протеомика	не рекомендован избыточное алкилирование по N-концевой аминогруппе, Met, Hys, Lys, Thr, C-term [16,23,24,33,38,39], Ser, Asp, Glu [24]	не рекомендован избыточное алкилирование по N-концевой аминогруппе, Met, Hys, Lys, Thr, C-term возможно использование с гашением избытка после алкилирования	рекомендован [24]	рекомендован [24] возможно избыточное алкилирование при повышенной температуре [33], но с меньшей интенсивностью, чем IAA, IAM и AA [24]. может вызывать усиленное окисление Met и Thr [51]	возможно избыточное алкилирование при повышенной температуре [37], недостаточная эффективность алкилирования Cys [23]	рекомендован [62] обратимое связывание, исключено гашение избытка после алкилирования	рекомендован возможно избыточное алкилирование по N-концу Lys и Hys [63]
протеогеномика – поиск одноаминокислотных вариантов	не рекомендован способствует преобразованию Met в isoThr [58]	не рекомендован способствует преобразованию Met в isoThr [17]					
исследования ПТМ:							
фосфорилирование		применение исключено алкилирует Ser и Thr [40]					
убиквитинирование		применение исключено алкилирует Lys с образованием двойных аддуктов, не отличимых от сайтов убиквитинирования [29]		рекомендован при проведении экспериментов при 21°C [33]			рекомендован [33]
редокс-протеомика		используется в качестве необратимо связывающегося агента в комбинации с MMTS [60]				используется в качестве обратимо связывающегося агента в комбинации с NEM [60]	используется в качестве необратимо связывающегося агента в комбинации с MMTS [60]

Внося в качестве постоянных модификаций те, в которых исследователь уверен (например, алкилирование цистеина, если оно проводилось специально), можно увидеть, какие ещё 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 переменного параметра при основном протеомном поиске и получить большее количество идентифицированных пептидов, то есть более качественно интерпретированные результаты. Применение открытого поиска было продемонстрировано на протеомных данных [68, 69]. Было показано, что учёт модификаций, обнаруженных при открытом поиске, позволяет увеличить количество идентифицированных пептидов при основном поиске [68]. Кроме того, профиль модификаций, полученный при помощи открытого поиска, можно использовать для сравнения протеомов разных состояний [69].

Для открытого поиска разработаны специальные инструменты, например, поисковая программа MSFragger [47]. Кроме того, недавно разработанный на её основе инструмент AA_stat [68] 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.

ЗАКЛЮЧЕНИЕ

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 вызывает слишком много нежелательных побочных реакций, и его предпочтительно заменить аналогом – хлорацетамидом или акриламидом. Рациональный выбор алкилирующего агента особенно важен при исследовании посттрансляционных модификаций, а также в протеоминых 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.

ФИНАНСИРОВАНИЕ

Работа К.Г. Кузнецовой, Е.М. Соловьевой и А.В. Кузикова поддержана грантом Российского фонда фундаментальных исследований № 18-315-00226 (руководитель К.Г. Кузнецова). Вклад в статью С.А. Мошковского и М.В. Горшкова поддержан грантом Российского научного фонда № 17-15-01229 (руководитель С.А. Мошковский).

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит каких-либо исследований с использованием людей или с использованием животных в качестве объектов.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА

1. Wilkins M.R., Sanchez J.C., Gooley A.A., Appel R.D., Humphery-Smith I., Hochstrasser D.F., Williams K.L. (1996) *Biotechnol. Gen. Eng. Revs.*, **13**(1), 19-50.
2. Edman P., Högfeldt E., Sillén L.G., Kinell P.-O. (1950) *Acta Chemica Scandinavica*, **4**, 283-293.
3. Laemli U.K. (1970) *Nature*, **227**, 680-685.
4. O'Farrell P.H. (1975) *J. Biol. Chem.*, **250**(10), 4007-4021.
5. Fenn J.B. (2003) *Angewandte Chemie - International Edition*, **42**(33), 3871-3894.
6. Washburn M.P., Wolters D., Yates J.R. (2001) *Nature Biotechnol.*, **19**(3), 242-247.
7. Pitt J.J. (2009) *Clinical Biochemist Revs.*, **30**(1), 19-34.
8. Verheggen K., Martens L., Berven F.S., Barsnes H., Vaudel M. (2016) *Adv. Exper. Med. Biol.*, **919**, 147-156.
9. Silva A.S., Bouwmeester R., Martens L., Degroove S. (2019) *Bioinformatics*, **35**(24), 5243-5248.
10. Tiwary S., Levy R., Gutenbrunner P., Salinas Soto F., Palaniappan K.K., Deming L., Berndt M., Brant A., Cimermancic P., Cox J. (2019) *Nature Methods*, **16**(6), 519-525.
11. Aebersold R., Mann M. (2016) *Nature*, **537**(7620), 347-355.
12. Borrás E., Sabido E. (2017) *Proteomics*, **17**(17-18), 1700180. DOI: 10.1002/pmic.201700180
13. Kelleher N.L. (2004) *Anal. Chem.*, **76**(11), 197A-203A. DOI: 10.1021/ac0415657
14. Moore S.M., Hess S.M., Jorgenson J.W. (2016) *J. Proteome Res.*, **15**(4), 1243-1252.
15. Sechi S., Chait B.T. (1998) *Anal. Chem.*, **70**(24), 5150-5158.
16. Crestfield A.M., Moore S., Stein W.H. (1963) *J. Biol. Chem.*, **238**(2), 622-627.
17. Hao P., Ren Y., Alpert A.J., Sze S.K. (2011) *Mol. Cell. Proteomics*, **10**(10), O111.009381. DOI: 10.1074/mcp.O111.009381
18. Boja E.S., Fales H.M. (2001) *Anal. Chem.*, **73**(15), 3576-3582.
19. Chernobrovkin A.L., Kopylov A.T., Zgoda V.G., Moysa A.A., Pyatnitskiy M.A., Kuznetsova K.G., Ilina I.Y., Karpova M.A., Karpov D.S., Veselovsky A.V., Ivanov M.V., Gorshkov M.V., Archakov A.I., Moshkovskii S.A. (2015) *J. Proteomics*, **120**, 169-178.

20. Ploug M., Stoffer B., Jensen A.L. (1992) Electrophoresis, **13**(1), 148-153.
21. Gundry R.L., White M.Y., Murray C.I., Kane L.A., Fu Q., Stanley B.A., van Eyk J.E. (2009) Curr. Prot. Mol. Biol., **77**(5), 342-355.
22. Cleland W.W. (1964) Biochemistry, **3**(4), 480-482.
23. Getz E.B., Xiao M., Chakrabarty T., Cooke R., Selvin P.R. (1999) Anal. Biochem., **273**(1), 73-80.
24. Rüegg U.T., Rudinger J. (1977) Methods Enzymology, **47**, 111-116.
25. Suttapitugsakul S., Xiao H., Smeekens J., Wu R. (2017) Molecular Biosystems, **13**(12), 2574-2582.
26. Müller T., Winter D. (2017) Mol. Cell. Proteomics, **16**(7), 1173-1187.
27. Manza L.L., Stamer S.L., Ham A.-J.L., Codreanu S.G., Liebler D.C. (2005) Proteomics, **5**(7), 1742-1745.
28. Wiśniewski J.R., Zougman A., Nagaraj N., Mann M. (2009) Nature Methods, **6**(5), 359-362.
29. Jiang X., Shamshurin D., Spicer V., Krokhin O.V. (2013) J. Chromatography B: Anal. Technol. Biomed. Life Sci., **915-916**, 57-63.
30. Pappin D.J.C., Hojrup P., Bleasby A.J. (1993) Curr. Biol., **3**(6), 327-332.
31. Nielsen M.L., Vermeulen M., Bonaldi T., Cox J., Moroder L., Mann M. (2008) Nature Methods, **5**(6), 459-460.
32. Sebastiano R., Citterio A., Lapadula M., Righetti P.G. (2003) Rapid Communications Mass Spectrometry, **17**(21), 2380-2386.
33. Brune D.C. (1992) Anal. Biochem., **207**(2), 285-290.
34. Karala A.R., Ruddock L.W. (2007) Antioxidants Redox Signaling, **9**(4), 527-531.
35. Paulech J., Solis N., Cordwell S.J. (2013) Biochim.Biophys. Acta - Proteins and Proteomics, **1834**(1), 372-379.
36. Xu P., Duong D.M., Seyfried N.T., Cheng D., Xie Y., Robert J., Rush J., Hochstrasser M., Finley D., Peng J. (2009) Cell, **137**(1), 133-145.
37. Kim S., Chen J., Cheng T., Gindulyte A., He J., He S., Li Q., Shoemaker B.A., Thiessen P.A., Yu B., Zaslavsky L., Zhang J., Bolton E.E. (2019) Nucl. Acids Res., **47**(D1), D1102-D1109.
38. Shevchenko A., Wilm M., Vorm O., Mann M. (1996) Anal. Chem., **68**(5), 850-858.
39. Guo M., Weng G., Yin D., Hu X., Han J., Du Y., Liu Y., Tang D., Pan Y. (2015) RSC Advances, **5**(125), 103662-103668.
40. Lundell N., Schreitmüller T. (1999) Anal. Biochem., **266**, 31-47.
41. Yang Z., Attygalle A.B. (2007) J. Mass Spectrometry, **42**(2), 233-243.
42. Woods A.G., Sokolowska I., Darie C.C. (2012) Biochem. Biophys. Res. Commun., **419**(2), 305-308.
43. Krüger R., Hung C.W., Edelson-Averbukh M., Lehmann W.D. (2005) Rapid Communications Mass Spectrometry, **19**(12), 1709-1716.
44. Lapko V.N., Smith D.L., Smith J.B. (2000) J. Mass Spectrometry, **35**(4), 572-575.
45. Schroeder M.J., Shabanowitz J., Schwartz J.C., Hunt D.F., Coon J.J. (2004) Anal. Chem., **76**(13), 3590-3598.
46. Chick J.M., Kolippakkam D., Nusinow D.P., Zhai B., Rad R., Huttlin E.L., Gygi S.P. (2015) Nature Biotechnol., **33**(7), 743-749.
47. Kong A.T., Leprevost F.V., Avtonomov D.M., Mellacheruvu D., Nesvizhskii A.I. (2017) Nature Methods, **14**(5), 513-520.
48. Verheggen K., Raeder H., Berven F.S., Martens L., Barsnes H., Vaudel M. (2017) Mass Spectrometry Revs., 1-15. DOI: 10.1002/mas.21543
49. Brot N., Weissbach H. (1983) Arch. Biochem. Biophys., **223**(1), 271-281.
50. Veredas F.J., Cantón F.R., Aledo J.C. (2017) Sci. Reps., **7**(1), 40403. DOI:10.1038/srep40403
51. Berlett B.S., Stadtman E.R. (1997) J. Biol. Chem., **272**(33), 20313-20316.
52. Lam X.M., Lai W.G., Chan E.K., Ling V., Hsu C.C. (2011) Pharmaceut. Res., **28**(10), 2543-2555.
53. Lagerwerf F.M., van de Weert M., Heerma W., Haverkamp J. (1996) Rapid Communications Mass Spectrometry, **10**(15), 1905-1910.
54. Hains P.G., Robinson P.J. (2017) J. Proteome Res., **16**(9), 3443-3447.
55. Paulo J.A. (2016) J. Biol. Methods, **3**(3), 45.
56. Udeshi N.D., Mertins P., Svinkina T., Carr S.A. (2013) Nature Protocols, **8**(10), 1950-1960.
57. Xu G., Jaffrey S.R. (2013) Biotechnol. Genetic Engineering Revs., **29**(1), 73-109.
58. Stes E., Laga M., Walton A., Samyn N., Timmerman E., de Smet I., Goormachtig S., Gevaert K. (2014) J. Proteome Res., **13**(6), 3107-3113.
59. Zhang B., Pirmoradian M., Chernobrovkin A., Zubarev R.A. (2014) Mol. Cell. Proteomics, **13**(11), 3211-3223.
60. Moshkovskii S.A., Ivanov M.V., Kuznetsova K.G., Gorshkov M.V. (2018) Biochemistry (Moscow), **83**(3), 250-258.
61. Moghaddas Gholami A., Hahne H., Wu Z., Auer F.J., Meng C., Wilhelm M., Kuster B. (2013) Cell Reports, **4**(3), 609-620.
62. Gundlach H.G., Moore S., Stein W.H. (1959) J. Biol. Chem., **234**(7), 1761-1764
63. Ghezzi P., Bonetto V. (2003) Proteomics, **3**(7), 1145-1153.
64. Wojdyla K., Rogowska-Wrzesinska A. (2015) Redox Biology, **6**, 240-252.
65. Hill B.G., Reily C., Oh J.Y., Johnson M.S., Landar A. (2009) Free Rad. Biol. Med., **47**(6), 675-683.
66. Scheerlinck E., Dhaenens M., van Soom A., Peelman L., de Sutter P., van Steendam K., Deforce D. (2015) Anal. Biochem., **490**, 14-19.
67. Ying J., Clavreul N., Sethuraman M., Adachi T., Cohen R.A. (2007) Free Radical Biol. Med., **43**(8), 1099-1108.
68. Bubis J.A., Levitsky L.I., Ivanov M.V., Gorshkov M.V. (2018) Proteomics, **18**(23), 1800117. DOI: 10.1002/pmic.201800117
69. Tarasova I.A., Chumakov P.M., Moshkovskii S.A., Gorshkov M.V. (2019) J. Proteomics, **191**, 16-21.
70. Pilo A.L., McLuckey S.A. (2014) J. Am. Soc. Mass Spectrometry, **25**(6), 1049-1057.

Поступила в редакцию: 27. 11. 2019.
После доработки: 12. 12. 2019.
Принята к печати: 17. 12. 2019.

MODIFICATION OF CYSTEINE RESIDUES FOR MASS SPECTROMETRY-BASED PROTEOMIC ANALYSIS: FACTS AND ARTIFACTS

K.G. Kuznetsova^{1}, E.M. Solovyeva^{2,3}, A.V. Kuzikov^{1,4}, M.V. Gorshkov², S.A. Moshkovskii^{1,4}*

¹Institute of Biomedical Chemistry,

10 Pogodinskaya str., Moscow, 119121 Russia; *e-mail: kuznetsova.ks@gmail.com

²Talrose Institute for Energy Problems of Chemical Physics, Semenov Federal Research Center of Chemical Physics, Russian Academy of Sciences, 38 Leninsky ave., bld.2, Moscow, 119334 Russia

³Moscow Institute of Physics and Technology (State University),

9 Institutsky per., Dolgoprudny, Moscow region, 141700 Russia

⁴Pirogov Russian National Research Medical University,

1 Ostrovityanova str., Moscow, 117997 Russia

Mass spectrometric proteomic analysis at the sample preparation stage involves the artificial reduction of disulfide bonds in proteins formed between cysteine residues. Such bonds, when preserved in their native state, complicate subsequent enzymatic hydrolysis and interpretation of the research results. To prevent the re-formation of the disulfide bonds, cysteine residues are protected by special groups, most often by alkylation. In this review, we consider the methods used to modify cysteine residues during sample preparation, as well as possible artifacts of this stage. Particularly, adverse reactions of the alkylating agents with other amino acid residues are described. The most common alkylating compound used to protect cysteine residues in mass spectrometric proteomic analysis is iodoacetamide. However, an analysis of the literature in this area indicates that this reagent causes more adverse reactions than other agents used, such as chloroacetamide and acrylamide. The latter can be recommended for wider use. In the review we also discuss the features of the cysteine residue modifications and their influence on the efficiency of the search for post-translational modifications and protein products of single nucleotide substitutions.

Key words: proteomics; sample preparation; alkylation; cysteine; mass spectrometry

Funding. The work of K. Kuznetsova, E. Solovyeva and A. Kuzikov is supported by the Russian Foundation for Basic Research, grant No. 18-315-00226 (to K.G. Kuznetsova). A contribution of S. Moshkovskii and M. Gorshkov was supported by Russian Scientific Foundation, grant No. 17-15-01229 (to S.A. Moshkovskii).

Received: 27.11.2019, revised: 12.12.2019, accepted: 17.12.2019.