

О ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ МЕТОДА УЛЬТРАКОРОТКОГО ПОЛНОПРОТЕОМНОГО АНАЛИЗА DirectMS1 В ЗАДАЧАХ ХИМИЧЕСКОЙ ПРОТЕОМИКИ ПО ПОИСКУ МИШЕНЕЙ ЛЕКАРСТВЕННОГО ВОЗДЕЙСТВИЯ

© 2022 Е.М. Соловьева¹, Ю.А. Бубис¹, И.А. Тарасова¹, А.А. Лобас¹, М.В. Иванов¹,
А.А. Назаров², И.А. Шутков², М.В. Горшков^{1*}

¹ Федеральный исследовательский центр химической физики им. Н.Н. Семенова РАН,
Институт энергетических проблем химической физики им. В.Л. Тальрозе,
119334 Москва, Россия; электронная почта: mike.gorshkov@gmail.com

² Московский государственный университет имени М.В. Ломоносова, химический факультет,
119991 Москва, Россия

Поступила в редакцию 16.06.2022

После доработки 20.09.2022

Принята к публикации 03.10.2022

Анализ количественного содержания белков в клетках тканей или физиологических жидкостях, основанный на хроматомасс-спектрометрии, является одним из ключевых источников информации о механизмах жизнедеятельности клеток в условиях химиотерапевтического воздействия. Выявление значимых изменений экспрессии белков решается методами химической протеомики и требует анализа протеомов клеток, подверженных обработке лекарствами, а также разработки экспериментальных и биоинформатических методов и подходов к поиску лекарственных мишеней. При этом производительность полнопротеомного анализа, основанного на жидкостной хроматографии и тандемной масс-спектрометрии, является одним из основных факторов, ограничивающих масштаб таких исследований. Метод прямой масс-спектрометрической идентификации белков DirectMS1 является одним из разработанных в последние годы подходов, позволяющих осуществлять ультракороткий полнопротеомный анализ в режиме минутных градиентов разделения смесей протеолитических пептидов. Целями данной работы являлось выяснение возможностей и ограничений метода для идентификации мишеней лекарственного воздействия на уровне всего клеточного протеома, а также выявления активируемых воздействием клеточных процессов. В частности, в работе был проведен анализ литературных данных химической протеомики для большого набора онкопрепаратов, полученных ранее методом мультиплексного количественного протеомного профилирования, основанного на тандемной масс-спектрометрии в сочетании с высокоэффективной жидкостной хроматографией. Полученные результаты были сопоставлены с данными полнопротеомного анализа, полученными методом DirectMS1 с использованием ультракоротких режимов разделения протеолитических смесей для клеточных линий человека, с целью выяснения возможностей метода идентифицировать известные лекарственные мишени онкопрепаратов. Также, на примере клеточной линии A2780, было проведено полнопротеомное сравнение двух методик клеточного лизиса, включая используемый в химической протеомике лизис, основанный на заморозке-разморозке и стандартную в протеомных исследованиях методику на основе ультразвукового разрушения клеток. Кроме того, были получены результаты экспрессионной протеомики на основе ультракороткого полнопротеомного анализа методом DirectMS1 для клеточной линии A2780, обработанной онкопрепаратом лонидамин, с последующим анализом геномных онтологий, с целью выяснения возможностей метода для выявления регуляции белков в клеточных процессах, ассоциированных с лекарственным воздействием.

КЛЮЧЕВЫЕ СЛОВА: протеомика, масс-спектрометрия, белки, лонидамин, лизис, сигнальные каскады лонидамина.

DOI: 10.31857/S0320972522110094, **EDN:** LWGFCF

Принятые сокращения: A2780 – клеточная линия карциномы яичника человека; DirectMS1 – метод прямой хроматомасс-спектрометрической идентификации белков протеомов; HeLa – клеточная линия рака шейки матки, используемая в протеомике в качестве стандарта клеточной линии человека; GO – анализ геномных онтологий; FC – кратные изменения концентраций белков; FDR – в протеомике означает уровень (в %) присутствия ложно-положительных идентификаций среди всех репортируемых идентификаций белков или пептидов; FT – лизис клеток на основе заморозки-разморозки; MS1 – масс-спектры первого уровня; TMT – тандемные изотопные метки (tandem mass-tag); US – лизис клеток на основе ультразвукового разрушения.

* Адресат для корреспонденции.

ВВЕДЕНИЕ

Выяснение механизмов жизнедеятельности клеток живых организмов и их развития, включая изменения в результате лекарственного воздействия, влияния внешних факторов или физиологического состояния самого организма, является одной из основных задач современных постгеномных исследований и разработок в области «омиксных» технологий, а также их дальнейшего использования в биологии и клинической практике [1], что, в свою очередь, расширит существующие возможности для качественных изменений в подходах к решению фундаментальных и практических проблем медицины, например, ранней диагностики различных форм тяжёлых заболеваний или предсказания эффективности терапевтического воздействия [2, 3]. Одним из основных подходов к выявлению изменений в экспрессии белков, связанных с влиянием перечисленных выше факторов и играющих ключевую роль в патологическом развитии или отклике на воздействие, является качественный и количественный анализ протеомов методами высокоэффективной жидкостной хроматографии в сочетании с тандемной масс-спектрометрией (ВЭЖХ-МС/МС) [4, 5]. Следует отметить, что поиск и анализ регулируемых воздействием белков на уровне всего протеома стали возможными с развитием масс-анализаторов высокого разрешения [6, 7] и, если говорить о химической протеомике, с появлением таких новых методов, как температурное полнопротеомное профилирование (thermal proteome profiling, TPP) [8, 9]. Однако значительные затраты инструментального времени для осуществления полнопротеомного профилирования остаётся серьёзной и актуальной проблемой дальнейшего методического развития химической протеомики и её использования в биомедицинских исследованиях. Действительно, типичное время, затрачиваемое на полнопротеомный анализ клеточных лизатов современными методами ВЭЖХ-МС/МС, составляет от несколько часов до десятков часов на один образец, проанализированный в нескольких технических репликах и биоповторах. В первую очередь такие затраты инструментального времени связаны с длительным разделением анализируемых протеолитических смесей, необходимым для реализации тандемной масс-спектрометрии (МС/МС) и идентификации как можно большего количества компонент смесей [10, 11]. Соответственно, только для одной системы «лекарство–протеом», для которой, например, реализация метода TPP связана с проведением порядка 100 полно-

протеомных хроматомасс-спектрометрических анализов, потребуются существенные затраты инструментального времени, ограничивающие его широкое практическое применение. В этой связи роль быстрого анализа становится принципиальной для выявления регулируемых воздействием белков на больших выборках образцов и систем «лекарство–протеом». Проблема частично решается использованием мультиплексинга на основе изобарных меток (tandem mass tags, TMT) [12], что тем не менее не снимает актуальности задачи сокращения времени полнопротеомного анализа из-за цены такого решения.

Первые работы по разработке методов быстрого полнопротеомного хроматомасс-спектрометрического анализа для широкомасштабных скрининговых исследований в области сравнительной протеомики с целью поиска биомаркеров появились буквально в последние несколько лет, включая исследования, посвящённые проблеме повышения скорости такого анализа [13–16]. Метод прямой хроматомасс-спектрометрической идентификации DirectMS1 [17, 18] является одним из таких подходов к реализации быстрого количественного полнопротеомного анализа за счёт использования ультракоротких градиентов разделения протеолитических смесей. Так, метод позволяет идентифицировать более 2000 белков за время анализа в несколько минут при использовании масс-спектрометров высокого разрешения на основе орбитальной ионной ловушки Орбитрэп [19]. Понятие «прямая» масс-спектрометрическая идентификация подразумевает исключение из анализа стадии последовательного получения масс-спектров фрагментации протеолитических пептидов смеси, являющейся одной из основных причин увеличения времени анализа в стандартных протеомных методах. Идентификация белков осуществляется за счёт использования всей совокупности экспериментальных данных о пептидах, комплементарных к масс-спектрометрическим, в частности, хроматографическим временам, ионной подвижности и др. Важным преимуществом метода DirectMS1 является возможность идентификации белков с существенно большим покрытием последовательности детектируемыми протеолитическими пептидами, что даёт более точную количественную информацию [19]. Однако возможности использования метода DirectMS1 в задачах химической протеомики, в частности, идентификации белков-мишеней лекарственного воздействия и выявления регулируемых воздействием процессов, до недавнего времени не исследовались.

В работе нами был проведён цикл исследований возможности использования метода прямой масс-спектрометрической идентификации белков DirectMS1 с использованием ультракоротких градиентов разделения в химической протеомике. Конкретными задачами исследования были следующие: (1) проверка возможностей метода для мониторинга мишеней лекарственных воздействий; (2) сравнение наборов белков, идентифицируемых в рамках полнопротеомного анализа, для различных методик клеточного лизиса, включая наиболее широко используемый в химической протеомике лизис на основе заморозки-разморозки (FT, freeze-thaw lysis) [20] и стандартный лизис на основе ультразвукового разрушения клеток (US, ultrasound-induced lysis); (3) выявление белков, регулируемых лекарственным воздействием. Для реализации первой задачи использовались экспериментальные данные, полученные ранее в Каролинском институте Стокгольма для большого набора онкопрепаратов методом мультиплексного глубокого количественного протеомного профилирования клеточной линии A549 [21], а также данные полнопротеомного анализа, полученные методом ультракороткого DirectMS1 как ранее для коммерческого гидролизата клеточной линии HeLa [19], так и в данной работе для гидролизата клеточной линии рака яичников A2780. Для реализации второй задачи использовались данные, полученные в работе двумя методами полнопротеомного анализа, включая стандартный ВЭЖХ-МС/МС и DirectMS1 на основе ультракоротких градиентов, для гидролизатов клеточной линии A2780 с использованием двух методик клеточного лизиса. Наконец, третья задача исследования решалась на примере полнопротеомного анализа клеточной линии рака яичников A2780, обработанной препаратом лонидамин (lonidamine), с последующим анализом генных онтологий для выявления обогащённых клеточных компонентов, регулируемых лекарственным воздействием.

МАТЕРИАЛЫ И МЕТОДЫ

Культивирование клеток. Клетки карциномы яичников человека A2780, полученные из Европейской коллекции аутентичных клеточных культур (ECACC, Великобритания), культивировали в стандартной среде RPMI 1640 («Gibco™», Ирландия), содержащей 5%-ную эмбриональную сыворотку телят («Gibco™», Бразилия) и смесь антибиотиков (пенициллин – 100 мкг/мл, стрептомицин –

100 мкг/мл, «ПанЭко», Россия) при 37 °С в атмосфере 5% CO₂. В экспериментах были использованы клетки после 5–7 пассажей. Для получения клеточной массы клетки высевали в культуральные флаконы 25 см² («TPP», Швейцария) (1 × 10⁶ клеток в 5 мл культуральной среды) и инкубировали 24 ч. В день эксперимента среду заменяли на 5 мл среды с содержанием активного соединения лонидамин (полученный в МГУ по опубликованной ранее методике [22]) в концентрации 150 мкМ (IC₅₀). Далее, клетки инкубировали в течение 72 ч, среду удаляли, клетки отделяли от поверхности скребком, суспендировали в свежей культуральной среде, центрифугировали (1500 об./мин, 3 мин), а затем дважды промывали холодным 10 мМ PBS (pH 7,4). В качестве контроля использовали клетки во флаконах, не подвергавшиеся обработке лонидамином. Все эксперименты проводились в пяти повторах.

Клеточный лизис. Белки клеток выделялись двумя сравниваемыми методами: (1) с использованием ультразвука (US) и (2) методом заморозки-разморозки (FT). В первом методе использовался стандартный протокол лизиса, включающий лизирующий буфер – 50 мМ аммоний бикарбонат (ABC), 10% ацетонитрила, 0,1% Protease MAX Surfactant («Promega», США). На 1 млн клеток добавляли 100 мкл лизирующего буфера с последующей инкубацией на термошейкере (500 об./мин) при комнатной температуре в течение 30 мин. Дальнейший лизис проводили с использованием ультразвукового гомогенизатора Qsonica Q125 («Qsonica», США) в течение 3 мин (циклами по 10 с) на льду с амплитудой 25%. Для дальнейшего ферментативного гидролиза пробы были разбавлены в 2 раза 50 мМ ABC.

Для лизиса методом заморозки-разморозки использовали 1 × PBS (pH 7,4); 100 мкл буфера на 1 млн клеток. Пробу помещали в жидкий азот на 1 мин, затем нагревали до 25 °С на термошейкере (600 об./мин) в течение 1 мин. Цикл заморозки-разморозки повторяли 3 раза, после чего пробы центрифугировали (10 000 g) в течение 5 мин. Супернатант переносили в новые пробирки для дальнейшего ферментативного гидролиза.

Подготовка протеомных проб для хромато-масс-спектрометрического анализа. Для восстановления дисульфидных связей к пробам добавлялся дитиотреитол (DTT, финальная концентрация 10 мМ) с последующим нагреванием до 56 °С (25 мин), после чего проводили реакцию алкилирования с использованием йодацетамида (IAA, финальная концентрация 15 мМ) в отсутствии освещения при ком-

натной температуре в течение 30 мин. Для устранения возможного избытка алкилирующего реагента к пробам добавляли DTT в финальной концентрации 5 мМ. Для ферментативного гидролиза был использован трипсин (Sequencing Grade Modified Trypsin, «Promega») в соотношении 1/75, реакцию проводили при температуре 37 °С в течение 15 ч. Для прекращения гидролиза к пробам была добавлена уксусная кислота (3–5%, финальный pH < 3). Через 30 мин проводили центрифугирование проб (12 000 g, 10 мин). Супернатант перенесли в стеклянные пробирки и высушивали с использованием вакуумного центрифужного испарителя. Полученные пробы хранились при температуре –80 °С. Непосредственно перед масс-спектрометрическим анализом пробы растворяли в 0,1%-ном водном растворе муравьиной кислоты до концентрации пептидов 1 мкг/мкл.

Хроматомасс-спектрометрический анализ протеолитических смесей пептидов протеомов.

Полнопротеомный анализ клеточных лизатов осуществлялся в работе двумя методами: (1) стандартный полнопротеомный анализ «снизу-вверх» (bottom-up) [23], основанный на гидролизе белков, разделении получаемых смесей протеолитических пептидов обращённо-фазовой хроматографией и идентификации пептидов на основе MS/MS и (2) метод прямой хроматомасс-спектрометрической идентификации белков DirectMS1, основанный на ультракоротком жидкостно-хроматографическом разделении смесей протеолитических пептидов и измерении масс-спектров первого уровня (MS1) высокого разрешения ионов пептидов без использования фрагментации [18].

Протеомный анализ на основе MS/MS осуществляли в двух технических повторах с использованием хроматографической ВЭЖХ-системы Ultimate 3000 RSLCnano («Thermo Scientific», США), соединённой с масс-спектрометром Q-Exactive HFX («Thermo Scientific»). Пептидную смесь (500 нг) загружали на обогащающую колонку Acclaim μ -Precolumn (0,5 мм × 3 мм, размер частиц – 5 мкм, «Thermo Scientific») при потоке 10 мкл/мин в течение 4 мин в изократическом режиме (2% ацетонитрила, 0,1% муравьиной кислоты в деионизованной воде). Далее пептиды разделяли на ВЭЖХ-колонке C18 (75 мкм × 150 мм, фаза Reprosil-Pur, размер частиц – 3 мкм, размер пор – 100 Å) в градиентном режиме элюирования. Градиент формировали подвижной фазой А (0,1%-ная муравьиная кислота) и подвижной фазой Б (80% ацетонитрила, 0,1% муравьиной кислоты в деионизованной воде)

при скорости потока 400 нл/мин. Колонку уравнивали при 2% подвижной фазы Б в течение 3 мин, после чего линейно увеличивали концентрацию подвижной фазы Б до 35% за 28 мин, затем линейно увеличивали концентрацию фазы Б до 99% за 5 мин, после 5-минутной промывки при 99% фазы Б концентрацию буфера линейно снижали до исходных 2% подвижной фазы Б за 1 мин. Общая продолжительность анализа составляла 45 мин. Масс-спектрометрический анализ проводили на масс-спектрометре Q-Exactive HFX в режиме положительной ионизации с использованием источника NESI («Thermo Scientific»). Для масс-спектрометрического анализа были установлены следующие параметры настроек: напряжение на эмиттере – 2,3 кВ; температура капилляра – 285 °С. Панорамное сканирование проводили в диапазоне масс 350–1500 m/z при разрешении 120,000. При тандемном сканировании, используемом для наработки данных протеомного анализа методом DDA (data dependent acquisition) [24], масс-спектральное разрешение устанавливали 15 000 в диапазоне масс от 100 m/z до верхней границы, которая определяется автоматически, исходя из массы прекурсора, но не более 2000 m/z . Изоляцию прекурсорных ионов проводили в окне ± 1 Да. Максимальное число разрешённых для изоляции ионов в режиме получения спектров фрагментации (масс-спектры второго уровня MS/MS) – не более 20, нормализованная энергия соударения (normalized collision energy NCE) равнялась 29. Для прекурсорных ионов, отправляемых на фрагментацию, учитывали только ионы от $z = 2+$ до $z = 6+$ по зарядовому состоянию. Максимальное время накопления для прекурсорных ионов составляло 50 мс, для фрагментных ионов – 110 мс. Величину общего накапливаемого в ловушке заряда (automatic gain control, AGC) для прекурсоров и фрагментных ионов устанавливали соответственно $1 \cdot 10^6$ и $2 \cdot 10^5$. Все ранее проанализированные ионы-прекурсоры динамически исключались на 30 с из списка кандидатов, отбираемых на изоляцию и фрагментацию. Данные протеомного анализа гидролизатов белков анализируемых клеточных линий, полученные стандартным методом, описанным выше, обрабатывались с целью получения идентификаций белков.

Протеомный анализ пептидов с использованием метода DirectMS1 осуществляли для каждой биологической реплики в трёх технических повторах на хроматографической ВЭЖХ-системе Ultimate 3000 RSLCnano, соединённой с масс-спектрометром Q-Exactive HFX.

Пептидную смесь (200 нг) загружали на обогащающую колонку Acclaim μ -Precolumn (0,5 мм \times 3 мм, размер частиц – 5 мкм) при потоке 10 мкл/мин в течение 4 мин в изократическом режиме (2% ацетонитрила, 0,1% муравьиной кислоты в деионизованной воде). Далее пептиды разделяли на обращённо-фазовой ВЭЖХ-колонке C18 (75 мкм \times 150 мм, фаза Reprosil-Pur, размер частиц – 3 мкм, размер пор – 100 Å) в градиентном режиме элюирования. Градиент формировали подвижной фазой А (0,1% муравьиной кислоты) и подвижной фазой Б (80% ацетонитрила, 0,1% муравьиной кислоты в деионизованной воде) при скорости потока 1,5 мкл/мин. Хроматографический градиент был сформирован следующим образом: сначала колонку эквilibрировали при 8% подвижной фазы Б в течение 0,8 мин, после чего линейно увеличивали концентрацию подвижной фазы Б до 35% за 4,2 мин, затем линейно увеличивали концентрацию фазы Б до 99% за 0,9 мин, после 0,9-минутной промывки при 99% фазы Б, концентрацию буфера линейно снижали до исходных 8% подвижной фазы Б за 0,1 мин. Общая длительность анализа, включая стадии промывки и уравнивания колонки составляла 7,3 мин.

Идентификация белков. Во всех случаях, здесь и далее, при идентификации белков использовалась библиотека канонических последовательностей человека SwissProt от 30.07.2021 с ограничением по видовой принадлежности исследуемого организма *Homo sapiens*, содержащей 20 368 белков. Идентификацию белков в данных, полученных с использованием МС/МС, проводили при помощи поисковой протеомной машины IdentiPy v.0.3.3 [25] и алгоритмом послепоисковой валидации Scavager v.0.2.9 [26] с интегрированным в него по умолчанию алгоритмом NSAF (normalized spectral abundance factor) [27], выбранным ранее как наиболее точный алгоритм безметочного полуколичественного анализа на основе спектров МС/МС [28]. Для поиска протеолитических пептидов задавались следующие входные параметры: фермент – трипсин, ограничения на длину пептидов – 7–30 аминокислотных остатков, точность определения масс моноизотопных пептидов в спектрах МС1 – ± 5 ppm, точность определения масс в спектрах МС/МС – $\pm 0,01$ Да, а также возможность пропуска до двух сайтов гидролиза белка трипсином. Окисление метионина, *N*-концевое ацетилирование белка и модификация цистеина алкилирующим реагентом были учтены как возможные модификации пептидов. Для идентификации сопоставлений (образования пар) спектров и

пептидов (peptide-spectrum matches, PSM), а также пептидов и белковых групп использовали величину ложноположительных идентификаций (false discovery rate, FDR) не более 1,0%. Для сравнения белков, идентифицированных с использованием различных методов лизиса, технические и биологические реплики были объединены и совместно отфильтрованы до такого же уровня FDR (функция union алгоритма Scavager v.0.2.9) [26].

Идентификацию белков в данных, полученных методом DirectMS1, проводили при помощи программы ms1searchpy v.2.0.3 [17], усовершенствованной путём её интеграции с алгоритмами машинного обучения [18, 19]. Ложные последовательности для последующей валидации результатов генерировались путём полного перемешивания исходных белковых последовательностей. Были заданы следующие поисковые параметры: фермент – трипсин; минимум 3 скана на пептидный ионный кластер, формируемый распределением пептидов по изотопам ^{13}C ; количество пиков в кластере не менее 2, включая моноизотопный; начальная точность масс – 8 ppm; рассматривались пептиды без пропусков сайтов гидролиза длиной 7–30 аминокислотных остатков (484 765 пептидов), зарядовые состояния которых варьировались в диапазоне от 1+ до 6+. Список белков фильтровался до уровня ложноположительных идентификаций FDR не более 1,0%. Для сравнения различных методов лизиса были проанализированы все 5 биологических повторов, и использовались белки, идентифицированные как минимум в 4-х из них. Идентифицированные белки репортировались как лидеры белковых групп, в которые включались белки с пересекающимися пептидами. При этом в последующем полуколичественном анализе участвовали все белки группы. В частности, для МС1-данных в методе DirectMS1 количественный анализ выполняется следующим образом [18]: на первой стадии собирается список всех достоверно (с заданным уровнем FDR) идентифицируемых белков (лидеров белковых групп) для технических/биологических повторов и образцов, используемых в последующем количественном анализе; далее, для этих белков составляется список всех идентифицированных пептидов; и, на финальной стадии, для этих пептидов составляется список всех белков из базы данных, от которых эти пептиды могли образоваться. Таким образом, в полуколичественном анализе для статистических тестов участвовали не только лидеры белковых групп, но и все белки, пересекающиеся с ними по пептидам.

Анализ данных химической протеомики Каролинского института. Данные по химической протеомике для большого набора систем «лекарство–протеом», проанализированных ранее мультиплексным методом на основе тандемных изотопных меток (ТМТ-меток) в ЦКП молекулярной биометрии Каролинского института [21], были скачаны в сыром виде (доступны в хранилище данных консорциума ProteomeXchange под идентификатором PXD009775) и проанализированы с использованием поискового алгоритма IdentiPy [25]. Для поиска пептидов точности масс ионов-предшественников и фрагментов были установлены, соответственно, на уровне 10 ppm и 0,01 Да. Данные были получены с использованием мультиплексинга на основе ТМТ-меток (ТМТ 10) соответственно, модификация с массовым сдвигом 229,16 Да была задана, как фиксированная на всех *N*-концах пептидов и остатках лизина, так же, как и карбамидометилирование остатков цистеина. Окисление остатков метионина и фосфорилирование остатков серина, треонина и тирозина были заданы как потенциальные модификации. Результаты поиска для отдельных фракций были объединены и валидированы с использованием программы Scavenger [26]. Для оценки относительного содержания белка была использована сумма его интенсивностей во всех 10 ТМТ-каналах, усреднённая для трёх реплик – такая величина не отражает реальное содержание белка в конкретном образце, но характеризует его среднюю концентрацию, поэтому она была использована для ранжирования известных мишеней препаратов и отбора мишеней топ-15.

Отбор дифференциально регулированных белков. Для полуколичественного анализа идентификаций белков, полученных методом DirectMS1, использовался алгоритм Diffacto [29]. Интенсивности пептидов, на основе которых Diffacto выполняет статистический анализ, извлекались с помощью алгоритма Biosaur v.2.0 [30]. Значения *p*-value для белков рассчитывались с использованием метода PECA (Probe-level Expression Change Averaging), интегрированного в Diffacto, и корректировались на множественные сравнения методом Бенджамини–Хохберга. Отбор дифференциально регулированных белков был выполнен, согласно следующим критериям: $FDR_{BH} < 0,001$; $FC_{up} > 1,2$; $FC_{down} < 1/1,2$. При оценке ответа клеток A2780 на лонидамин методом DirectMS1 анализ генов онтологий (GO) выполнялся отдельно для белков с повышенной ($FC_{up} > 1,2$) и пониженной

($FC_{down} < 1/1,2$) регуляцией с помощью программы GOrilla [31]. Обогащения GO вычислялись против всех идентифицированных белков, попавших в результаты полуколичественного анализа (background).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

Для анализа возможности метода прямой масс-спектрометрической идентификации белков DirectMS1 на основе ультракоротких градиентов разделения использовались данные по химической протеомике для большого набора систем «лекарство–протеом», проанализированных ранее мультиплексным методом на основе ТМТ-меток в ЦКП молекулярной биометрии Каролинского института [21]. Данные были получены для нескольких десятков как широко используемых в клинической практике, так и перспективных, проходящих начальные стадии доклинических испытаний, противоопухолевых химиотерапевтических препаратов, которыми обрабатывались клетки линии A549 аденокарциномы лёгких человека коллекции ATCC. В табл. 1 представлен список использовавшихся лекарств и их мишени (коллонки «Лекарство» и «ID белков-мишеней»). Для сравнения в результатах мультиплексного протеомного ВЭЖХ-МС/МС-анализа были выбраны те из мишеней, которые входят в список 15 мишеней с наибольшими значениями относительных концентраций (топ-15) из всех найденных. В протеомных данных, полученных с использованием метода DirectMS1 и ультракоротких градиентов разделения осуществлялся поиск белков, также относящихся к списку мишеней топ-15. В частности, в табл. 1 представлены результаты DirectMS1-анализа как клеточной линии HeLa, полученные ранее с использованием масс-спектрометра Orbitrap Fusion («Thermo Scientific»), оборудованного системой разделения ионов пептидов по подвижности FAIMS [19], так и клеточной линии A2780, полученные в данной работе с использованием двух разных методик лизиса. Следует отметить, что результаты, полученные для стандартного гидролизата клеточной линии HeLa, являются примером наиболее глубокого протеома, достигаемого на сегодняшний день при использовании метода DirectMS1 и 5-минутных градиентов разделения. Как видно из представленных данных, при глубине анализа порядка 2000 белков, достижимого методом DirectMS1, идентифицируется большинство белков-мишеней из списка топ-15 для лекарств, представленных в табл. 1.

Таблица 1. Анализ данных по химической протеомике для наборов систем «лекарство–протеом», проанализированных методом мультиплексного ВЭЖХ-МС/МС анализа на основе ТМТ-меток [21], возможности метода DirectMS1 с использованием ультракоротких градиентов разделения на примере протеомного анализа стандартного гидролизата клеточной линии HeLa, а также сравнение чувствительности метода ультракороткого полнопротеомного количественного профилирования на примере клеточной линии A2780 при мониторинге белков-мишеней лекарственного воздействия при использовании различных методик клеточного лизиса

Лекарство	ID белков-мишеней	Методы протеомного анализа		Методы клеточного лизиса	
		мультиплексный (ТМТ-метки), 90 мин ВЭЖХ-МС/МС, гидролизат линии A549	5 мин ВЭЖХ-DirectMS1, гидролизат линии HeLa**	FT-лизис, 5 мин ВЭЖХ-DirectMS1, гидролизат линии A2780	US-лизис, 5 мин ВЭЖХ-DirectMS1, гидролизат линии A2780
Lapatinib	P04626, P00533	P00533	P00533		
Afatinib	P04626, Q15303, P00533	P00533	P00533		
Dasatinib	P00519, P12931, P29317, P06239, P07947, P10721, P09619, P51692, P42684, P06241, Q06187, Q92570, P11274, P41240, P54756, P54760, P09769, P42685, P11142, P07948, Q9NYL2, Q16539, Q06203	P11142, Q06203		P11142, Q06203	P07947, P11142, Q9NYL2, Q06203
Gefitinib	P00533	P00533	P00533		
OSI-420	P00533	P00533	P00533		
Fludarabine	P09884, P23921, P27707	P23921	P09884	P27707	P23921, P27707
8-Azaguanine	P00491	P00491	P00491	P00491	P00491
Vincristine	P07437, P68366	P07437, P68366	P07437	P07437	P07437
Paclitaxel	P10415, Q9H4B7, O75469, P27816, P11137, P10636	P27816, Q9H4B7		P27816	P27816
Docetaxel	P10415, Q9H4B7, O75469, P27816, P11137, P10636	P27816, Q9H4B7		P27816	P27816
Genistein	Q92731, P03372, P11388, Q14289, Q15596, Q15788	P11388	P11388		
Epirubicin	O14646, P11388	P11388	O14646, P11388		O14646
Doxorubicin	P11388	P11388	P11388		
Etoposide	P11388, Q02880	P11388, Q02880	P11388		Q02880
Idarubicin	P11388	P11388	P11388		
Teniposide	P11388	P11388	P11388		
Topotecan	Q969P6, P11387	P11387	P11387	P11387	P11387
Irinotecan	Q969P6, P11387	P11387	P11387	P11387	P11387
Camptothecin	P11387	P11387	P11387	P11387	P11387
Bortezomib	P28074, P20618	P20618	P28074	P28074, P20618	P28074, P20618
b-AP15	Q9Y5K5, P54578	P54578	Q9Y5K5	Q9Y5K5, P54578	Q9Y5K5, P54578
TRi-1	Q16881	Q16881	Q16881	Q16881	Q16881
TRi-2	Q16881	Q16881	Q16881	Q16881	Q16881
OSW-1	P22059	P22059	P22059		P22059
Lonidamin*	P19367, P13569	P19367			

Примечание. В таблице представлены известные белки-мишени различных онкопрепаратов (вторая колонка), а также те из них, которые «видны» тем или иным методом протеомного анализа и/или лизиса.

* Препарат лонидамин не анализировался в работе Saei et al. [21].

** Использовались экспериментальные данные для клеточной линии HeLa, полученные ранее методом DirectMS1 [19] (идентификатор данных PXD023977 в депозитарии PRIDE).

Ещё одной задачей исследования было сравнение различных методик клеточного лизиса, включая методику на основе ультразвукового (US) разрушения клеток, а также методику, основанную на криогенной заморозке клеток с последующей разморозкой (FT). Сравнение протоколов лизиса было проведено для необработанного образца клеточной линии A2780 с использованием как стандартного протеомного анализа на основе тандемной масс-спектрометрии ВЭЖХ-МС/МС с 30-минутными градиентами разделения, так и DirectMS1 с 5-минутными градиентами разделения (см. выше). Протокол лизиса FT используется в задачах химической протеомики при

поиске мишеней лекарственного воздействия с целью реализации взаимодействия лекарственных препаратов с белками в их нативной форме [8]. На рис. 1 представлены результаты сравнения двух методик лизиса с использованием протеомного анализа на основе тандемной масс-спектрометрии (рис. 1, а) и метода DirectMS1 (рис. 1, б). Как видно из полученных результатов, использование лизиса FT даёт несколько меньшую глубину анализа (белковые группы, теряемые в МС/МС-анализе при использовании лизиса FT, приведены в табл. S1 в Приложении). Однако следует отметить, что наблюдаемое уменьшение глубины анализа ожидаемо происходит в основном за счёт

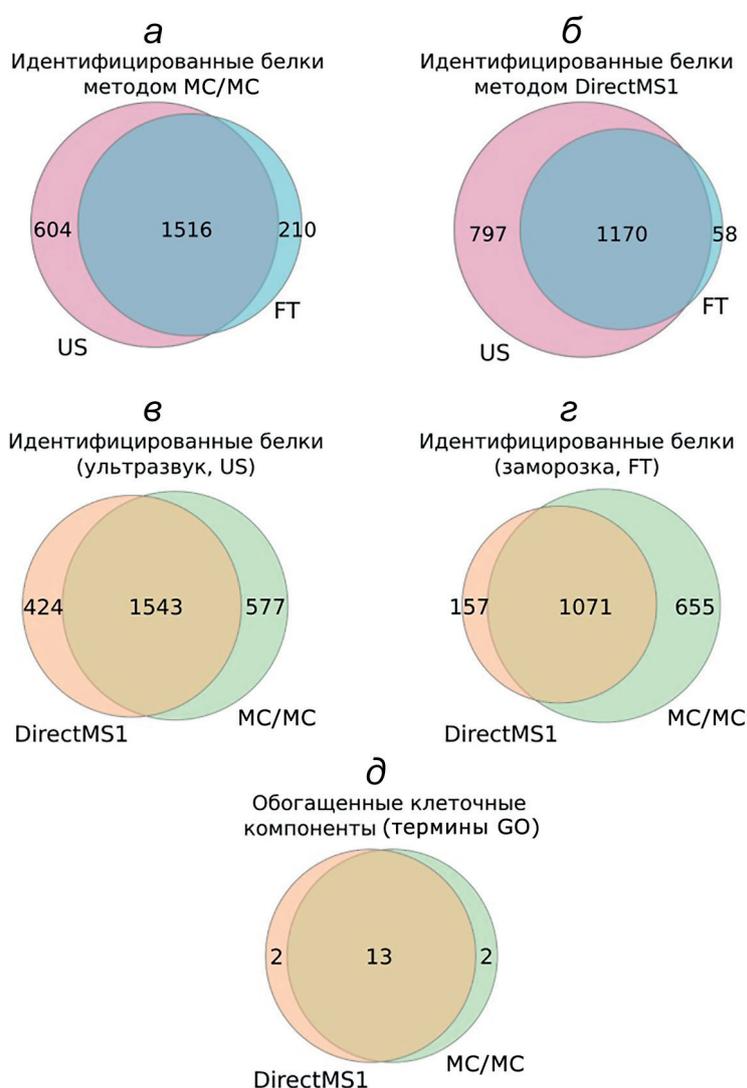


Рис. 1. Результаты сравнения различных методик лизиса, включая стандартный протокол на основе ультразвукового разрушения клеток (US) и протокол, основанный на заморозке клеток в жидком азоте с последующей медленной разморозкой (FT). Сравнение глубины полнопротеомного анализа клеточной линии A2780 при использовании двух протоколов клеточного лизиса с использованием протеомного анализа на основе тандемной масс-спектрометрии МС/МС (а) и метода DirectMS1 (б). Сравнение идентифицированных белков (лидеров белковых групп), полученных методами DirectMS1 и МС/МС с использованием US-лизиса (в) и FT-лизиса (г). д – Пересечение 15 наиболее статистически значимых обогащённых клеточных компонентов для белков, теряемых при использовании протокола лизиса FT относительно контроля (подробнее см. табл. 2). Анализ проведён с использованием утилиты GOzilla

мембранных белков. Действительно, согласно анализу обогащения по функциональной принадлежности с использованием утилиты GOrilla [31], подавляющее большинство белков, идентифицированных исключительно с использованием классического метода лизиса US и «теряемых» при переходе на методику FT, относятся к мембранным клеточным компонентам, таким как клеточная стенка и митохондриальная мембрана (табл. 2). Пересечение идентифицированных белков (в данном случае лидеров белковых групп) между методами на основе MC/MC и DirectMS1 составило порядка 80% (рис. 1, в и г), и вне зависимости от метода анализа наблюдается обогащение мембранных компонент (рис. 1, д). Стоит отметить, что белки, уникальные для метода лизиса FT, имеют в основном цитоплазматическую локализацию и не относятся к определённому клеточному компартменту. Таким образом, их идентификация, вероятно, связана с уменьшением сложности белковой смеси за счёт отсутствия значительного количества мембранных белков. Следует отметить, что анализ генных онтологий (Gene Ontology, GO) [32], результаты которого приведены в табл. 2, осуществлялся для данных, полученных стандартным ВЭЖХ-MC/MC, с использованием только лидеров белковых групп, представляющих собой наиболее вероятные белки, присутствующие в образце. В то же время количественный анализ в алгоритме Diffacto, используемом в методе DirectMS1, осуществляется для всех белков группы [18, 19]. Однако такое различие в

методах не повлияло на результаты сравнения, поскольку подавляющее большинство теряемых при использовании методики FT-лизиса белковых групп состояло, по результатам MC/MC-анализа, из одного белка (80% белковых групп) либо из двух гомологичных белков (10% белковых групп), имеющих одинаковые термины GO.

Исследование возможностей развиваемого метода прямой масс-спектрометрической идентификации белков DirectMS1 на основе ультракоротких градиентов разделения для решения задач выявления каскадов внутриклеточных взаимодействий, регулируемых лекарственным воздействием, осуществлялось на примере полнопротеомного анализа клеточной линии рака яичников A2780, обработанной препаратом лонидамин. Считается, что метаболизм раковых клеток характеризуется повышенным поглощением и потреблением глюкозы – феноменом, известным как эффект Варбурга (Warburg effect). Постоянная активация анаэробного гликолиза в раковых клетках может быть связана с активацией онкогенов или потерей опухолевых супрессоров, тем самым существенно ускоряя прогрессирование рака. В этом отношении ингибирование гликолитической способности в злокачественных клетках может усиливать эффект противоопухолевой комбинированной химиотерапии [33]. Лонидамин является производным индазола, и считается, что механизм его действия основан на ингибировании ассоциированной с митохондриями гексокиназы-II [34, 35]. Также ряд

Таблица 2. Анализ обогащения по функциональной принадлежности (клеточным компонентам) белков, теряемых при использовании протокола лизиса на основе заморозки-разморозки относительно контроля

Термины GO	Описание процесса	<i>q</i> -value, DirectMS1	<i>q</i> -value, MC/MC
GO:0016021	компоненты мембраны	2,93E-42	1,9E-60
GO:0031224	внутренние компоненты мембраны	1,31E-41	9,55E-60
GO:0044425	части мембраны	3,53E-19	6,01E-40
GO:0031090	мембрана (органелла)	7,88E-11	8,92E-17
GO:0031966	митохондриальная мембрана	1,56E-9	7,89E-16
GO:0005789	мембрана эндоплазматического ретикулума	2,65E-8	1,18E-11
GO:0044455	части митохондриальной мембраны	4,76E-6	1,02E-16
GO:0005743	внутренняя митохондриальная мембрана	8,48E-6	5,67E-12
GO:0019866	внутренняя мембрана органелл	1,41E-5	4,04E-10
GO:0098796	мембранный белковый комплекс	1,55E-5	2,63E-10

Примечание. В таблице приведены 10 наиболее значимых компонентов (терминов GO), демонстрирующих обогащение в двух методах анализа (на основе тандемной масс-спектрометрии MC/MC и метода DirectMS1 с использованием ультракоротких градиентов, рис. 1, в). Представленные в таблице *q*-value были рассчитаны утилитой GOrilla и представляют собой *p*-value обогащённых процессов, скорректированных на множественные сравнения.

более ранних исследований показал, что лонидамин ингибирует и гликолиз, что приводит к снижению синтеза внутриклеточного АТФ [36]. Перечисленные возможные механизмы действия лонидамина обозначили его, как один из компонентов комбинированной химиотерапии солидных опухолей, включая рак яичников [37, 38]. В данном контексте в работе был проведён сравнительный анализ протеомов клеточной линии рака яичников, обработанной лонидамином (в концентрации IC₅₀), относительно контроля с использованием развиваемого метода ультракороткого полнопротеомного анализа DirectMS1. Иными словами, интерес представляло увидеть регуляцию белков, характерных для метабо-

лических путей, связанных с гликолизом и синтезом АТФ для относительно хорошо охарактеризованной системы «лекарство–рак яичников» [39]. На рис. 2 представлены результаты анализа протеомов клеточной линии рака яичников A2780, обработанной лонидамином, в сравнении с контрольной, не обработанной лекарством, линией с использованием 5-минутного ВЭЖХ-DirectMS1-метода. На основе полученных диаграмм рассеяния количественных изменений содержания белков, идентифицированных с использованием метода DirectMS1 и геновых онтологий, можно сделать вывод о том, что выявленные регулируемые белки, статистически значимо повысившие концентрацию, как и ожидалось, связаны

Таблица 3. Анализ геновых онтологий с использованием утилиты GOzilla для белков, статистически значимо повысивших концентрацию после обработки лекарством (белки в правом верхнем квадранте диаграммы рассеяния на рис. 2)

Термины GO	Описание процесса	<i>q</i> -value	Степень обогащения
GO:0006735	восстановление кофермента дегидрогеназ (NADP)	0,044	10,04
GO:0006007	катаболический процесс разложения глюкозы	0,041	10,04
GO:0061718	катаболический процесс разложения глюкозы с образованием пирувата	0,039	10,04
GO:0061621	канонический гликолиз	0,037	10,04
GO:0061620	гликолитический процесс через глюкозо-6-фосфата	0,050	9,32
GO:0061615	гликолитический процесс через фруктозо-6-фосфата	0,048	9,32
GO:0006757	получение АТФ из ADP	0,036	8,7
GO:0006096	гликолитический процесс	0,032	8,7
GO:0006094	глюконеогенез	0,035	8,24
GO:0019320	катаболический процесс разложения гексозы	0,033	8,24
GO:0042866	процесс биосинтеза пирувата	0,031	8,24
GO:0006165	фосфорилирование нуклеозиддифосфата	0,015	8,03
GO:0046939	фосфорилирование нуклеотидов	0,013	7,73
GO:0019319	процесс биосинтеза гексозы	0,043	7,46
GO:0009206	процесс биосинтеза пуририбонуклеозидтрифосфата	0,018	7,2
GO:0006754	процесс биосинтеза АТФ	0,050	7,12
GO:0046031	метаболический процесс ADP	0,048	7,12
GO:0046364	процесс биосинтеза моносахаридов	0,030	7,03
GO:0009145	процесс биосинтеза пуририбонуклеозидтрифосфата	0,019	6,96
GO:0016051	процесс биосинтеза углеводов	0,028	6,91
GO:0009201	процесс биосинтеза рибонуклеозидтрифосфата	0,021	6,73
GO:0009132	процесс метаболизма нуклеозиддифосфата	0,033	6,14
GO:0009142	процесс биосинтеза нуклеозидтрифосфата	0,034	5,8
GO:0016310	фосфорилирование	0,030	2,73

Примечание. Приведены процессы с индексом достоверности *q*-value < 0,05 в порядке убывания степени обогащения. Анализ геновых онтологий показал регулирование белков, связанных с метаболическими путями гликолиза и синтеза АТФ, ассоциированных ранее с механизмом действия лонидамина (ячейки, выделенные серым фоном).

Таблица 4. Анализ генных онтологий с использованием утилиты GOzilla для белков, статистически значимо понизивших концентрацию после обработки лекарством (левый верхний квадрант на рис. 2)

Термины GO	Описание процесса	q-value	Степень обогащения
GO:0045814	отрицательная эпигенетическая регуляция экспрессии генов	0,040	8,02
GO:0043484	регуляция сплайсинга РНК	0,001	5,33
GO:0048024	регуляция сплайсинга мРНК через сплайсосому	0,039	4,94
GO:0000377	сплайсинг РНК посредством реакций переэтерификации с аденозином в качестве нуклеофила	0,001	3,56
GO:0000398	сплайсинг мРНК с участием сплайсосомы	0,001	3,56
GO:0000375	сплайсинг РНК посредством реакций переэтерификации	0,001	3,5
GO:0008380	сплайсинг РНК	0,001	3,35
GO:0006397	процессинг мРНК	0,000	3,22
GO:0016071	метаболизм мРНК	0,001	2,55
GO:0006396	процессинг РНК	0,003	2,34
GO:0016070	метаболизм РНК	0,002	2,03
GO:0051252	регуляция метаболизма мРНК	0,006	1,89
GO:0090304	метаболизм нуклеиновых кислот	0,004	1,82
GO:0019219	регуляция метаболизма азотистых соединений	0,008	1,8
GO:0010468	регуляция экспрессии генов	0,010	1,68

Примечание. Приведены процессы с индексом достоверности q -value < 0,05 в порядке убывания степени обогащения. Анализ генных онтологий выявил обогащение биологических процессов, связанных метаболизмом РНК и сплайсингом.

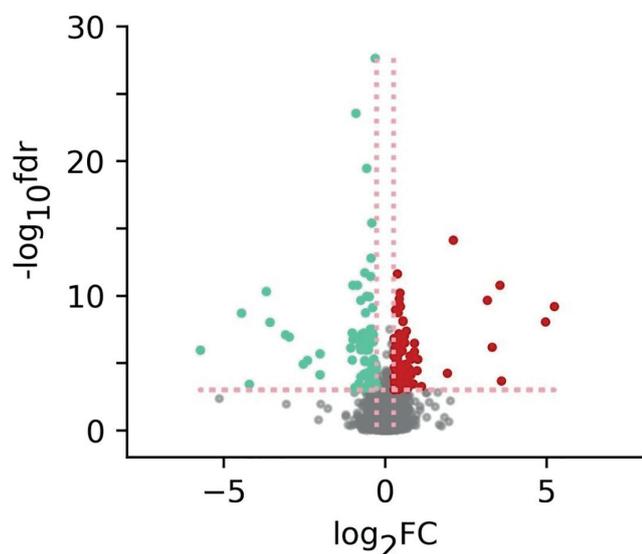


Рис. 2. Диаграмма рассеяния белков, идентифицированных с использованием метода DirectMS1 для клеточной линии A2780, обработанной лонидамином в координатах кратных изменений (FC) измеренных концентраций, относительно контроля и статистической значимости индекса достоверности этих изменений (p -value). Эксперименты проводились на ВЭЖХ-МС-системе Orbitrap Q Exactive HF, адаптированной под метод DirectMS1 на основе ультракоротких (5 мин) ВЭЖХ градиентов

с метаболическими путями гликолиза и синтеза АТФ (выделено серым фоном в табл. 3), ассоциированных ранее с известным механизмом действия лонидамина. Для оптимизации статистического порога функционально значимых изменений концентрации белков протеомов в условиях воздействия лонидамином использовался развитый в проекте подход, основанный на методе Бенджамини–Хохберга, — анализ плотности распределения точек на диаграммах рассеяния и генных онтологий [40]. Анализ генных онтологий для выявленных дифференциально регулируемых белков, статистически значимо понизивших концентрацию после обработки, выявил обогащение биологических процессов, связанных метаболизмом РНК и сплайсингом (табл. 4).

ЗАКЛЮЧЕНИЕ

В целом, исследования показали принципиальную возможность использования ультракороткого полнопротеомного полуквантитативного анализа DirectMS1 в задачах химической протеомики, связанных с поиском

мишеней химиотерапевтического воздействия. В частности, в работе были проанализированы данные химической протеомики, полученные ранее методом глубокой экспрессионной протеомики для большого набора известных химиотерапевтических препаратов, и показано, что DirectMS1 позволяет в рамках 5-минутных полнопротеомных анализов идентифицировать и измерять количественные изменения основных белков-мишеней лекарственного воздействия для большинства из представленных систем «лекарство—протеом». Более того, даже при недостаточной глубине анализа для прямой количественной идентификации мишеней, как в случае с проанализированной в работе клеточной линии рака яичников A2780, обработанной лонидамином, метод позволяет правильно выявлять основные каскады внутриклеточных взаимодействий, активируемых лекарственным воздействием. Полученные результаты позволяют предположить, что новый метод ультракороткого полнопротеомного полуколичественного анализа найдёт широкое применение в задачах, требующих использования больших количеств образцов для получения биологически значимой информации, таких как клиническая и популяционная протеомика, валидация панелей белковых биомаркеров заболеваний и другие.

Вклад авторов. Е.М. Соловьева — проведение экспериментов по методикам лизиса и анализ экспериментальных данных; Ю.А. Бубис,

М.В. Иванов — проведение экспериментов и анализ данных, полученных с использованием метода DirectMS1; И.А. Тарасова — статистический анализ результатов экспрессионной протеомики, полученных для клеточной линии A2780, обработанной лонидамином; А.А. Лобас — анализ данных химической протеомики Каролинского института, полученных для клеточной линии A549, обработанной 56 известными онкопрепаратами; А.А. Назаров, И.А. Шутков — клеточные работы; М.В. Горшков — общее руководство работами и написание статьи.

Благодарности. Авторы благодарят Центр коллективного пользования «Протеом Человека» Института биомедицинской химии им. В.Н. Ореховича и проф. В.Г. Згуду за помощь в реализации метода прямой масс-спектрометрической идентификации белков DirectMS1 для полнопротеомного анализа и наработку экспериментальных данных для данного исследования.

Финансирование. Работа выполнена при финансовой поддержке Российского научного фонда (грант № 20-14-00229).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм. Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

Дополнительные материалы. Приложение к статье опубликовано на сайте журнала «Биохимия» (<https://biochemistrymoscow.com>).

СПИСОК ЛИТЕРАТУРЫ

1. Olivier, M., Asmis, R., Hawkins, G. A., Howard, T. D., and Cox, L. A. (2019) The need for multi-omics biomarker signatures in precision medicine, *Int. J. Mol. Sci.*, **20**, 4781, doi: 10.3390/ijms20194781.
2. Ibrahim, R., Pasic, M., and Yousef, G. M. (2016) Omics for personalized medicine: defining the current we swim in, *Expert Rev. Mol. Diagn.*, **16**, 719-722, doi: 10.1586/14737159.2016.1164601.
3. Kamel, H. F. M., and Al-Amodi, H. S. A. B. (2017) Exploitation of gene expression and cancer biomarkers in paving the path to era of personalized medicine, *Genom. Proteom. Bioinform.*, **15**, 220-235, doi: 10.1016/j.gpb.2016.11.005.
4. Aebersold, R., and Mann, M. (2016) Mass-spectrometric exploration of proteome structure and function, *Nature*, **537**, 347-355, doi: 10.1038/nature19949.
5. Beck, M., Schmidt, A., Malmstroem, J., Claassen, M., Ori, A., et al. (2011) The quantitative proteome of a human cell line, *Mol. Syst. Biol.*, **7**, 549, doi: 10.1038/msb.2011.82.
6. Saei, A. A., Sabatier, P., Tokat, Ü. G., Chernobrovkin, A., Pirmoradian, M., et al. (2018) Comparative proteomics of dying and surviving cancer cells improves the identification of drug targets and sheds light on cell life/death decisions, *Mol. Cell Proteomics*, **17**, 1144-1155, doi: 10.1074/mcp.RA118.000610.
7. Gaetani, M., Sabatier, P., Saei, A. A., Beusch, C. M., Yang, Z., et al. (2019) Proteome integral solubility alteration: a high-throughput proteomics assay for target deconvolution, *J. Proteome Res.*, **18**, 4027-4037, doi: 10.1021/acs.jproteome.9b00500.
8. Savitski, M. M., Reinhard, F. B., Franken, H., Werner, T., Savitski, M. F., et al. (2014) Tracking cancer drugs in living cells by thermal profiling of the proteome, *Science*, **346**, 1255784, doi: 10.1126/science.1255784.
9. Mateus, A., Kurzawa, N., Perrin, J., Bergamini, G., and Savitski, M. M. (2022) Drug target identification

- in tissues by thermal proteome profiling, *Annu. Rev. Pharmacol. Toxicol.*, **62**, 465-482, doi: 10.1146/annurev-pharmtox-052120-013205.
10. Baker, E. S., Liu, T., Petyuk, V. A., Burnum-Johnson, K. E., Ibrahim, Y. M., et al. (2012) Mass spectrometry for translational proteomics: progress and clinical implications, *Genome Med.*, **4**, 63, doi: 10.1186/gm364.
 11. Gillet, L. C., Leitner, A., and Aebersold, R. (2016) Mass spectrometry applied to bottom-up proteomics: entering the high-throughput era for hypothesis testing, *Annu. Rev. Anal. Chem.*, **9**, 449-472, doi: 10.1146/annurev-anchem-071015-041535.
 12. Zhang, L., and Elias, J. E. (2017) Relative protein quantification using tandem mass tag mass spectrometry, *Methods Mol. Biol.*, **1550**, 185-198, doi: 10.1007/978-1-4939-6747-6_14.
 13. Bekker-Jensen, D. B., Kelstrup, C. D., Batth, T. S., Larsen, S. C., Haldrup, C., et al. (2017) An optimized shotgun strategy for the rapid generation of comprehensive human proteomes, *Cell Syst.*, **4**, 587-599.e4, doi: 10.1016/j.cels.2017.05.009.
 14. Meier, F., Geyer, P. E., Virreira Winter, S., Cox, J., and Mann, M. (2018) BoxCar acquisition method enables single-shot proteomics at a depth of 10,000 proteins in 100 minutes, *Nat. Methods*, **15**, 440-448, doi: 10.1038/s41592-018-0003-5.
 15. Bache, N., Geyer, P. E., Bekker-Jensen, D. B., Hoerning, O., Falkenby, L., et al. (2018) Novel LC system embeds analytes in pre-formed gradients for rapid, ultra-robust proteomics, *Mol. Cell. Proteomics*, **17**, 2284-2296, doi: 10.1074/mcp.TIR118.000853.
 16. Meier, F., Brunner, A. D., Koch, S., Koch, H., Lubeck, M., et al. (2018) Online Parallel Accumulation-Serial Fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer, *Mol. Cell. Proteomics*, **17**, 2534-2545, doi: 10.1074/mcp.TIR118.000900.
 17. Ivanov, M. V., Tarasova, I. A., Levitsky, L. I., Solovyeva, E. M., Pridatchenko, M. L., et al. (2017) MS/MS-free protein identification in complex Mixtures using multiple enzymes with complementary specificity, *J. Proteome Res.*, **16**, 3989-3999, doi: 10.1021/acs.jproteome.7b00365.
 18. Ivanov, M. V., Bubis, J. A., Gorshkov, V., Tarasova, I. A., Levitsky, L. I., et al. (2020) DirectMS1: MS/MS-free identification of 1000 proteins of cellular proteomes in 5 minutes, *Anal. Chem.*, **92**, 4326-4333, doi: 10.1021/acs.analchem.9b05095.
 19. Ivanov, M. V., Bubis, J. A., Gorshkov, V., Abdrakhimov, D. A., Kjeldsen, F., et al. (2021) Boosting MS1-only proteomics with machine learning allows 2000 protein identifications in single-shot human proteome analysis using 5 min HPLC gradient, *J. Proteome Res.*, **20**, 1864-1873, doi: 10.1021/acs.jproteome.0c00863.
 20. Tansey, W. P. (2006) Freeze-thaw lysis for extraction of proteins from Mammalian cells, *CSH Protoc.*, **2006**, pdb.prot4614, doi: 10.1101/pdb.prot4614.
 21. Saei, A. A., Beusch, C. M., Chernobrovkin, A., Sabatier, P., Zhang, B., et al. (2019) ProTargetMiner as a proteome signature library of anticancer molecules for functional discovery, *Nat. Commun.*, **10**, 5715, doi: 10.1038/s41467-019-13582-8.
 22. Duan, J.-X. (2005) Method for synthesis of lonidamine and related indazole derivatives, *Patent WO2005120498A2*, publication date 22.12.2005.
 23. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C., and Yates, J. R. 3rd. (2013) Protein analysis by shotgun/bottom-up proteomics, *Chem. Rev.*, **113**, 2343-2394, doi: 10.1021/cr3003533.
 24. Bateman, N. W., Goulding, S. P., Shulman, N. J., Gaddok, A. K., Szumlinski, K. K., et al. (2014) Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA), *Mol. Cell. Proteomics*, **13**, 329-338, doi: 10.1074/mcp.M112.026500.
 25. Levitsky, L. I., Ivanov, M. V., Lobas, A. A., Bubis, J. A., Tarasova, I. A., et al. (2018) IdentiPy: an extensible search engine for protein identification in shotgun proteomics, *J. Proteome Res.*, **17**, 2249-2255, doi: 10.1021/acs.jproteome.7b00640.
 26. Ivanov, M. V., Levitsky, L. I., Bubis, J. A., and Gorshkov, M. V. (2019) Scavenger: a versatile postsearch validation algorithm for shotgun proteomics based on gradient boosting, *Proteomics*, **19**, e1800280, doi: 10.1002/pmic.201800280.
 27. Zybailov, B., Mosley, A. L., Sardi, M. E., Coleman, M. K., Florens, L., et al. (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*, *J. Proteome Res.*, **5**, 2339-2347, doi: 10.1021/pr060161n.
 28. Bubis, J. A., Levitsky, L. I., Ivanov, M. V., Tarasova, I. A., and Gorshkov, M. V. (2017) Comparative evaluation of label-free quantification methods for shotgun proteomics, *Rapid Commun. Mass Spectrom.*, **31**, 606-612, doi: 10.1002/rcm.7829.
 29. Zhang, B., Pirmoradian, M., Zubarev, R., and Käll, L. (2017) Covariation of peptide abundances accurately reflects protein concentration differences, *Mol. Cell. Proteomics*, **16**, 936-948, doi: 10.1074/mcp.O117.067728.
 30. Abdrakhimov, D. A., Bubis, J. A., Gorshkov, V., Kjeldsen, F., Gorshkov, M. V., et al. (2021) Biosaur: An open-source Python software for liquid chromatography-mass spectrometry peptide feature detection with ion mobility support, *Rapid Commun. Mass Spectrom.*, e9045, doi: 10.1002/rcm.9045.
 31. Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists, *BMC Bioinformatics*, **10**, 48, doi: 10.1186/1471-2105-10-48.
 32. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, *Nat. Genet.*, **25**, 25-29, doi: 10.1038/75556.

33. Zhao, Y., Butler, E. B., and Tan, M. (2013) Targeting cellular metabolism to improve cancer therapeutics, *Cell Death Dis.*, **4**, e532, doi: 10.1038/cddis.2013.60.
34. Shang, C., Hou, Y., Meng, T., Shi, M., and Cui, G. (2021) The anticancer activity of indazole compounds: a mini review, *Curr. Top. Med. Chem.*, **21**, 363-376, doi: 10.2174/1568026620999201124154231.
35. Peng, J., Cui, Y., Xu, S., Wu, X., Huang, Y., et al. (2021) Altered glycolysis results in drug-resistant in clinical tumor therapy, *Oncol. Lett.*, **21**, 369, doi: 10.3892/ol.2021.12630.
36. Pelicano, H., Martin, D. S., Xu, R. H., and Huang, P. (2006) Glycolysis inhibition for anticancer treatment, *Oncogene*, **25**, 4633-4646, doi: 10.1038/sj.onc.1209597.
37. De Lena, M., Lorusso, V., Latorre, A., Fanizza, G., Gargano, G., et al. (2001) Paclitaxel, cisplatin and lonidamine in advanced ovarian cancer. A phase II study, *Eur. J. Cancer*, **37**, 364-368, doi: 10.1016/s0959-8049(00)00400-7.
38. Nath, K., Guo, L., Nancolas, B., Nelson, D. S., Shestov, A. A., et al. (2016) Mechanism of anti-neoplastic activity of lonidamine, *Biochim. Biophys. Acta*, **1866**, 151-162, doi: 10.1016/j.bbcan.2016.08.001.
39. Shen, Y. A., Chen, C. C., Chen, B. J., Wu, Y. T., Juan, J. R., et al. (2021) Potential therapies targeting metabolic pathways in cancer stem cells, *Cells*, **10**, 1772, doi: 10.3390/cells10071772.
40. Gabdrakhmanov, I. T., Gorshkov, M. V., and Tarasova, I. A. (2021) Proteomics of cellular response to stress: taking control of false positive results, *Biochemistry (Moscow)*, **86**, 338-349, doi: 10.1134/S0006297921030093.

ON THE FEASIBILITY OF USING AN ULTRA-FAST DirectMS1 METHOD OF PROTEOME-WIDE ANALYSIS FOR SEARCHING DRUG TARGETS IN CHEMICAL PROTEOMICS

E. M. Solovyeva¹, J. A. Bubis¹, I. A. Tarasova¹, A. A. Lobas¹, M. V. Ivanov¹, A. A. Nazarov², I. A. Shutkov², and M. V. Gorshkov^{1*}

¹ V. L. Talrose Institute for Energy Problems of Chemical Physics, N.N. Semenov Federal Research Center for Chemical Physics, Russian Academy of Sciences, 119334 Moscow, Russia; e-mail: mike.gorshkov@gmail.com

² Faculty of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia

Protein quantitation in tissue cells or physiological fluids based on liquid chromatography/mass spectrometry is one of the key sources of information on the mechanisms of cell functioning during chemotherapeutic treatment. Information on significant changes in protein expression upon treatment can be obtained by chemical proteomics and requires analysis of the cellular proteomes, as well as development of experimental and bioinformatic methods for identification of the drug targets. Low throughput of whole proteome analysis based on liquid chromatography and tandem mass spectrometry is one of the main factors limiting the scale of these studies. The method of direct mass spectrometric identification of proteins, DirectMS1, is one of the approaches developed in recent years allowing ultrafast proteome-wide analyses employing minute-scale gradients for separation of proteolytic mixtures. Aim of this work was evaluation of both possibilities and limitations of the method for identification of drug targets at the level of whole proteome and for revealing cellular processes activated by the treatment. Particularly, the available literature data on chemical proteomics obtained earlier for a large set of onco-pharmaceuticals using multiplex quantitative proteome profiling were analyzed. The results obtained were further compared with the proteome-wide data acquired by the DirectMS1 method using ultrashort separation gradients to evaluate efficiency of the method in identifying known drug targets. Using ovarian cancer cell line A2780 as an example, a whole-proteome comparison of two cell lysis techniques was performed, including the freeze-thaw lysis commonly employed in chemical proteomics and the one based on ultrasonication for cell disruption, which is the widely accepted as a standard in proteomic studies. Also, the proteome-wide profiling was performed using ultrafast DirectMS1 method for A2780 cell line treated with lonidamine, followed by gene ontology analyses to evaluate capabilities of the method in revealing regulation of proteins in the cellular processes associated with drug treatment.

Keywords: proteomics, mass spectrometry, proteins, lonidamine, lysis, signaling pathways