

ОБЗОРЫ

©Коллектив авторов

ПРОТЕОМИКА АНТИТЕЛ

Л.Ш. Казиева, Т.Е. Фарафонова, В.Г. Згода*

Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича,
119121, Москва, Погодинская ул., 10; *эл. почта: victor.zgoda@gmail.com

Антитела являются важным компонентом гуморального иммунитета, в связи с чем их изучение имеет колоссальное значение для молекулярной биологии и медицины. Уникальная способность антител специфично распознавать и связываться с молекулярной мишенью — антигеном — обуславливает их широкое применение в терапии и диагностике заболеваний, а также в лабораторной и биотехнологической практике. Высокая специфичность и аффинность антител обеспечивается наличием в их структуре переменных участков, которые не закодированы в геноме человека, а уникальны для каждого конкретного клона В-лимфоцитов — продуцентов антител. Как следствие, информация о последовательностях переменных участков антител в базах данных почти отсутствует. Это отличает изучение первичной структуры антител от других протеомных исследований, поскольку подразумевает расшифровку генома В-лимфоцитов либо *de novo* секвенирование аминокислотной последовательности антител. В данном обзоре на примере ряда протеомных и протео-геномных исследований рассмотрены существующие подходы, которые протеомика может предложить для изучения антител, в частности, для расшифровки их первичной структуры, оценки посттрансляционных модификаций и оптимизации биоинформатических инструментов с учётом специфики анализа антител.

Ключевые слова: антитела; масс-спектрометрия; секвенирование антител *de novo*; репертуар антител

DOI: 10.18097/PBMC20236901005

ВВЕДЕНИЕ

Антитела являются важным компонентом гуморального иммунитета, в связи с чем их изучение имеет колоссальное значение для молекулярной биологии и медицины. Синтез антител к различным экзо- и эндогенным антигенам указывает на наличие иммунного ответа, что является индикатором того или иного патологического процесса в организме. Это позволяет применять антитела в качестве биомаркеров для диагностики различных инфекционных, аутоиммунных и онкологических заболеваний [1-3]. Кроме диагностики антитела применяются в качестве молекулярных инструментов для идентификации, выделения/очистки и меченых белков (иммуноферментный анализ, радиоиммунный анализ, Вестерн-блоттинг, аффинная хроматография, цитофлуориметрия) [4-7].

Уникальная способность антител специфично распознавать и связываться с молекулярной мишенью — антигеном — и в дальнейшем элиминировать его обеспечивает возможность создания препаратов для таргетной терапии различных заболеваний.

Понимание молекулярных механизмов взаимосвязи специфичности антител с их строением позволит открыть новые возможности в лечении и диагностике заболеваний. Функцию антител во многом определяет аминокислотная последовательность их белковой молекулы. Одним из методов определения аминокислотной последовательности белков является метод масс-спектрометрического анализа с последующей интерпретацией масс-спектров по базам данных, содержащим сиквенс белков. Это возможно благодаря тому, что геномы многих

организмов, включая человека, на сегодняшний день расшифрованы.

Особенностью строения антител является наличие в их структуре переменных участков, аминокислотная последовательность которых уникальна для каждого конкретного клона В-лимфоцита [8]. Из чего следует, что стандартный подход интерпретации масс-спектров по существующим базам данных не применим. Это обуславливает необходимость секвенировать переменную часть антитела с помощью анализа генома активированных В-лимфоцитов, либо проводить аминокислотное секвенирование с использованием масс-спектрометрических методов.

Целью данного обзора было рассмотрение возможностей протеомики и разнообразия существующих на данный момент методических подходов для изучения структуры антител и получения их полной аминокислотной последовательности, оценка перспектив применения антител в терапии и диагностике, а также обозначение существующих проблем в области протеомных исследований антител и возможных путей их решения.

1. СТРУКТУРА И ФУНКЦИИ АНТИТЕЛ

В организме большинства млекопитающих продуцируются антитела пяти классов — А, М, G, Е и D, из которых иммуноглобулины класса G (IgG) наиболее представлены в сыворотке. Молекула IgG состоит из четырёх полипептидных цепей: двух тяжёлых цепей (HC) и двух лёгких (LC). Каждая лёгкая цепь состоит из переменного (VL) и константного (CL) доменов, а тяжёлая цепь — из одного переменного (VH) и трёх константных (CH1-3) доменов [9, 10].

В структуре антитела выделяют так называемые Fc- (Fragment crystallizable) и Fab- (Fragment antigen-binding) фрагменты (рис. 1). Fc-фрагмент состоит из CH-доменов и обеспечивает эффекторные свойства антител, способность же взаимодействовать с различными антигенами обеспечивается Fab-фрагментом, а именно его антиген-связывающими участками (CDR1-3, complementarity determining regions 1-3). Благодаря разнообразию аминокислотных последовательностей в CDRs, антитела потенциально способны защитить организм от внешних воздействий, а также внутренних сбоев в геноме. Количество возможных последовательностей переменных доменов огромно и, по некоторым оценкам, достигает от 10^8 до 10^{13} вариантов [11]. Наибольшая вариативность наблюдается в участке CDR3 тяжёлой цепи (CDRH3). Предполагается, что именно CDRH3 вносит основной вклад во взаимодействие с антигеном [12-14]. По этой причине внимание многих исследователей приковано к CDRH3.

В целом представление об аминокислотной последовательности антиген-связывающего участка антитела является ключевым шагом в понимании его функции и терапевтического потенциала.

2. ЗНАЧЕНИЕ АНТИТЕЛ В БИОЛОГИИ И МЕДИЦИНЕ

С момента открытия гибридных технологий антитела прочно вошли в медицинскую и научную практику в качестве терапевтических, диагностических и лабораторных инструментов. Начиная с регистрации первого терапевтического моноклонального антитела муромонаб (от англ. murine monoclonal antibody — мышьеиное моноклональное антитело) в 1986 году [15], препараты данной группы стали одним

из самых быстроразвивающихся направлений в фармацевтической промышленности и широко применяются в лечении заболеваний различных нозологических форм [16]. На сегодняшний день в терапии и визуализации патологических процессов применяются как моноклональные антитела, так и различные терапевтические молекулы на их основе: антитела, конъюгированные с активными частицами — токсинами, радиоизотопами или цитокинами [17], видоизменённые антитела и их фрагменты — аффитела [18, 19], нанотела [20], биспецифические антитела [21, 22]. Препараты на основе моноклональных антител зарекомендовали себя в терапии онкологических и аутоиммунных заболеваний [23-25].

Поскольку антитела являются компонентом гуморального иммунитета, их содержание, а также изменения уровня их посттрансляционных модификаций позволяют отследить протекающие в организме реакции иммунной системы на внешнюю или внутреннюю угрозу. Это обуславливает закрепившееся в клинической практике использование антител к экзо- и эндогенным антигенам в качестве биомаркеров различных патологических процессов злокачественной [26-28], аутоиммунной [29-32] и инфекционной природы [33-36].

3. СЕКВЕНИРОВАНИЕ АНТИТЕЛ

Наиболее значимой для протеомики особенностью антител является то, что антигенсвязывающая часть не закодирована в геноме организма, а определяется в ходе рекомбинации определённых участков ДНК в каждой индивидуальной В-клетке и окончательно формируется в результате соматических мутаций при созревании активированных В-лимфоцитов [8].

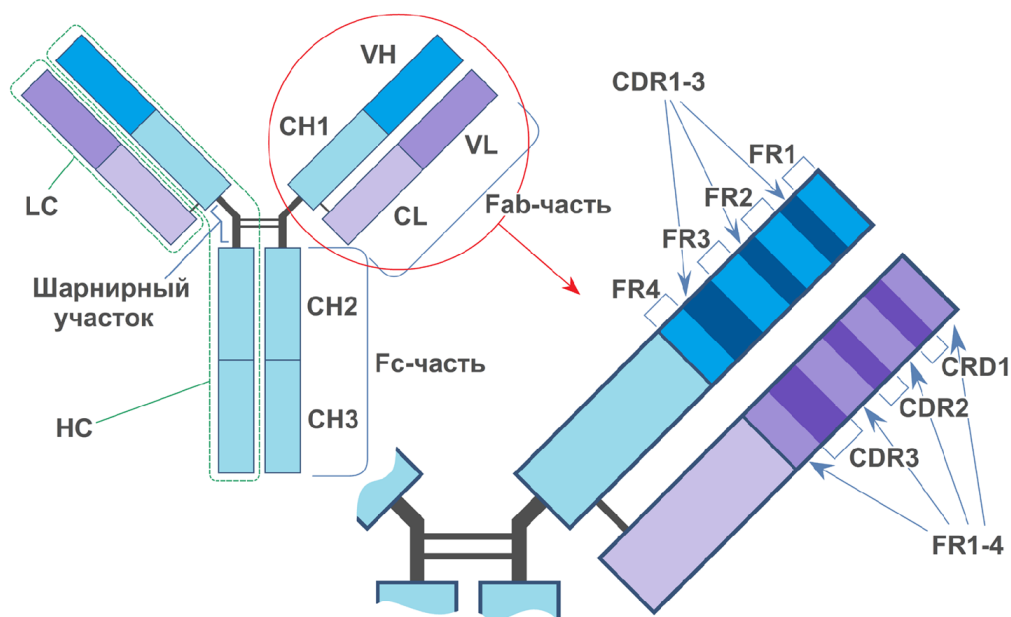


Рисунок 1. Строение иммуноглобулина G: слева показана доменная структура антитела, LC – лёгкая цепь, HC – тяжёлая цепь. Лёгкая цепь состоит из вариабельного и константного доменов VL и CL, тяжёлая – из одного вариабельного домена VH и трёх константных доменов CH1-3. Справа на увеличенной схеме Fab-части изображено подробное строение вариабельных доменов антитела, имеющих в составе четыре каркасных участка FR1-4 и три гипервариативных участка CDR1-3.

Ввиду отсутствия искоемых последовательностей в геномных и протеомных базах данных, для исследования первичной структуры антител необходимо либо получить данные о генах, которые их кодируют, либо секвенировать молекулы антител *de novo* (рис. 2).

Таким образом, определение переменных последовательностей антител на сегодняшний день возможно двумя способами:

- секвенированием генома циркулирующих в крови активированных В-лимфоцитов либо генов В-клеточного рецептора (BCR) и создание базы данных последовательностей антител конкретного организма для дальнейших протеомных исследований;
- *de novo* секвенированием аминокислотной последовательности антитела хромато-масс-спектрометрическими методами (LC-MS и LC-MS/MS — хромато-масс-спектрометрия и тандемная хромато-масс-спектрометрия).

В арсенале современной протеомики можно выделить три основных хромато-масс-спектрометрических подхода анализа первичной структуры белков — top-down, middle-down и bottom-up.

Подход top-down представляет собой масс-спектрометрический анализ целого белка (рис. 3). Преимуществами данного подхода являются возможность измерить молекулярную массу интактного белка вместе с посттрансляционными

модификациями (ПТМ) и возможным процессингом С- и N-концевых аминокислот, а также возможность определить наличие протеоформ белка [37]. Используя различные режимы фрагментации белковой молекулы в ходе тандемного масс-спектрометрического анализа (MS/MS), можно провести секвенирование аминокислотной последовательности целого антитела. Одним из основных ограничений top-down подхода является низкая производительность анализа, так как метод мало пригоден для работы со сложными белковыми смесями и требует значительных количеств очищенного препарата антитела. При этом получение тандемных спектров фрагментации, которые позволят расшифровать первичную структуру антитела, почти всегда требует индивидуального подхода.

Подход bottom-up заключается в предварительном гидролизе белка протеолитическими ферментами и дальнейшем хромато-масс-спектрометрическом (LC-MS/MS) анализе полученной смеси пептидов. Пептиды, по сравнению с целым белком, обладают более высокой растворимостью в полярных растворителях, лучше разделяются на фракции и ионизируются, что значительно упрощает их анализ (рис. 4). Для bottom-up подхода разработано множество биоинформатических инструментов, позволяющих идентифицировать белки в образце по белковым базам данных, а также составлять на основе MS/MS-данных аминокислотную последовательность *de novo* [38-45].

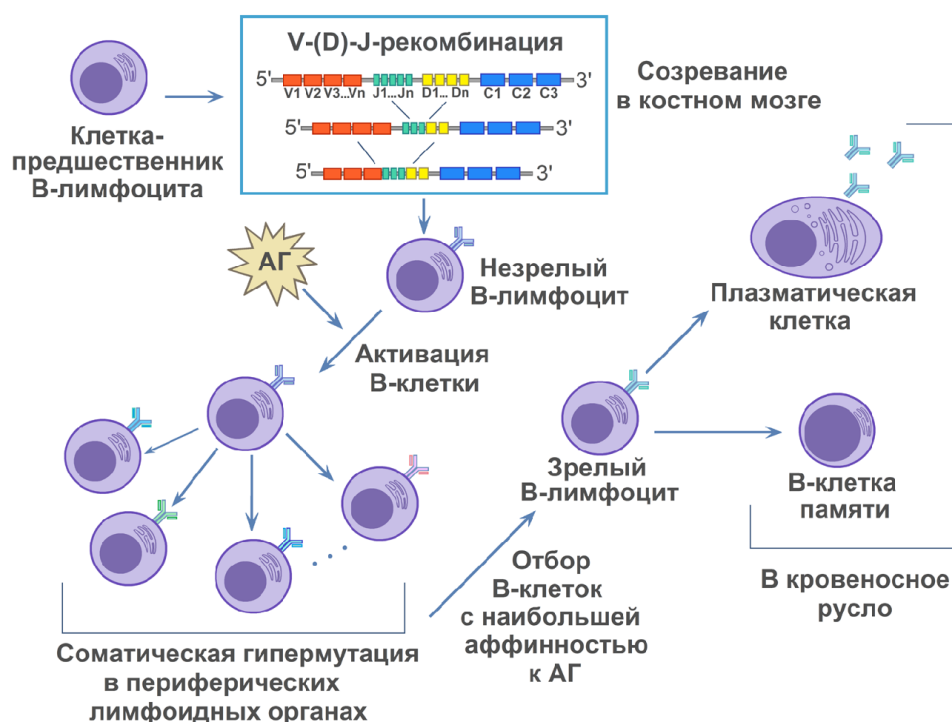


Рисунок 2. Механизм обеспечения разнообразия антител. В верхней части — совокупность предшественников В-лимфоцитов созревает в костном мозге, проходя процесс V-(D)-J-рекомбинации кластеров генов тяжелой и легкой цепей антител. Получившиеся в результате этого незрелые В-клетки активируются при встрече с антигеном (АГ) и мигрируют в периферические органы лимфатической системы для дальнейшего созревания. В нижней части рисунка — соматическая гипермутация в клонах активированного незрелого В-лимфоцита позволяет достичь огромного разнообразия возможных вариантов антител, из которых происходит отбор клона с наибольшей аффинностью к АГ. Этот клон В-клетки является зрелым В-лимфоцитом, который далее дифференцируется либо в плазматические клетки, активно производящие антитела, либо в циркулирующие в крови В-клетки памяти.

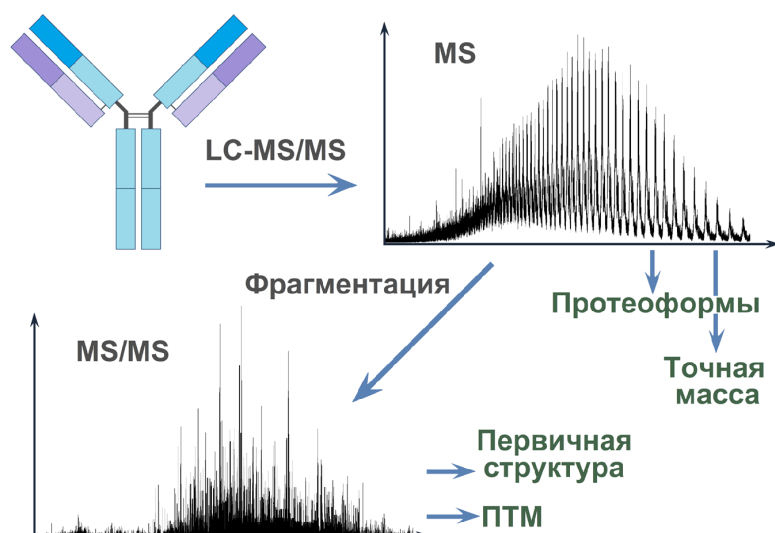


Рисунок 3. Схема подхода top-down. В верхней части рисунка – масс-спектр целого антитела, полученный с помощью масс-спектрометра высокого разрешения. В нижней части – фрагментация антитела с получением тандемного масс-спектра для определения первичной последовательности и расположения посттрансляционных модификаций (ПТМ).

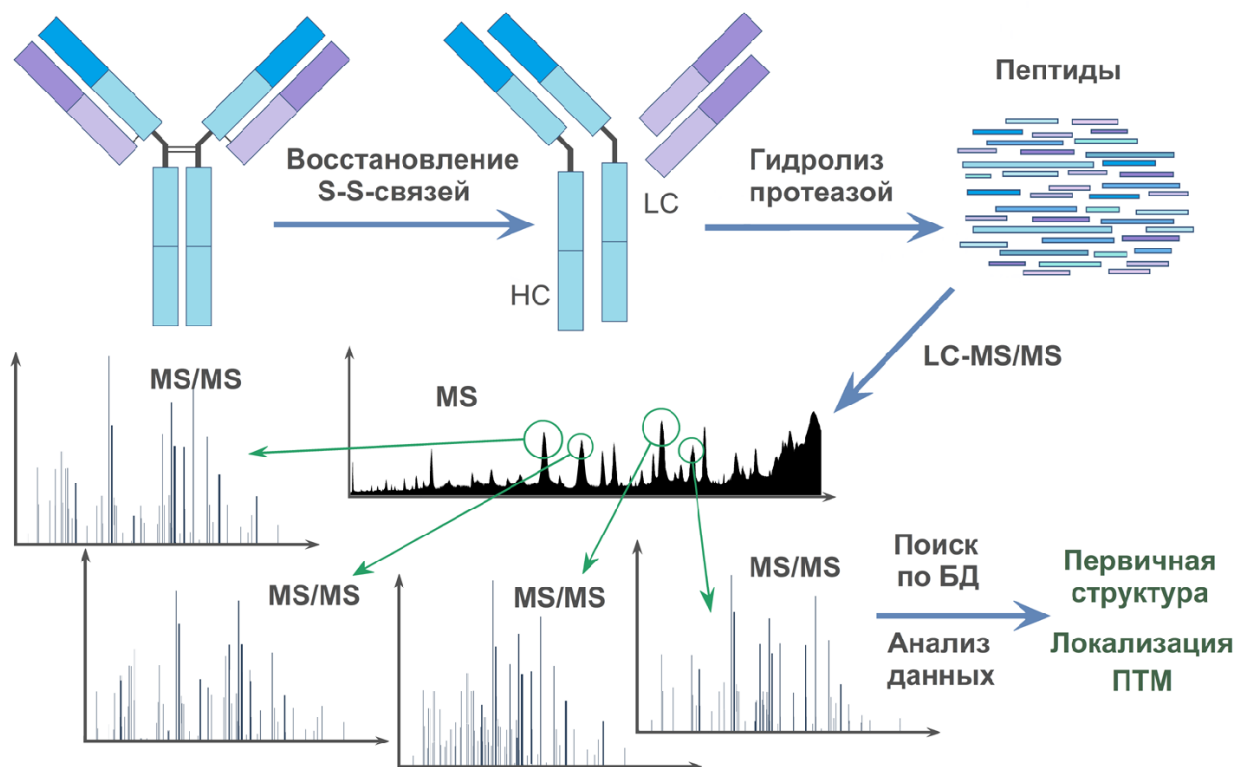


Рисунок 4. Схема подхода bottom-up. В верхней части рисунка – восстановление антитела для разделения его на лёгкие (LC) и тяжёлые (HC) цепи и последующего гидролиза протеазой. Внизу – тандемный хромато-масс-спектрометрический анализ полученных пептидов, хроматограмма пептидов и тандемные масс-спектры (MS/MS) их фрагментов. Биоинформатический анализ полученных данных позволяет определить первичную структуру исследуемого белка и локализацию ПТМ.

Одним из основных недостатков bottom-up является то, что некоторые из полученных в ходе гидролиза пептидов могут выйти за диапазон детектируемых значений соотношения массы к заряду (m/z), приводя к потере информации о части аминокислотной последовательности и ПТМ. В случае секвенирования антител это является серьёзной проблемой, поскольку получение

информации о последовательностях CDRs критически значимо. К счастью, данный недостаток нивелируется использованием комбинаций различных протеаз (трипсин, химо tripsин, LysN, LysC, GluC, AspN, термолизин, эластаза), так как за счёт их специфичности к различным сайтам расщепления можно получить пептиды с перекрывающимися участками последовательностей [46, 47].

Подход middle-down является производным top-down и представляет собой промежуточную версию между top-down и bottom-up (рис. 5). Он заключается в том, что перед масс-спектрометрическим анализом белок подвергается частичному расщеплению, например, такому как восстановление S-S связей с последующим разделением антитела на лёгкие и тяжёлые цепи. Для получения фрагментов белка также применяют гидролиз протеазами, специфичными к редким аминокислотам (GluC, LysN, LysC, AspN) либо к определённому участку белка (папаин, IdeS), или же неэнзиматический гидролиз в кислой среде [48]. Подход middle-down находит своё применение при изучении различных белков, в том числе и антител. Благодаря тому, что анализируются достаточно крупные фрагменты антитела, middle-down позволяет не только повысить степень их ионизации по сравнению с целой молекулой, но и снизить потерю данных о ПТМ, определить порядок расположения гипервариабельных участков, а также даёт возможность сопоставить лёгкие и тяжёлые цепи в смесях антител [49].

Одной из трудностей в секвенировании белков методом LC-MS/MS является то, что фрагментация родительских ионов пептидов не позволяет различить пространственные изомеры с одинаковой молекулярной массой — лейцин и изолейцин. В случае *de novo* секвенирования вариабельных участков антител невозможность достоверного определения всех аминокислот в последовательности представляет собой проблему в дальнейшем синтезе моноклональных антител и может привести к снижению их аффинности. Данная задача может быть решена комбинированием нескольких методов фрагментации ионов — многоэтапной тандемной масс-спектрометрии

со столкновительной фрагментацией с высокой энергией (HCD-MSⁿ) и многоэтапной MS/MS (MSⁿ) с фрагментацией с переносом электрона и анализа столкновительной фрагментацией с высокой энергией (ETD-HCD MS³) [50].

3.1. Полное секвенирование лёгкой и тяжёлой цепей

Полное секвенирование лёгкой и тяжёлой цепей антитела необходимо при изучении модификаций концевых последовательностей и ПТМ антитела, таких как гликозилирование, окисление метионина, дезамидирование глутамата и аспартата, в особенности при проведении контроля качества терапевтических моноклональных антител (МАт).

Для решения данной задачи широко применяется масс-спектрометрический подход top-down. Мао и соавт. [51] проводили top-down секвенирование нативного терапевтического МАт методом ECD-MS/MS (тандемная MS с диссоциацией с захватом электрона) на собранной исследователями FTICR-MS (МС ионного циклотронного резонанса с преобразованием Фурье) установке. Авторы исследования смогли продемонстрировать данные по 25% покрытию лёгкой и тяжёлой цепей при изоляции прекурсора одного зарядового состояния, сайтов фрагментации наблюдалось больше в участках без S-S-линкеров. Для MS/MS всех зарядовых состояний без изоляции прекурсора покрытие составило 32% лёгкой и 35% тяжёлой цепей соответственно.

Fornelli и соавт. [52] продемонстрировали применение подхода middle-down для анализа первичной последовательности IgG на примере трёх терапевтических МАт. Для проведения анализа антитела гидролизовали ферментом IdeS

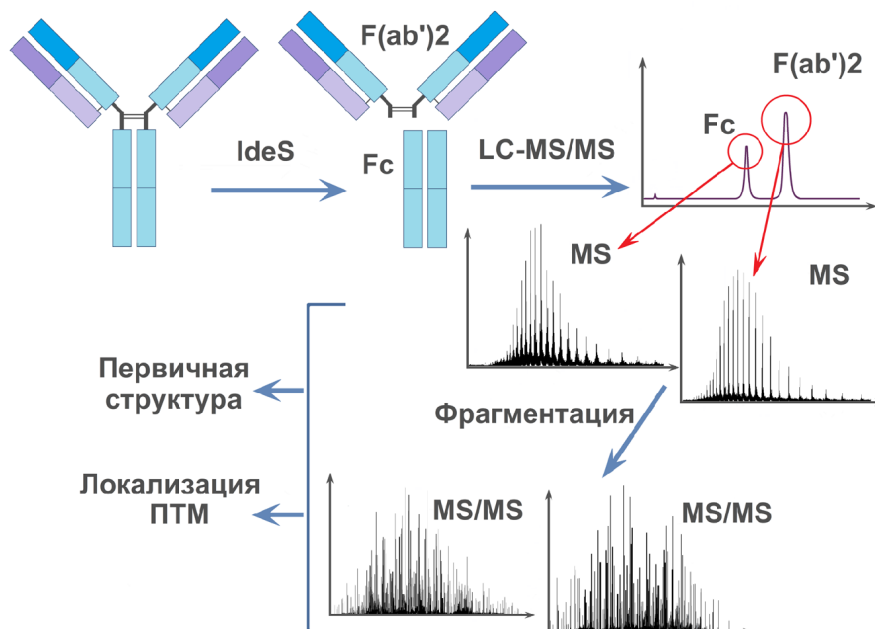


Рисунок 5. Схема подхода middle-down на примере анализа антитела с использованием фермента IdeS. В верхней части рисунка – гидролиз антитела IdeS, расщепляющим тяжёлую цепь на границе между шарнирным участком и CH₂-доменом. Полученные F(ab')₂- и Fc-фрагменты анализируют хромато-масс-спектрометрически. В нижней части рисунка – масс-спектры F(ab')₂- и Fc-частей антитела (MS) и тандемные масс-спектры их фрагментации (MS/MS). Дальнейший анализ масс-спектрометрических данных позволяет определить первичную структуру и локализацию ПТМ.

(IgG-degrading enzyme of *Streptococcus pyogenes*), после восстановления и алкилирования фрагменты массой ~25 кДа анализировали хромато-масс-спектрометрически. С помощью LC-MS анализа были определены точные массы фрагментов антител после восстановления и алкилирования. Последующая расшифровка MS/MS-данных позволила определить аминокислотную последовательность антитела, в том числе гипервариабельных участков, а также подтвердить такие ПТМ, как сайты гликозилирования в Fc-фрагменте и отщепление C-концевого лизина тяжёлой цепи. Одно из антител подвергалось химическому окислению в мягких условиях до гидролиза IdeS с целью моделирования поиска окислительных ПТМ с помощью middle-down методики. Исследователи смогли определить сайты окисления метионина в тяжёлой цепи. Также был проведён LC-MS анализ смеси трёх антител, гидролизованных IdeS. Несмотря на хроматографическую коэлюцию нескольких различных фрагментов от разных антител, их масс-спектры позволяли установить массы ионов-прекурсоров (родительских ионов), которые могут быть использованы для дальнейшего таргетного MS/MS-анализа и установления первичной структуры каждого антитела в смеси.

В другой работе Fornelli и соавт. [49] на примере изучения антитела ритуксимаб показали возможности различных методов фрагментации для top-down и middle-down анализа антител. Моноклональное антитело было проанализировано как в нативном, так и в гидролизованном виде. Для гидролиза применялись два фермента — гингипаин К и IdeS,

которые гидролизуют антитело на фрагменты массой 50 кДа в случае гингипаина К и 25 кДа после восстановления и алкилирования в случае IdeS соответственно. Применив в отдельных экспериментах фрагментацию белковых ионов методами диссоциации с переносом электрона высокой энергии (ETD-HD), фотодиссоциации ультрафиолетом (UVPD) и диссоциации с переносом электронов/столкновительной диссоциацией с повышенной энергией (EThcD), исследователи получили покрытие 40% сиквенса нативного антитела, 20% при анализе фрагментов массой 50 кДа (гингипаин К), более 70% для Fd-фрагмента и до 90% для LC и Fc/2 (восстановленный Fc-фрагмент) (IdeS).

В ряде исследований top-down подход нашёл своё применение в комбинации с хроматографией ультравысокого давления (1D UPLC-TDMS) для разделения и анализа Fab-фрагментов антител сыворотки крови [53, 54].

На начальном этапе bottom-up протеомные исследования антител, как правило, включают создание референтной геномной базы данных. Однако сейчас ведутся разработки методов, основанных на МС без опоры на базы данных геномов В-лимфоцитов. Так, Cheng и соавт. [55] продемонстрировали на примере терапевтического МАт методику секвенирования, основанную на капиллярном электрофорезе (CE) триптических пептидов с последующей тандемной масс-спектрометрией (рис. 6). Исследователям удалось добиться 100% покрытия лёгкой цепи и 96% покрытия тяжёлой. Также они предложили использование

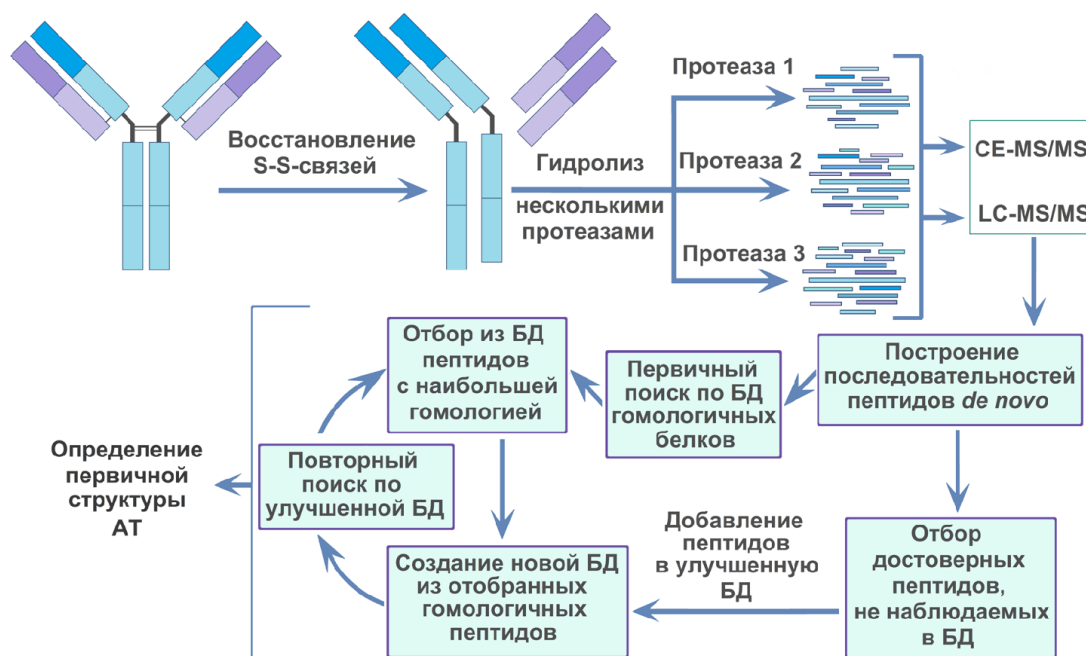


Рисунок 6. Схема исследования Cheng и соавт. [55]. В верхней части рисунка показаны пробоподготовка и получение экспериментальных данных методами капиллярного электрофореза – тандемной масс-спектрометрии (CE-MS/MS) и тандемной хромато-масс-спектрометрии (LC-MS/MS). В нижней части отображён биоинформатический анализ полученных данных: после построения последовательностей *de novo* и первичного поиска по базе данных (БД) гомологичных белков составляли новую улучшенную базу данных с добавленными в неё пептидами, полученными *de novo* с высокой достоверностью и не встречающимися в исходной БД. Путём повторения такого цикла оптимизации БД составляется аминокислотная последовательность исследуемого антитела (АТ).

комбинации методов LC-MS/MS и CE-MS/MS. Такой смешанный подход был продемонстрирован на примере секвенирования нового терапевтического анти-CD176-антитела и позволил авторам первыми предложить аминокислотную последовательность для этого антитела [55].

3.2. Секвенирование антигенсвязывающих участков

В силу того, что константная часть антител видоспецифична, зачастую исследователей интересует именно секвенирование варибельной части и CDRs. Секвенирование гиперварибельных участков антител также используется для изучения клонального состава (репертуара) антител при различных иммунных процессах.

Например, Wheatley и соавт. [56] провели количественную оценку антител, специфичных к S-белку вируса SARS-CoV-2, и клеточных иммунных реакций в серии образцов крови, взятых у пациентов в течение первых четырёх месяцев после выздоровления от COVID-19. Объединение геномного секвенирования рецепторов В-клеток и протеомных методов позволили идентифицировать и количественно определить циркулирующие антитела к S1-субдомену S-белка у выздоровевших от COVID-19 пациентов. В данной работе S1-специфический IgG был выделен из плазмы крови выздоравливающих пациентов с COVID-19 с помощью магнитных частиц с нанесённым антигеном (АГ), а тяжёлые цепи были подвергнуты LC-MS/MS. Для определения клонотипов на основе аминокислотной последовательности CDRH3 пептидные спектры сравнивали с геномными последовательностями BCR, полученными из единичных В-клеток памяти от тех же людей. Затем для относительного количественного параллельного мониторинга реакции (PRM) клонотип-специфические пептиды использовали в качестве штрих-кодов и идентифицированные пептиды отслеживали в образцах плазмы в течение четырёх месяцев. Количественное определение пептидов, проведённое на основе хроматографических кривых, показало более быстрый спад нейтрализующего ответа в течение первых 70 дней после заражения [56].

Lavinder и соавт. [57] проанализировали репертуар антител после ревакцинации вакциной от столбняка. Исследователи изолировали антитела из сыворотки крови добровольцев после иммунизации вакциной и гидролизировали их на F(ab')₂- и Fc-части, аффинно выделив АГ-специфичные F(ab')₂-части. С помощью bottom-up подхода были получены масс-спектрометрические данные пептидов, которые были идентифицированы по индивидуальным базам данных VH и VL генов В-клеток пациентов. Было установлено, что репертуар антител к столбнячному анатоксину после ревакцинации состоит из приблизительно 100 клонотипов антител, 3 из которых составляют 40% иммунного ответа. Для проверки корректности секвенирования были синтезированы 13 рекомбинантных антител, 5 из которых по результатам оценки аффинности к АГ показали связывание с эпитопом столбнячного

токсина, отвечающего за проникновение в клетку. Также было выявлено, что из разнообразия клонотипов плазматических клеток на пике реакции на вакцину только <5% детектируются на девятый месяц после вакцинации. Это и другие подобные исследования помогают глубже взглянуть на механизмы иммунного ответа при вакцинации, а также оценить эффективность вакцин и необходимость их доработки.

Guthals и соавт. [58] разработали методику секвенирования антител к цитомегаловирусу (ЦМВ) с использованием bottom-up протеомики и LC-MS нативных антител. Поликлональные антитела к ЦМВ были аффинно выделены из плазмы крови инфицированного пациента, подвергнуты восстановлению и алкилированию. Полученные лёгкие и тяжёлые цепи антител были разделены с помощью SDS-PAGE и затем гидролизованы в геле различными протеазами. Гидролизаты лёгких и тяжёлых цепей анализировали методом LC-MS/MS. Также исследователи измерили точные массы анти-ЦМВ антител, предварительно проведя их дегликозилирование. Совокупность полученных масс-спектрометрических данных была использована для построения аминокислотной последовательности антител *de novo*. На основе полученных данных были синтезированы рекомбинантные анти-ЦМВ антитела, а их аффинность подтверждена методом ИФА с иммобилизованным АГ. Исследователи установили, что эти анти-ЦМВ антитела, составившие 1-2% от общего количества IgG, не были обнаружены при параллельном секвенировании ДНК В-клеток того же пациента.

Ещё одним примером применения комбинированного протео-геномного подхода для изучения репертуаров антител и их динамических изменений служит исследование Lee и соавт. [59], посвященное изменениям репертуара IgG к вирусу гриппа. Объектом исследования были образцы крови, многократно взятые у пациента в течение нескольких лет как в стадиях активного иммунного ответа после вакцинации или инфицирования, так и в период выздоровления. Разнообразие антител определяли по последовательностям гиперварибельного участка тяжёлой цепи CDRH3. Для оценки репертуара антитела выделяли из крови пациента аффинной хроматографией с АГ (гемагглютинин), гидролизировали трипсином и анализировали хромато-масс-спектрометрически. Для идентификации пептидов была использована индивидуальная база данных VH-генов периферических В-клеток пациента. Изучив клональный состав антител и динамику его изменений на протяжении пяти лет, исследователи выявили всего 210 уникальных клонотипов. Состав поликлональной смеси анти-гемагглютининовых антител в целом сохранялся в практически неизменном виде с превалированием небольшого числа клонотипов, устойчиво сохраняющихся в течение исследуемого периода времени и обладающих значительной аффинностью к высококонсервативным участкам гемагглютинина вируса гриппа.

Bondt и соавт. [60] изучили Fab-фрагменты секреторного IgA1 из грудного молока, полученного от двух здоровых доноров в течение 16 недель. Хромато-масс-спектрометрическое исследование показало, что репертуар антител у каждого донора не меняется, по крайней мере, на протяжении 16 недель; при этом перекрытие клональных репертуаров у двух отдельных доноров практически отсутствует. Однако ограничение middle-down подхода, по мнению авторов, заключается в том, что реальная структурная и функциональная уникальность каждого клона антител не определяется.

3.3. Сопоставление лёгких и тяжёлых цепей антитела

При изучении поликлональных антител перед исследователями возникает ряд трудностей, концептуально отличающихся от исследования моноклонального антитела. Поскольку при секвенировании генома используется поликлональная фракция В-клеток, одной из таких трудностей является корректное сопоставление лёгких и тяжёлых цепей антитела. Решить данную проблему позволяет комбинирование геномных и протеомных методов — секвенирование генов В-клеток для создания базы данных BCR-генов, опираясь на которую, можно более достоверно секвенировать антитела с помощью подхода bottom-up.

Одними из первых, кто предложил совместное применение масс-спектрометрии и геномных технологий для секвенирования антител, были исследователи группы Georgiou. Они применили созданную методику, представляющую собой комбинацию технологий секвенирования генома В-клеточных рецепторов и bottom-up хромато-масс-спектрометрии, для эффективной оценки репертуаров антител в сыворотке крови после иммунного ответа [61]. Для получения поликлональной смеси антител был смоделирован иммунный ответ *in vivo* путём иммунизации кроликов гемоцианином моллюска *Concholepas concholepas*. Выделенные из костного мозга и крови иммунизированных кроликов АГ-специфичные В-клетки были отсекарованы, полученные последовательности генов варибельной части антител собраны в VH- и VL-базы данных. Из крови посредством аффинной хроматографии также была выделена совокупность всех IgG, которые гидролизвали пепсином, и из полученных F(ab')₂-фрагментов с помощью аффинной хроматографии с АГ была выделена АГ-специфичная фракция. Далее отобранные F(ab')₂-фрагменты гидролизвали трипсином и проанализировали гидролизат методом LC-MS/MS. Идентификацию пептидов проводили по собранной исследователями базе данных VH- и VL-генов с добавлением всех белок-кодирующих последовательностей генома кролика (база данных OryCun2). При оценке клонального состава информативными считали те уникальные пептиды, которые соответствовали только одному CDRH3-участку. Сопоставив геномные и протеомные данные, исследователи выявили группу CDRH3, для которых соответствующие пептиды

встречались на порядок чаще в АГ-специфичной фракции F(ab')₂-фрагментов по сравнению с неспецифичной фракцией. Для дальнейшего сопоставления пар лёгких и тяжёлых цепей были использованы семь VH-генов, содержащих наиболее представленные CDRH3-последовательности. Для каждого из выбранных VH-генов составляли библиотеку из комбинаций со всеми последовательностями лёгких цепей, обнаруженными в геномах АГ-специфичных В-клеток. Полученные однопечечные VH-VL-сшивки (scFv, single-chain Fragment variable) экспрессировали с помощью фагового дисплея и посредством инкубации фаговых библиотек с иммобилизованным АГ (метод “биопэннинга”) отобрали scFv с наибольшей аффинностью к антигену. Далее были отсекарованы бактериофаги, содержащие гены сшивок с высокой аффинностью к АГ. На основе полученных пар VH- и VL- генов были синтезированы рекомбинантные антитела. Для проверки правильности пар аминокислотных последовательностей лёгкой и тяжёлой цепей была определена аффинность рекомбинантных антител методом конкурентного ИФА к АГ.

Использование подхода middle-down служит ещё одним методом решения данной задачи. Благодаря возможности анализировать крупные фрагменты белка, этот подход позволяет исследовать не восстановленные F(ab')₂-фрагменты антител [48], за счёт чего в образце не происходит разобщения лёгкой цепи и фрагмента тяжёлой цепи, содержащего VH домен. Как показывают данные работ [49, 62, 63], применение middle-down подхода, в особенности в комбинациях с другими хромато-масс-спектрометрическими подходами и с несколькими методами фрагментации ионов, позволяет достоверно получить информацию о парах лёгкая-тяжёлая цепи в структуре антител.

3.4. Особенности обработки данных для секвенирования антител

Для эффективного секвенирования *de novo* моно- и поликлональных антител большое значение имеют не только экспериментальные подходы к секвенированию и масс-спектрометрии, но и эффективные биоинформатические инструменты для обработки данных, учитывающие специфику строения антител. Оптимизация анализа протеомных и протеогеномных данных при изучении антител ставит перед исследователями ряд задач:

- уменьшение числа ложноположительных идентификаций пептидов в ходе анализа масс-спектров пептидов антител;
- оптимизация и автоматизация процесса обработки масс-спектрометрических данных пептидов антител в ходе секвенирования;
- создание методик *de novo* секвенирования антител, не требующих получения базы данных генов В-клеточных рецепторов, а опирающихся на биоинформатическую реконструкцию процесса V-D-J-рекомбинации;

• вариабельная часть AT, в частности CDRs, менее протяжённая по сравнению константной частью, следовательно при гидролизе AT пептиды вариабельной части будут представлены меньше, чем пептиды константной части. Это означает, что необходимы методы, позволяющие максимизировать получение информации обо всех пептидах, присутствующих в образце.

Примером решения одной из таких задач служит исследование Boutz и соавт. [64], в котором были секвенированы VH- и VL-гены В-клеток иммунизированных кроликов (АГ — гемоцианин моллюска *Concholepas concholepas*) и человека (АГ — столбнячный/дифтерийный анатоксин). Антитела, выделенные из сыворотки аффинной хроматографией с протеином А (протеином G для человеческих АТ), гидролизовали пепсином, а полученные F(ab')₂-фрагменты подвергали трипсинолизу с последующим анализом пептидов методом LC-MS/MS. Основной задачей работы была разработка способа снижения числа ложноположительных identifications пептидов CDRH3. Наличие в структуре Fab-фрагментов каркасных участков (framework regions), обладающих значительно меньшей вариабельностью по сравнению с CDRs, приводит к тому, что при протеолитическом расщеплении смеси поликлональных Fab-фрагментов возникает пул пептидов, которые дают совпадения с высоким индексом достоверности с последовательностями сразу нескольких V-генов В-лимфоцитов. Это явление создаёт трудности при анализе масс-спектров пептидов, поскольку стандартный метод моделирования ошибок с применением случайно сгенерированных последовательностей (десоу-последовательностей) не учитывает данную особенность антител и приводит к высокому уровню ложноположительных identifications. Поэтому необходимо использование методов со строгим отбором для анализа спектров пептидов. Для определения процента некорректных identifications и нахождения характеристик, по которым можно различить истинные и ложные совпадения, авторы [64] применили дифференциальное мечение остатков цистеина. Это позволило составить список identifications с наибольшим соответствием данным о мечении и выявить ложноположительные identifications с высоким индексом достоверности. Образцы F(ab')₂-фрагментов кроликов были разделены на две части и алкилированы двумя различными агентами — йодацетамидом и йодэтанолам. В результате этого в обеих пробах содержались одинаковые по аминокислотной последовательности пептиды, но с разницей масс в 13 Да по модифицированным остаткам цистеина, совокупность которых служила для определения истинных identifications. Проанализировав списки ложных и истинных identifications с высоким индексом достоверности, авторы смогли добиться повышения достоверности идентификации пептидов за счёт более строгих параметров среднего отклонения массы. Они показали, что при пороге среднего

отклонения в 1 ppm наблюдается улучшение точности поиска на 87% без потери истинных identifications [64].

При идентификации результатов LC-MS/MS эксперимента по базам данных для белков исследователи зачастую сталкиваются с тем, что часть масс-спектров хорошего качества остаётся неидентифицированной. Это приводит к тому, что некоторые белки или пептиды, присутствующие в образце, могут остаться незамеченными. Для увеличения эффективности анализа масс-спектров, Lundström и соавт. [65] предложили новый подход — SpotLight Proteomics. Подход SpotLight идейно схож со стандартным bottom-up анализом антител, но имеет ряд особенностей: (i) пептиды антител анализируют методом LC-MS/MS, применяя два различных способа фрагментации (HCD и ETD); (ii) на основании расшифровки полученных MS данных с использованием *de novo* алгоритмов создают базу последовательностей антител образца; (iii) далее анализируют как исходный образец сыворотки, так и её очищенную фракцию с высоким содержанием IgG. Для идентификации пептидов по масс-спектрам исследователи объединили референсный протеом человека (Uniprot) и секвенированные *de novo* последовательности в единую SpotLight-базу данных. По полученной базе данных во всех образцах была проведена идентификация пептидов с количественной оценкой с помощью алгоритма DeMix-Q [66]. Этот подход был использован для сравнительного анализа профилей антител у пациентов с ранними стадиями болезни Альцгеймера и деменции с тельцами Леви. В исходном протеоме сыворотки были количественно определены 156 белков, в протеоме IgG-обогащённого образца — 81, из которых 23 не были обнаружены в исходном образце сыворотки. На уровне пептидов в интактной сыворотке 18% из 2112 уникальных последовательностей были ассоциированы с IgG. В образце, обогащённом антителами, долю пептидов, гомологичных IgG, составлял уже 31% из 2077 последовательностей. В исходных образцах 4 белка и 47 пептидов значительно различались у пациентов с болезнью Альцгеймера и деменцией с тельцами Леви, в IgG-обогащённых образцах различие в количестве наблюдалось для 49 пептидов, 23 из которых относятся к белкам, существенно различающимся по представленности в сравниваемых группах.

Sen и соавт. [45] продемонстрировали разработанный ими алгоритм Supernovo, который позволяет автоматизировать секвенирование моноклональных антител (рис. 7). Данный алгоритм использует MS/MS-данные пептидов антитела, полученные методом bottom-up с применением нескольких протеаз. Первым этапом является поиск пептидов по базе данных последовательностей иммуноглобулинов зародышевых линий IMGT (международная информационная система ImMunoGeneTics) и последующая рекомбинация V-, J- и C-экзонов *in silico* для создания заготовки последовательностей лёгкой и тяжёлой цепей. Далее полученная заготовка дорабатывается путём

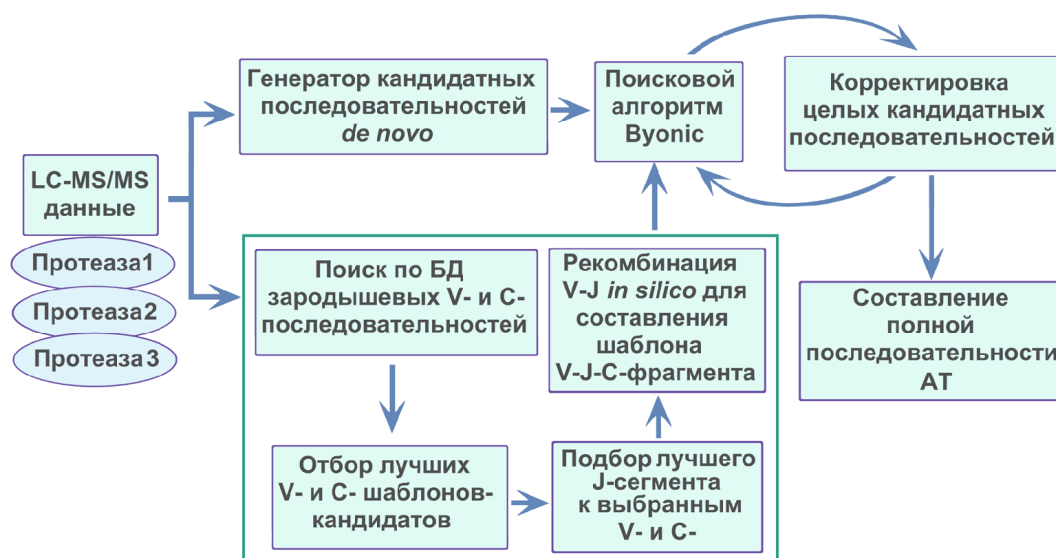


Рисунок 7. Схема исследования Sen и соавт. [45]. Слева направо: (i) по полученным экспериментальным данным одновременно проводили генерацию последовательностей пептидов *de novo* и поиск по базе данных генов иммуноглобулинов зародышевых линий ImMunoGeneTics (IMGT) для отбора последовательностей-шаблонов V- и C-сегментов с наибольшей гомологией к исследуемому антителу; (ii) после подбора J-сегмента к выбранным V- и C- проводили рекомбинацию *in silico* для составления шаблона V-J-C-сегмента последовательности антитела; (iii) на основе полученных шаблонов и *de novo* последовательностей с помощью алгоритма Byonic [67] составляли кандидатные последовательности пептидов; (iv) после нескольких циклов корректировки пептидов-кандидатов составляли полную последовательность антитела.

сопоставления с результатами *de novo* секвенирования пептидов по масс-спектрам и её корректировки с помощью пакета программ Byonic [67]. В результате нескольких повторений данного процесса собирается итоговая последовательность антитела.

Автоматизация процесса секвенирования антител имеет большое значение для фармации и биохимии, поскольку это позволит снизить временные и финансовые затраты на данный этап разработки новых терапевтических, диагностических и лабораторных МАТ.

Недавно Schulte и соавт. предложили алгоритм Stitch для оптимизации обработки данных при секвенировании антител [68]. Данный алгоритм использует последовательности пептидов, полученные *de novo* на основе MS/MS-спектров, для их картирования посредством настраиваемых шаблонов аминокислотных последовательностей (рис. 8). С помощью базы данных IMGT пептиды группируют по принадлежности к лёгким и тяжёлым цепям с помощью локального выравнивания. Также IMGT применяется в данном алгоритме для создания настраиваемых шаблонов, по которым выравниваются короткие сиквенсы, полученные *de novo* из LC-MS/MS экспериментов. Для реконструкции последовательности гипервариабельного участка CDRH3 алгоритм сопоставляет перекрывающиеся участки V- и J-частей, выравнивая их по пользовательским последовательностям-шаблонам, таким образом реконструируя не покрываемые базой данных гипервариабельные участки.

Используя эту методику, исследователи продемонстрировали применение Stitch на примере секвенирования двух моноклональных антител,

отдельного Fab-фрагмента в смеси поликлональных антител плазмы, секвенирования лёгких цепей в моче пациентов с множественной миеломой и оценки репертуара антител у людей, госпитализированных с COVID-19.

ЗАКЛЮЧЕНИЕ

Антитела неоспоримо имеют большое значение для современной молекулярной биологии, медицины и фармации. Их уникальные свойства, позволяющие специфично связывать ту или иную молекулярную мишень, дали толчок развитию множества технологий и методик, применяемых в терапии, диагностике, производстве лекарственных препаратов и в биомедицинских лабораториях. LC-MS/MS зарекомендовала себя как высокоэффективный метод для получения полной аминокислотной последовательности моноклональных антител и определения имеющихся посттрансляционных модификаций. Изучение разнообразия антител в поликлональных смесях представляет собой более сложную задачу из-за малой доли различающихся между клонотипами последовательностей в молекуле антитела. При этом использование только секвенирования генов BCR в В-клетках крови недостаточно информативно, поскольку периферические В-клетки не отражают репертуар антител в крови. Это делает масс-спектрометрический анализ незаменимым методом в изучении поликлональных антител.

Технологическое развитие масс-спектрометрического оборудования расширяет возможности применения МС для *de novo* секвенирования антител, в частности

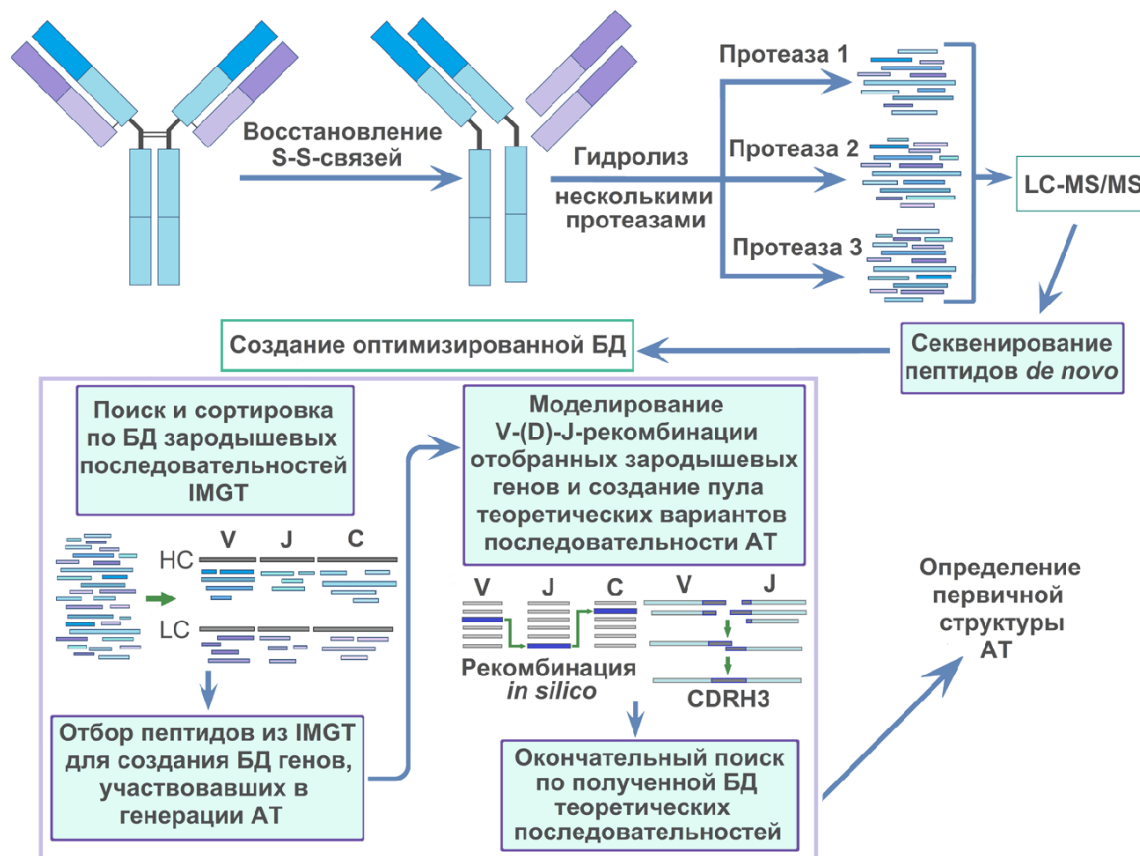


Рисунок 8. Схема исследования Schulte и соавт. [68]. В верхней части рисунка отражены пробоподготовка и тандемный хромато-масс-спектрометрический анализ пептидов антитела. В нижней части показан биоинформатический анализ полученных масс-спектров. Для идентификации пептидов и составления последовательности антитела (АТ) создавали оптимизированную базу данных (БД): для полученных *de novo* пептидов проводили поиск по базе данных генов иммуноглобулинов зародышевых линий ImMunoGeneTics (IMGT), пептиды АТ рассортировывали по принадлежности к V-, J- и C-участкам тяжёлой (HC) и лёгкой (LC) цепей. Последовательности из БД IMGT, наиболее гомологичные исследуемым пептидам, отбирали для моделирования V-(D)-J-рекомбинации *in silico*, реконструкции гипервариабельного участка тяжёлой цепи CDRH3 и создания БД теоретических последовательностей АТ. По полученной БД проводили повторный поиск для составления аминокислотной последовательности АТ.

за счёт возможности совмещения нескольких методов фрагментации ионов в одном приборе. Несмотря на то, что эффективная фрагментация целого антитела либо его крупных частей всё ещё остаётся сложной задачей, вероятно, дальнейший технологический прогресс позволит усовершенствовать масс-спектрометры в достаточной степени, чтобы внедрить подходы top-down и middle-down в рутинную протеомную практику наравне с bottom-up.

ФИНАНСИРОВАНИЕ

Работа выполнена в рамках Программы фундаментальных научных исследований в Российской Федерации на долгосрочный период (2021-2030 годы) (№122030100168-2).

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Данная работа не содержит каких-либо исследований с использованием людей и животных в качестве объектов исследования.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА

1. Yman V, Tuju J, White M.T, Kamuyu G, Mwai K, Kibinge N, Asghar M, Sundling C, Sonden K, Murungi L. et al. (2022) Distinct kinetics of antibodies to 111 *Plasmodium falciparum* proteins identifies markers of recent malaria exposure. Nat. Commun., **13**(1), 331-331. DOI: 10.1038/s41467-021-27863-8
2. Monroy-Iglesias M.J., Crescioli S., Beckmann K., Le N., Karagiannis S.N., van Hemelrijck M., Santaolalla A. (2022) Antibodies as biomarkers for cancer risk: A systematic review. Clin. Exper. Immunol., **209**(1), 46-63. DOI: 10.1093/cei/uxac030
3. Satoh M., Ceribelli A., Hasegawa T., Tanaka S. (2022) Clinical significance of antinucleolar antibodies: Biomarkers for autoimmune diseases, malignancies, and others. Clin. Rev. Allergy Immunol., **63**(2), 210-239. DOI: 10.1007/s12016-022-08931-3

4. *Alsaed O.S., Alamlah L.I., Al-Radideh O., Chandra P., Alemadi S., Al-Allaf A.-W.* (2021) Clinical utility of ANA-ELISA vs ANA-immunofluorescence in connective tissue diseases. *Sci. Rep.*, **11**(1), 8229-8229. DOI: 10.1038/s41598-021-87366-w
5. *Musharova O., Medvedeva S., Klimuk E., Guzman N.M., Titova D., Zgoda V., Shiriaeva A., Semenova E., Severinov K., Savitskaya E.* (2021) Prespacers formed during primed adaptation associate with the Cas1-Cas2 adaptation complex and the Cas3 interference nuclease-helicase. *Proc. Nat. Acad. Sci. USA*, **118**(22), e2021291118. DOI: 10.1073/pnas.2021291118
6. *Neubert H., Shuford C.M., Olah T.V., Garofolo F., Schultz G.A., Jones B.R., Amaravadi L., Laterza O.F., Xu K., Ackermann B.L.* (2020) Protein biomarker quantification by immunoaffinity liquid chromatography-tandem mass spectrometry: Current state and future vision. *Clinical Chemistry*, **66**(2), 282-301. DOI: 10.1093/clinchem/hvz2022
7. *Brummelman J., Haftmann C., Núñez N.G., Alvisi G., Mazza E.M.C., Becher B., Lugli E.* (2019) Development, application and computational analysis of high-dimensional fluorescent antibody panels for single-cell flow cytometry. *Nature Protocols*, **14**(7), 1946-1969. DOI: 10.1038/s41596-019-0166-2
8. *David M.P.C., Asprer J.J.T., Ibana J.S.A., Concepcion G.P., Padlan E.A.* (2007) A study of the structural correlates of affinity maturation: antibody affinity as a function of chemical interactions, structural plasticity and stability. *Mol. Immunol.*, **44**(6), 1342-1351. DOI: 10.1016/J.MOLIMM.2006.05.006
9. *Stanfield R.L., Wilson I.A.* (2014) Antibody structure. *Microbiology Spectrum*, **2**(2), DOI: 10.1128/microbiolspec.AID-0012-2013
10. *Snapkov I., Chernigovskaya M., Sinitcyn P., Lê Quý K., Nyman T.A., Greiff V.* (2022) Progress and challenges in mass spectrometry-based analysis of antibody repertoires. *Trends Biotechnol.*, **40**(4), 463-481. DOI: 10.1016/J.TIBTECH.2021.08.006
11. *Greiff V., Miho E., Menzel U., Reddy S.T.* (2015) Bioinformatic and statistical analysis of adaptive immune repertoires. *Trends Immunol.*, **36**(11), 738-749. DOI: 10.1016/j.it.2015.09.006
12. *Xu J.L., Davis M.M.* (2000) Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity*, **13**(1), 37-45. DOI: 10.1016/S1074-7613(00)00006-6
13. *Khass M., Vale A.M., Burrows P.D., Schroeder H.W.* (2018) The sequences encoded by immunoglobulin diversity (DH) gene segments play key roles in controlling B-cell development, antigen-binding site diversity, and antibody production. *Immunol. Rev.*, **284**(1), 106-119. DOI: 10.1111/imr.12669
14. *Robinson S.A., Raybould M.I.J., Schneider C., Wong W.K., Marks C., Deane C.M.* (2021) Epitope profiling using computational structural modelling demonstrated on coronavirus-binding antibodies. *PLoS Comput. Biol.*, **17**(12), e1009675. DOI: 10.1371/journal.pcbi.1009675
15. *Todd P.A., Brogden R.N.* (1989) Muromonab CD3. *Drugs*, **37**(6), 871-899. DOI: 10.2165/00003495-198937060-00004
16. *Lu R.-M., Hwang Y.-C., Liu I.-J., Lee C.-C., Tsai H.-Z., Li H.-J., Wu H.-C.* (2020) Development of therapeutic antibodies for the treatment of diseases. *J. Biomed. Sci.*, **27**(1), 1-1. DOI: 10.1186/s12929-019-0592-z
17. *Awwad S., Angkawinitwong U.* (2018) Overview of antibody drug delivery. *Pharmaceutics*, **10**(3), 83-83. DOI: 10.3390/pharmaceutics10030083
18. *Frejd F.Y., Kim K.-T.* (2017) Affibody molecules as engineered protein drugs. *Exper. Mol. Med.*, **49**(3), e306-e306. DOI: 10.1038/emmm.2017.35
19. *Brasino M., Roy S., Erbse A.H., He L., Mao C., Park W., Cha J.N., Goodwin A.P.* (2018) Anti-EGFR affibodies with site-specific photo-cross-linker incorporation show both directed target-specific photoconjugation and increased retention in tumors. *J. Am. Chem. Soc.*, **140**(37), 11820-11828. DOI: 10.1021/jacs.8b07601
20. *Altunay B., Morgenroth A., Beheshti M., Vogg A., Wong N.C.L., Ting H.H., Biersack H.-J., Stickler E., Mottaghy F.M.* (2021) HER2-directed antibodies, affibodies and nanobodies as drug-delivery vehicles in breast cancer with a specific focus on radioimmunotherapy and radioimmunomaging. *Eur. J. Nucl. Med. Mol. Imaging*, **48**(5), 1371-1389. DOI: 10.1007/s00259-020-05094-1
21. *Brinkmann U., Kontermann R.E.* (2021) Bispecific antibodies. *Science*, **372**(6545), 916-917. DOI: 10.1126/science.abg1209
22. *Lee S.C., Ma J.S.Y., Kim M.S., Laborda E., Choi S.-H., Hampton E.N., Yun H., Nunez V., Muldong M.T., Wu C.N., Ma W. et al.* (2021) A PSMA-targeted bispecific antibody for prostate cancer driven by a small-molecule targeting ligand. *Science Advances*, **7**(33), eabi8193. DOI: 10.1126/sciadv.abi8193
23. *Oldham R.K., Dillman R.O.* (2008) Monoclonal antibodies in cancer therapy: 25 years of progress. *J. Clin. Oncol.*, **26**(11), 1774-1777. DOI: 10.1200/JCO.2007.15.7438
24. *Hafeez U., Gan H.K., Scott A.M.* (2018) Monoclonal antibodies as immunomodulatory therapy against cancer and autoimmune diseases. *Curr. Opin. Pharmacol.*, **41**, 114-121. DOI: 10.1016/j.coph.2018.05.010
25. *Melero I., Hervas-Stubbs S., Glennie M., Pardoll D.M., Chen L.* (2007) Immunostimulatory monoclonal antibodies for cancer therapy. *Nature Reviews. Cancer*, **7**(2), 95-106. DOI: 10.1038/nrc2051
26. *Fujiwara S., Wada H., Kawada J., Kawabata R., Takahashi T., Fujita J., Hirao T., Shibata K., Makari Y., Iijima S. et al.* (2013) NY-ESO-1 antibody as a novel tumour marker of gastric cancer. *Br. J. Cancer*, **108**(5), 1119-1125. DOI: 10.1038/bjc.2013.51
27. *Lee Z.J.O., Eslick G.D., Edirimanne S.* (2020) Investigating antithyroglobulin antibody as a prognostic marker for differentiated thyroid cancer: A meta-analysis and systematic review. *Thyroid*, **30**(11), 1601-1612. DOI: 10.1089/thy.2019.0368
28. *Bogdanos D.P., Gkoutzouvelas A., Papadopoulos V., Liaskos C., Patrikiou E., Tsigalou C., Saratziotis A., Hajiioannou J., Scheper T., Meyer W., Sakkas L.I., Papandreou C.* (2021) Anti-Ro52 antibody is highly prevalent and a marker of better prognosis in patients with ovarian cancer. *Clin. Chim. Acta*, **521**, 199-205. DOI: 10.1016/j.cca.2021.07.006
29. *Scherer H.U., Huizinga T.W.J., Krönke G., Schett G., Toes R.E.M.* (2018) The B cell response to citrullinated antigens in the development of rheumatoid arthritis. *Nature Reviews. Rheumatology*, **14**(3), 157-169. DOI: 10.1038/nrrheum.2018.10
30. *Darrah E., Giles J.T., Ols M.L., Bull H.G., Andrade F., Rosen A.* (2013) Erosive rheumatoid arthritis is associated with antibodies that activate PAD4 by increasing calcium sensitivity. *Sci. Transl. Med.*, **5**(186), 186ra65. DOI: 10.1126/scitranslmed.3005370
31. *Zhou X., Motta F., Selmi C., Ridgway W.M., Gershwin M.E., Zhang W.* (2021) Antibody glycosylation in autoimmune diseases. *Autoimmun. Rev.*, **20**(5), 102804-102804. DOI: 10.1016/j.autrev.2021.102804

32. Fechtner S., Berens H., Bemis E., Johnson R.L., Guthridge C.J., Carlson N.E., Demoruelle M.K., Harley J.B., Edison J.D., Norris J.A. et al. (2022) Antibody responses to Epstein-Barr virus in the preclinical period of rheumatoid arthritis suggest the presence of increased viral reactivation cycles. *Arthritis Rheumatology*, **74**(4), 597-603. DOI: 10.1002/art.41994
33. Lu L.L., Chung A.W., Rosebrock T.R., Ghebremichael M., Yu W.H., Grace P.S., Schoen M.K., Tafesse F., Martin C., Leung V. et al. (2016) A functional role for antibodies in tuberculosis. *Cell*, **167**(2), 433-443.e14. DOI: 10.1016/j.cell.2016.08.072
34. Alter G., Ottenhoff T.H.M., Joosten S.A. (2018) Antibody glycosylation in inflammation, disease and vaccination. *Semin. Immunol.*, **39**, 102-110. DOI: 10.1016/j.smim.2018.05.003
35. Offersen R., Yu W.-H., Scully E.P., Julg B., Euler Z., Sadanand S., Garcia-Dominguez D., Zheng L., Rasmussen T.A., Jennewein M.F. et al. (2020) HIV antibody Fc N-linked glycosylation is associated with viral rebound. *Cell Reports*, **33**(11), 108502-108502. DOI: 10.1016/j.celrep.2020.108502
36. Bournazos S., Vo H.T.M., Duong V., Auerswald H., Ly S., Sakuntabhai A., Dussart P., Cantaert T., Ravetch J.V. (2021) Antibody fucosylation predicts disease severity in secondary dengue infection. *Science*, **372**(6546), 1102-1105. DOI: 10.1126/science.abc7303
37. Toby T.K., Fornelli L., Srzentić K., de Hart C.J., Levitsky J., Friedewald J., Kelleher N.L. (2019) A comprehensive pipeline for translational top-down proteomics from a single blood draw. *Nature Protocols*, **14**(1), 119-152. DOI: 10.1038/s41596-018-0085-7
38. Hirose M., Hoshida M., Ishikawa M., Toya T. (1993) MASCO: Multiple alignment system for protein sequences based on three-way dynamic programming. *Bioinformatics*, **9**(2), 161-167. DOI: 10.1093/bioinformatics/9.2.161
39. Tyanova S., Temu T., Cox J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols*, **11**(12), 2301-2319. DOI: 10.1038/nprot.2016.136
40. Barsnes H., Vaudel M. (2018) SearchGUI: A highly adaptable common interface for proteomics search and *de novo* engines. *J. Proteome Res.*, **17**(7), 2552-2555. DOI: 10.1021/acs.jproteome.8b00175
41. Vaudel M., Burkhart J.M., Zahedi R.P., Oveland E., Berven F.S., Sickmann A., Martens L., Barsnes H. (2015) PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nature Biotechnology*, **33**(1), 22-24. DOI: 10.1038/nbt.3109
42. Zhang J., Xin L., Shan B., Chen W., Xie M., Yuen D., Zhang W., Zhang Z., Lajoie G.A., Ma B. (2012) PEAKS DB: *De novo* sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics*, **11**(4), M111.010587-M111.010587. DOI: 10.1074/mcp.M111.010587
43. Frank A., Pevzner P. (2005) PepNovo: *De novo* peptide sequencing via probabilistic network modeling. *Anal. Chem.*, **77**(4), 964-973. DOI: 10.1021/ac048788h
44. Jeong K., Kim S., Pevzner P.A. (2013) UniNovo: A universal tool for *de novo* peptide sequencing. *Bioinformatics*, **29**(16), 1953-1962. DOI: 10.1093/bioinformatics/btt338
45. Sen K.L., Tang W.H., Nayak S., Kil Y.J., Bern M., Ozoglu B., Ueberheide B., Davis D., Becker C. (2017) Automated antibody *de novo* sequencing and its utility in biopharmaceutical discovery. *J. Am. Soc. Mass Spectrom.*, **28**(5), 803-810. DOI: 10.1007/s13361-016-1580-0
46. Zhang Y., Fonslow B.R., Shan B., Baek M.-C., Yates J.R. (2013) Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.*, **113**(4), 2343-2394. DOI: 10.1021/cr3003533
47. Peng W., Pronker M.F., Snijder J. (2021) Mass spectrometry-based *de novo* sequencing of monoclonal antibodies using multiple proteases and a dual fragmentation scheme. *J. Proteome Res.*, **20**, 3559-3566. DOI: 10.1021/acs.jproteome.1c00169
48. Cristobal A., Marino F., Post H., van den Toorn H.W.P., Mohammed S., Heck A.J.R. (2017) Toward an optimized workflow for middle-down proteomics. *Anal. Chem.*, **89**(6), 3318-3325. DOI: 10.1021/acs.analchem.6b03756
49. Fornelli L., Srzentić K., Huguet R., Mullen C., Sharma S., Zabrouskov V., Fellers R.T., Durbin K.R., Compton P.D., Kelleher N.L. (2018) Accurate sequence analysis of a monoclonal antibody by top-down and middle-down orbitrap mass spectrometry applying multiple ion activation techniques. *Anal. Chem.*, **90**(14), 8421-8429. DOI: 10.1021/acs.analchem.8b00984
50. Xiao Y., Vecchi M.M., Wen D. (2016) Distinguishing between leucine and isoleucine by integrated LC-MS analysis using an orbitrap fusion mass spectrometer. *Anal. Chem.*, **88**(21), 10757-10766. DOI: 10.1021/acs.analchem.6b03409
51. Mao Y., Valeja S.G., Rouse J.C., Hendrickson C.L., Marshall A.G. (2013) Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation-fourier transform ion cyclotron resonance-mass spectrometry. *Anal. Chem.*, **85**(9), 4239-4246. DOI: 10.1021/ac303525n
52. Fornelli L., Ayoub D., Aizikov K., Beck A., Tsybin Y.O. (2014) Middle-down analysis of monoclonal antibodies with electron transfer dissociation orbitrap fourier transform mass spectrometry. *Anal. Chem.*, **86**(6), 3005-3012. DOI: 10.1021/ac4036857
53. Wang Z., Liu X., Muther J., James J.A., Smith K., Wu S. (2019) Top-down mass spectrometry analysis of human serum autoantibody antigen-binding fragments. *Sci. Rep.*, **9**(1), 2345. DOI: 10.1038/s41598-018-38380-y
54. Rosati S., Yang Y., Barendregt A., Heck A.J.R. (2014) Detailed mass analysis of structural heterogeneity in monoclonal antibodies using native mass spectrometry. *Nature Protocols*, **9**(4), 967-976. DOI: 10.1038/nprot.2014.057
55. Cheng J., Wang L., Rive C.M., Holt R.A., Morin G.B., Chen D.D.Y. (2020) Complementary methods for *de novo* monoclonal antibody sequencing to achieve complete sequence coverage. *J. Proteome Res.*, **19**(7), 2700-2707. DOI: 10.1021/acs.jproteome.0c00223
56. Wheatley A.K., Juno J.A., Wang J.J., Selva K.J., Reynaldi A., Tan H.-X., Lee W.S., Wragg K.M., Kelly H.G., Esterbauer R., Davis S.K. et al. (2021) Evolution of immune responses to SARS-CoV-2 in mild-moderate COVID-19. *Nat. Commun.*, **12**(1), 1162-1162. DOI: 10.1038/s41467-021-21444-5
57. Lavinder J.J., Wine Y., Giesecke C., Ippolito G.C., Horton A.P., Lungu O.I., Hoi K.H., de Kosky B.J., Murrin E.M., Wirth M.M. et al. (2014) Identification and characterization of the constituent human serum antibodies elicited by vaccination. *Proc. Nat. Acad. Sci. USA*, **111**(6), 2259-2264. DOI: 10.1073/pnas.1317793111
58. Guthals A., Gan Y., Murray L., Chen Y., Stinson J., Nakamura G., Lill J.R., Sandoval W., Bandeira N. (2017) *De novo* MS/MS sequencing of native human antibodies. *J. Proteome Res.*, **16**(1), 45-54. DOI: 10.1021/acs.jproteome.6b00608
59. Lee J., Paparoditis P., Horton A.P., Frühwirth A., McDaniel J.R., Jung J., Boutz D.R., Hussein D.A., Tanno Y., Pappas L. et al. (2019) Persistent antibody clonotypes dominate the serum

- response to influenza over multiple years and repeated vaccinations. *Cell Host Microbe*, **25**(3), 367-376.e5. DOI: 10.1016/j.chom.2019.01.010
60. Bondt A., Dingess K.A., Hoek M., van Rijswijk D.M.H., Heck A.J.R. (2021) A direct MS-based approach to profile human milk secretory immunoglobulin A (IgA1) reveals donor-specific clonal repertoires with high longitudinal stability. *Front. Immunol.*, **12**, 789748. DOI: 10.3389/fimmu.2021.789748
 61. Wine Y., Boutz D.R., Lavinder J.J., Miklos A.E., Hughes R.A., Hoi K.H., Jung S.T., Horton A.P., Murrin E.M., Ellington A.D. et al. (2013) Molecular deconvolution of the monoclonal antibodies that comprise the polyclonal serum response. *Proc. Nat. Acad. Sci.*, **110**(8), 2993-2998. DOI: 10.1073/pnas.1213737110
 62. Shaw J.B., Liu W., Vasil'ev Y.V., Bracken C.C., Malhan N., Guthals A., Beckman J.S., Voinov V.G. (2020) Direct determination of antibody chain pairing by top-down and middle-down mass spectrometry using electron capture dissociation and ultraviolet photodissociation. *Anal. Chem.*, **92**(1), 766-773. DOI: 10.1021/acs.analchem.9b03129
 63. Srzentić K., Nagornov K.O., Fornelli L., Lobas A.A., Ayoub D., Kozhinov A.N., Gasilova N., Menin L., Beck A., Gorshkov M.V., Aizikov K., Tsybin Y.O. (2018) Multiplexed middle-down mass spectrometry as a method for revealing light and heavy chain connectivity in a monoclonal antibody. *Anal. Chem.*, **90**(21), 12527-12535. DOI: 10.1021/acs.analchem.8b02398
 64. Boutz D.R., Horton A.P., Wine Y., Lavinder J.J., Georgiou G., Marcotte E.M. (2014) Proteomic identification of monoclonal antibodies from serum. *Anal. Chem.*, **86**(10), 4758-4766. DOI: 10.1021/ac4037679
 65. Lundström S.L., Zhang B., Rutishauser D., Aarsland D., Zubarev R.A. (2017) SpotLight proteomics: Uncovering the hidden blood proteome improves diagnostic power of proteomics. *Sci. Rep.*, **7**(1), 41929-41929. DOI: 10.1038/srep41929
 66. Zhang B., Pirmoradian M., Chernobrovkin A., Zubarev R.A. (2014) DeMix workflow for efficient identification of cofragmented peptides in high resolution data-dependent tandem mass spectrometry. *Mol. Cell. Proteomics*, **13**(11), 3211-3223. DOI: 10.1074/mcp.O114.038877
 67. Bern M., Kil Y.J., Becker C. (2012) Byonic: Advanced peptide and protein identification software. *Curr. Protoc. Bioinformatics*, **40**(1), 13.20.1-13.20.14. DOI: 10.1002/0471250953.bi1320s40
 68. Schulte D., Peng W., Snijder J. (2022) Template-based assembly of proteomic short reads for *de novo* antibody sequencing and repertoire profiling. *Anal. Chem.*, **94**(29), 10391-10399. DOI: 10.1021/acs.analchem.2c01300

Поступила в редакцию: 08. 11. 2022.
После доработки: 26. 12. 2022.
Принята к печати: 28. 12. 2022.

ANTIBODY PROTEOMICS

*L.Sh. Kazieva, T.E. Farafonova, V.G. Zgoda**

Institute of Biomedical Chemistry,
10 Pogodinskaya str., Moscow, 119121 Russia; *e-mail: victor.zgoda@gmail.com

Antibodies represent an essential component of humoral immunity; therefore their study is important for molecular biology and medicine. The unique property of antibodies to specifically recognize and bind a certain molecular target (an antigen) determines their widespread application in treatment and diagnostics of diseases, as well as in laboratory and biotechnological practices. High specificity and affinity of antibodies is determined by the presence of primary structure variable regions, which are not encoded in the human genome and are unique for each antibody-producing B cell clone. Hence, there is little or no information about amino acid sequences of the variable regions in the databases. This differs identification of antibody primary structure from most of the proteomic studies because it requires either B cell genome sequencing or *de novo* amino acid sequencing of the antibody. The present review demonstrates some examples of proteomic and proteogenomic approaches and the methodological arsenal that proteomics can offer for studying antibodies, in particular, for identification of primary structure, evaluation of posttranslational modifications and application of bioinformatics tools for their decoding.

The whole English version is available at <http://pbmc.ibmc.msk.ru>.

Key words: antibodies; mass-spectrometry; *de novo* antibody sequencing; antibody repertoire

Funding. The work was performed within the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021-2030) (No. 122030100168-2).

Received: 08.11.2022; revised: 26.12.2022; accepted: 28.12.2022.