

ТЕХНОЛОГИИ БЕЛКОВОЙ ИНТЕРАКТОМИКИ

© 2011 г. А. С. Иванов[#], В. Г. Згода, А. И. Арчаков

НИИ биомедицинской химии им. В.Н. Ореховича РАМН, 119121, Москва,
Погодинская ул., 10

Поступила в редакцию 08.06.2010 г. Принята к печати 04.08.2010 г.

Исследования белок-белковых взаимодействий обусловлены необходимостью более глубокого изучения функций белков, которые, как правило, работают в живых системах за счет взаимодействия друг с другом в составе стабильных или динамических белковых комплексов. В представляемом обзоре дается краткое описание и сравнительная оценка наиболее часто используемых в белковой интерактомике методов.

Ключевые слова: интерактомика; белок-белковые взаимодействия; биоинформатика; протеомика; геномные технологии; молекулярный фишинг; оптический биосенсор; белковая масс-спектрометрия.

ВВЕДЕНИЕ

Биологические процессы в живых системах в большинстве своем происходят с участием огромного количества различных белков. При этом функционирование белковых молекул часто осуществляется за счет их взаимодействия друг с другом в составе стабильных или динамических белковых комплексов. Число и разнообразие белок-белковых взаимодействий (ББВ) у высших организмов настолько велико, что их графическое представление имеет вид крайне сложных и запутанных сетей [1–8]. Данные сети приобретают еще большую сложность, когда привлекается дополнительная информация из геномных и протеомных исследований [9, 10].

Экспериментальные исследования ББВ *in vitro* являются популярным подходом в относительно молодой науке, названной интерактомикой. В многоклеточном организме может существовать большое число разных интерактомонов в зависимости от типа ткани, биологического

Сокращения: ББВ – белок-белковые взаимодействия; DBD – ДНК-связывающий домен фактора транскрипции (DNA-Binding Domain); AD – активационный домен фактора транскрипции (Activation Domain); HTS – высокопроизводительные технологии скрининга (High-Throughput Screening); ORFs – открытые рамки считывания (Open Read Frames); SGA – синтетические генетические матрицы (Synthetic Genetic Array); GST – глутатион-S-трансфераза (Glutathione S-Transferase); 6xHis – специфическая пептидная метка в виде 6 остатков гистидина; NTA – нитрилтриуксусная кислота; белок-«наживка» – (от англ. bait protein) белок, который используется в экспериментальных методах для «вылавливания» и идентификации неизвестных белков-партнеров по ББВ; белок-«добыча» – (от англ. prey protein) неизвестный белок, который экспериментально идентифицируют как белок, взаимодействующий с белком-«наживкой»; TAP – тандемная аффинная очистка целевых белков (Tandem Affinity Purification); CBP – кальмодулинсвязывающий пептид (Calmodulin Binding Peptide); TEV – протеиназа вируса гравировки табака (Tobacco etch virus); EGTA – комплексон (этиленгликольтетраацетат); SPR – эффект поверхностного плазмонного резонанса (Surface Plasmon Resonance); 2DE – двумерный электрофорез (two-dimensional electrophoresis); 1DE – одномерный электрофорез (one-dimensional electrophoresis).

[#] Автор для связи (тел.: (499) 246-36-93; эл. почта: alexei.ivanov@ibmc.msk.ru).

состояния и так далее. Поэтому информация о ББВ крайне важна для понимания большинства сложных молекулярных механизмов функционирования живых систем. Интерактомные данные также крайне необходимы в протеомных исследованиях для выяснения функции белков, которые идентифицируются геномными и/или протеомными методами, но не могут быть аннотированы из-за отсутствия известных белков-гомологов. Подсказку о функции неизвестного белка можно получить, исследуя его взаимодействие с другими белками, функции которых уже известны. Этот подход базируется на концепции о том, что если функция одного белка известна, то функция его партнера по ББВ, вероятно, будет с ней связана. Более того, наблюдающаяся у разных организмов консервативность определенных частей интерактома может давать дополнительную информацию о эволюции живых систем.

Необходимо также отметить, что, поскольку большинство клеточных процессов регулируется мультимерными белковыми комплексами, то нарушение ББВ может иметь далекие последствия, вплоть до развития специфических заболеваний. Например, онкогенные аденовирусы вызывают быструю безостановочную пролиферацию клетки хозяина путем нарушения взаимодействий между белками, регулирующими клеточный цикл [11]. Идентификация подобных белков представляет большой интерес, так как они могут рассматриваться как потенциальные мишени для создания новых типов лекарств, действующих через регуляцию ББВ. Поэтому в практическом плане результаты интерактомных исследований позволяют разрабатывать инновационные подходы к созданию лекарств нового поколения, действующих не на отдельные белки-мишени, а на ББВ, блокируя формирование функционально активных белковых комплексов [12–17].

Существующее разнообразие методических подходов, применяемых для изучения ББВ, обусловлено тем, что белковая интерактомика находится на стыке различных научных дисциплин, таких, как биохимия, геномика, биоинформатика, компьютерное молекулярное моделирование, клеточная биология, молекулярная биология, биофизика и др. Условно все подходы в интерактомике можно разделить на две группы: биоинформационные и экспериментальные. В свою очередь каждая из этих групп включает в себя много разнообразных подходов в исследовании ББВ (таблица). При таком разнообразии невозможно в одной работе описать все существующие подходы и методы белковой интерактомики, поэтому цель данного обзора – краткое ознакомление с рядом подходов, наиболее часто используемых в данной области.

1. БИОИНФОРМАЦИОННЫЕ ПОДХОДЫ

Данная группа объединяет многочисленные и достаточно разнообразные вычислительные методы [18, 19], которые ставят своей целью анализ имеющихся геномных, протеомных и интерактомных экспериментальных данных о ББВ и предсказание новых взаимодействий на основе выявляемых закономерностей.

Биоинформационные подходы могут быть классифицированы в пять базовых категорий (см. таблицу) [20], основанных на: 1) геномной информации, 2) эволюционных отношениях, 3) трехмерной структуре белка, 4) белковых доменах, 5) первичной структуре белка.

1.1. Подходы, основанные на геномной информации

Методы, базирующиеся на результатах полного геномного секвенирования, используют для предсказания ББВ информацию о всех имеющихся генах и их организации в геноме, что позволяет выдвигать предположения о функциональной связанности ряда генов. Для этого используется не только анализ сходства последовательностей соответствующих генов (или их продуктов) [21–24], но и оценка функциональных связей между парами или группами совместно расположенных генов (метод позиционно связанных генов).

Однако этот подход не может быть использован в случае эукариот, так как совместное регулирование генов не применимо на уровне структуры генома [25]. Иногда в ходе эволюции соседствующие функционально взаимосвязанные гены сливаются друг с другом. Продуктом трансляции такого гена является мультидоменный белок. В результате в одних организмах существуют отдельные функционально связанные белки, а в других – единый белок слияния (подобные белки еще называют белками «Розеттского камня» [26–28]). Слияние белковых доменов указывает на то, что белки, соответствующие отдельным доменам, вероятнее всего образуют комплекс друг с другом.

Был разработан ряд вычислительных методов поиска таких слитных белков одновременно во многих геномах и с их помощью были предсказаны тысячи ББВ. Многие из этих предполагаемых взаимодействий были позже подтверждены экспериментально. Поэтому данный анализ слияния генов был включен в алгоритм предсказания ББВ и слияния белков [26–28].

1.2. Подходы, основанные на эволюционных взаимосвязях белков

Методы, базирующиеся на эволюционной взаимосвязи, используют для предсказания ББВ или функциональной взаимосвязи белков так называемые филогенетические профили [29, 30]. Филогенетический профиль белка описывает наличие гомологов в серии организмов. Белки, имеющие сходные профили, могут быть функционально связаны [31]. Например, белки, входящие в один и тот же мультимерный белковый комплекс, или белки из одного биохимического пути имеют сходные филогенетические профили. С помощью сравнительного геномного филогенетического анализа были предсказаны сотни пар ББВ в *E. coli* и тысячи в дрожжах [32].

1.3. Подходы, основанные на трехмерной структуре белка

Методы предсказания ББВ, базирующиеся на данных о трехмерной структуре белков, в последнее время приобретают все большее значение из-за быстрого роста числа исследованных пространственных структур белков [33–37]. Было обнаружено, что гомологичные белки (>30% идентичности последовательностей) участвуют в одних и тех же или сходных ББВ [37]. Метод оценивает возможность формирования молекулярного комплекса с известной пространственной структурой при участии двух потенциально взаимодействующих белков и предсказывает, как ББВ может произойти. Было показано, что аминокислотные остатки в зоне контакта белков (контактный интерфейс) являются структурно консервативными [38]. Описание структуры контактного интерфейса осуществляется путем анализа известной пространственной структуры белкового комплекса, а затем делается оценка возможности формирования аналогичных комплексов белками, которые гомологичны тем, что формируют комплекс. Все возможные пары ББВ между двумя белковыми семьями могут быть смоделированы, наиболее вероятные ББВ определены и даже ранжированы по возможной специфичности взаимодействия [37]. Данный метод доступен в виде Интернет-сервиса InterPreTS (Interaction Prediction through Tertiary Structure) (<http://www.russell.embl-heidelberg.de/interprets/>) [39], который позволяет идентифицировать гомологичные пары белков по базе данных взаимодействующих доменов DBID (Database of interacting domains) из известных структур белковых комплексов, оценить степень сохранения структуры контактных интерфейсов и визуализировать молекулярные детали предсказанного ББВ.

Алгоритм многомерного тридинга (MULTIPROSPECTOR) [33] способен узнавать белки-партнеры, вовлеченные в ББВ, и корректно предсказывать значительное число белковых пар, которые уже описаны у дрожжей в базе данных DIP (Database of Interacting Proteins, <http://dip.doe-mbi.ucla.edu>). Метод правильно узнает большинство гомо- и гетеро-димеров белков, которые имеются в белковом банке PDB (Protein Data Bank, <http://www.rcsb.org/pdb/home/home.do>) [40].

1.4. Подходы, основанные на информации о белковых доменах

Существует довольно много вычислительных методов, которые базируются исключительно на анализе консерватизма белковых доменов и оценке вероятности взаимодействий между каждой парой доменов из базы данных Pfam (protein families database, <http://pfam.sanger.ac.uk/>) [41–45].

Эти рассчитанные взаимодействия доменов далее используются для предсказания ББВ.

Другой алгоритм предсказания ББВ в системе PreSPI (Prediction System for Protein Interaction, <http://prespi.icu.ac.kr/>), основанный на консерватизме домен-доменных взаимодействий [46–49], предсказывает вероятность ББВ как результат взаимодействия многих пар доменов. Для предсказания ББВ на основе информации о парах возможно взаимодействующих доменов может быть использована база данных потенциально взаимодействующих пар доменов PID (Potentially

Interacting Domain pairs, <http://www.biointeraction.org/>) [50, 51].

1.5. Подходы, основанные на первичной структуре белков

Подходы, основанные на анализе первичной структуры белков, базируются на гипотезе, что ББВ могут быть выявлены путем нахождения нескольких коротких специфических полипептидных последовательностей (сигнатурные последовательности). Эти последовательности не соответствуют целым белковым доменам, но часто обнаруживаются в клеточных белках. Существует подход [52–54], объединяющий предсказания, полученные с помощью различных вычислительных методов, и экспериментальные данные. Было показано, что характерные для ББВ пары сигнатурных последовательностей могут быть идентифицированы при анализе первичных структур белков в базе данных экспериментально определенных ББВ.

Разрабатываются подходы автоматического распознавания и предсказания ББВ на основе анализа первичных структур и ассоциированных физико-химических свойств с использованием метода опорных векторов (support vector machine) [55–58], ядерных методов (kernel methods) [59, 60] или сервера PIPE (Protein-Protein Interaction Prediction Engine, <http://pipe.cgmlab.org>), использующего данные о первичной структуре белков из баз данных MIPS (Munich Information Center for Protein Sequences, <http://mips.gsf.de>) и DIP.

2. ЭКСПЕРИМЕНТАЛЬНЫЕ ПОДХОДЫ В ИССЛЕДОВАНИИ ББВ

В последнее время в белковой интерактомике все большее распространение получают экспериментальные подходы, что привело к быстрому росту объема фактических данных о ББВ и соответственно начало тормозить развитие вычислительных (биоинформационных) подходов [20]. Однако последние не утрачивают своего значения, так как являются дополняющими к экспериментальным подходам и обеспечивают системный анализ экспериментальных данных, а также предсказание новых ББВ и планирование новых экспериментальных исследований для их обнаружения.

Экспериментальные методы исследования ББВ базируются на одном общем принципе, который очень схож с принципом рыбной ловли, что нашло свое отражение в используемой терминологии: белок-«наживка» (bait protein), белок-«добыча» (prey protein), молекулярная «рыбалка» (molecular fishing). Белком-наживкой называют известный белок, который экспериментатор использует для «вылавливания» и идентификации одного или нескольких неизвестных белков-партнеров по ББВ. Соответственно последние называются белками-«добычей».

2.1. Геномные подходы

Дрожжевая двугибридная система

Основным подходом в этой группе является двугибридная дрожжевая система (система

двойных гибридов дрожжей), которая была первоначально разработана как генетический тест для детекции ББВ в клеточной регуляции [61]. Данный подход базируется на обнаруженном факте, что многие факторы транскрипции эукариот могут быть разделены на два функционально различных домена, один из которых контролирует связывание ДНК, а другой – активацию транскрипции. В классическом варианте двугибридной дрожжевой системы (рис. 1), белок-"наживка" (bait protein) представляет собой белок слияния (fusion protein), состоящий из целевого белка X и ДНК-связывающего домена (DNA-binding domain, DBD) фактора транскрипции (DBD-X). Белок-"добыча" (prey protein) также является белком слияния, состоящим из белка Y и активационного домена (activation domain, AD) того же фактора транскрипции (AD-Y). Оба белка совместно коэкспрессируются в дрожжевых клетках. Белок DBD-X связывается с сайтом оператора в промоторной области, но не активирует транскрипцию гена-репортера из-за отсутствия AD домена. В случае взаимодействия белков X и Y происходит формирование из доменов DBD и AD функционально активного фактора транскрипции, что приводит к включению транскрипции гена-репортера, в качестве которого обычно используется ген *lacZ*, кодирующий бактериальную β -галактозидазу. Активность *lacZ* может быть легко определена с помощью колориметрии.

Простота и быстрота двугибридной дрожжевой системы сделали ее наиболее часто используемой для обнаружения новых ББВ [62, 63]. Так как это генетическая система, то она хорошо подходит для использования в высоко производительных технологиях скрининга (high-throughput screening, HTS) для идентификации ББВ между всеми белками, экспрессируемыми в определенном типе клеток или в целом организме. В результате применения таких подходов произошел бурный рост числа работ, которые породили многочисленные и крайне сложные интерактомные карты полных геномов разных организмов. Эта объемная информация хранится в различных базах данных типа IntAct [64] и DIP [65–67].

Однако в последнее время все чаще высказываются вполне обоснованные критические оценки эффективности и достоверности подобных результатов. Отмечается крайне высокий уровень ложноположительных определений ББВ. В результате среднее число ББВ для каждого белка достигает 5–8 и более [68]. По-видимому, это обусловлено плохим распознаванием специфических и не специфических взаимодействий.

Оказалось также, что результаты исследования ББВ с помощью двугибридных дрожжевых систем зависят от конкретной экспериментальной системы. Так например, анализ одних и тех же 6000 открытых рамок считывания (ORFs) с помощью двух систем с разными парами DBD и AD показал всего лишь 20% совпадение данных о ББВ [69–71]. Отмечается, что фон не воспроизводящихся ложноположительных результатов может достигать даже 90% от всех обнаруженных ББВ [72, 73]. Аналогично, при сравнении данных широкомасштабных исследований ББВ в дрожжах, которые были выполнены двумя группами исследователей, был

получен низкий уровень совпадений. При этом были выявлены как ложноположительные, так и ложноотрицательные результаты [74]. Позже было также показано плохое совпадение результатов анализа ББВ в различных организмах [75].

Высокий уровень ложных результатов анализа ББВ (как ложноотрицательных, так и ложноположительных) при использовании двугибридной дрожжевой системы крайне осложняет интерпретацию биологической значимости ББВ.

Возможными причинами ложноотрицательных результатов (то есть не обнаружение реально существующих ББВ) могут быть: 1) «неправильные» конструкции белков слияния (например, для реального ББВ необходимо, чтобы были свободны *N*- и/или *C*-концы белков-партнеров) или возникновение ошибок при широкомасштабном конструировании белков слияния; 2) по сравнению со свободными белками-партнерами, экспрессия и/или фолдинг белков слияния происходят по-другому; 3) возникновение мутаций при выполнении ПЦР при амплификации дрожжевых ORFs; 4) использование объединенных образцов, содержащих смеси белков-слияния (такой прием часто используется в широкомасштабных HTS-исследованиях для сокращения времени и стоимости анализа), что при избыточности объединения приводит к ложноотрицательным результатам из-за снижения чувствительности [76].

Еще более серьезной проблемой является крайне высокий уровень ложноположительных результатов (то есть обнаружение реально не существующих ББВ). Источники ложноположительных данных могут быть сгруппированы в два класса: 1) экспериментальные ошибки (например, плохая воспроизводимость анализа из-за чрезмерной экспрессии белка-наживки; регистрация не специфических взаимодействий белков слияния как позитивный результат); 2) биологическое несоответствие (ряд ББВ, обнаруженных с помощью двугибридных систем и физически подтвержденных другими методами, не являются биологически значимыми), что обусловлено гетеротопической экспрессией белков, которые в природе никогда не встречаются из-за различия профилей экспрессии и/или субклеточной локализации.

Ложноположительное срабатывание системы при не специфическом взаимодействии белков приводит к регистрации в интерактомных данных так называемых «липких» белков, которые не специфически взаимодействуют с большим числом разнообразных белков. На интерактомных картах такие белки, как правило, образуют главные узлы сети, соединенные с огромным числом белков-партнеров. Если это не ложноположительный результат, то с физической точки зрения такие белки должны иметь или очень большие размеры, чтобы всем белкам-партнерам хватило места в мультимерном комплексе, или их взаимодействия должны быть разнесены по времени и механизмам. Однако наиболее вероятно, что в большинстве случаев такие ББВ являются ложноположительным результатом.

Синтетические генетические матрицы

Помимо двугибридных систем, к группе геномных подходов относятся так называемые

синтетические генетические матрицы (Synthetic Genetic Array, SGA), которые представляют собой геномный подход к широкомасштабному анализу генетических взаимосвязей путем систематического конструирования двойных мутантов [77–79]. В основу данного подхода легло наблюдение [77], что в дрожжах *Saccharomyces cerevisiae* среди примерно 6200 генов более 80% не являются жизненно важными (штамм продолжает жить при удалении по одиночке любого из этих генов), что говорит о высокой степени буферизации геномных пертурбаций. Авторы разработали метод анализа функциональной взаимосвязи белков путем проверки жизнеспособности штаммов с одновременным сочетанием двух мутаций. Типичный анализ SGA состоит из матрицы пересечений мутации целевого белка с приблизительно 5000 жизнеспособных мутантных штаммов с одиночными делециями генов с целью выявления двойных мутантов с дефектом роста. Нарушение роста конкретного двойного мутанта свидетельствует о функциональной взаимосвязи генов (белков). В плане анализа ББВ с помощью данной технологии, предполагается, что полученные позитивные результаты анализа функциональной взаимосвязи генов (белков) частично обусловлены нарушением прямых ББВ.

Корреляционный профиль экспрессии мРНК

К той же группе геномных подходов можно отнести так называемый корреляционный профиль экспрессии мРНК, основанный на анализе корреляций транскриптомных и интерактомных данных [80]. Была разработана стратегия картирования корреляций транскриптома и интерактома путем сравнения взаимодействий между белками, которые кодированы генами из общих кластеров профилей экспрессии, и белками, которые кодированы генами из разных кластеров. На примере *S. cerevisiae* было показано, что гены со сходными профилями экспрессии чаще кодируют взаимодействующие белки. Таким образом, анализ корреляции транскриптомных и интерактомных данных может быть использован для улучшения качества гипотез, основанных на информации от обоих подходов.

2.2. Биохимические подходы

Методы на основе твердофазной аффинной хроматографии

Для обнаружения конкретного целевого белка в сложной биологической системе необходим высокоаффинный реагент, узнающий данный белок с высокой селективностью. В качестве таких реагентов чаще всего используются антитела, узнающие определенные антигенные детерминанты природных или гетерологических белков. Для исследования ББВ часто используются антитела для выделения целевых белков в комплексе с белками-партнерами [81]. Такой подход называется коиммунопреципитацией. Для этих целей могут быть использованы и другие аффинные реагенты, например, хроматографическая смола с конъюгированным глутатионом для выделения белков-партнеров в комплексе с целевым белком с меткой в виде глутатион-S-трансферазы (glutathione S-transferase, GST), или металл-хелатный сорбент с ионами никеля для выделения белковых комплексов с белком, имеющим метку в виде б остатков

гистидина (6xHis).

Коиммунопреципитация. В этом подходе применяются специфические антитела для выделения из клеточного лизата целевых белков (bait-белки) в комплексе с белками-партнерами [82–85]. Далее используются различные методы разделения и анализа белков-партнеров. Выбор антител является ключевым звеном в данной технологии. Антитела к целевым белкам должны быть высокоаффинными (для достаточно прочного связывания антигена и его выделения из смеси) и высокоспецифичными (для минимизации не специфических взаимодействий). Для этих целей используются как моноклональные, так и поликлональные антитела. Антитела обоих типов имеют свои достоинства и недостатки: первые – более специфичны, но менее толерантны к изменениям антигена, так как связываются только с одним эпитопом, который может быть частично или полностью закрыт в комплексе с белком-партнером; вторые представляют собой смесь антител, которые взаимодействуют со множеством эпитопов, что делает их более толерантными к изменениям антигена, но менее специфичными.

Описываемый подход начинается с добавления антител в клеточный лизат и их связывания с целевым белком в составе комплекса с белками-партнерами (рис. 2). Преципитация образовавшихся комплексов антитело/целевой белок/белки-партнеры осуществляется на сефарозе, конъюгированной с белком А (из *S. aureus*) или белком G (из *Streptococcus*), специфически связывающимися с Fc-доменом иммуноглобулинов. Такой вариант представляется наиболее простым и предпочтительным, так как используется нативный белок-наживка и нет необходимости прибегать к экспрессии его генно-модифицированных вариантов. К сожалению, далеко не всегда для целевых белков существуют антитела, и это заставляет исследователей все-таки идти на модификацию целевых белков методами генной инженерии. Чаще всего это делается путем слияния целевого белка с какой-либо специфической меткой в виде пептидного эпитопа, на который имеются селективные антитела, например, Мус (пептидный эпитоп из фактора транскрипции с-Мус), FLAG (октапептид N-DYKDDDDK-C) или гемагглютинин (HA). Такие модифицированные белки экспрессируются в клеточной культуре и потом выделяются из клеточного лизата в комплексе с белками-партнерами с помощью коиммунопреципитации, то есть в клеточный лизат добавляется соответствующее антитело и образовавшиеся комплексы выделяются на сефарозе с белком А или белком G.

Аффинная хроматография белков со специфической меткой. В отличие от коиммунопреципитации, в данном подходе отсутствует этап добавления антител в клеточный лизат. Целевой белок в комплексе с белками-партнерами сразу выделяется из лизата с помощью аффинного сорбента. Как и в случае коиммунопреципитации, если к целевому белку есть специфическое антитело, то метод может быть реализован с нативным, не модифицированным белком. В противном случае (при отсутствии такого антитела), используются генно-модифицированные целевые белки с включенными в их состав различными метками: эпитопы

типа Мус, FLAG или HA; GST; 6xHis. Соответственно аффинный сорбент представляет собой носитель (сефароза, ионообменная смола или микрочастицы) с закрепленным на нем аффинным реагентом (антитела для выделения белков, меченных пептидными эпитопами; глутатион для выделения белков, меченных GST; нитрилотриуксусная кислота (NTA) для выделения белков с 6xHis-меткой).

Использование GST-метки обеспечивает высокоспецифическое и прочное связывание меченого белка на сорбенте с глутатионом [86–89]. Это обусловлено специфичностью и прочностью связывания субстрата (глутатион) в активном центре фермента (GST).

Использование 6xHis-метки – крайне популярный прием для экспрессии и выделения целевых белков из-за своей простоты и надежности (рис. 3). Аффинное выделение 6xHis-меченых белков (в комплексе с белками-партнерами) может быть осуществлено на аффинном сорбенте с NTA в присутствии ионов никеля (6xHis образует три хелатных комплекса с NTA при участии соответственно трех ионов Ni^{2+}) [90–94] или на сорбенте со специфическими антителами на 6xHis-эпитоп [95, 96].

Тандемная аффинная очистка (Tandem Affinity Purification, TAP). Данный подход был разработан для повышения эффективности выделения белковых комплексов из клеточного лизата путем одновременного использования двух аффинных меток в составе целевого белка. Протокол аффинного выделения и очистки белковых комплексов по технологии TAP содержит соответственно две стадии с использованием двух аффинных сорбентов [97–102]. Типичный вариант TAP основан на использовании генно-модифицированного целевого белка с двумя аффинными метками, разделенными специфическим сайтом TCS (TEV cleavage site, пептидная последовательность ENLYFQG) разрезания белковой цепи протеиназой вируса гравировки табака (Tobacco Etch Virus, TEV). В качестве меток используются: белок А из *S. aureus* (или его IgG-связывающий домен) и кальмодулинсвязывающий пептид (CBP) (рис. 4). На первой стадии выделяется комплекс целевого белка с белками-партнерами за счет связывания первой метки (белок А) с аффинным сорбентом, содержащим IgG (рис. 4-1) и отмывки (рис. 4-2). Далее целевой белок разрезается TEV-протеиназой (рис. 4-3) с отделением белка А, который остается на сорбенте, и открытием доступа к CBP. При этом белковый комплекс смывается с сорбента (рис. 4-4). На второй стадии аффинной очистки (рис. 4-5) белковый комплекс выделяется в присутствии ионов кальция на микрочастицах с иммобилизованным кальмодулином (СМ). После аккуратной отмывки (рис. 4-6) осуществляется элюция белкового комплекса (рис. 4-7) путем связывания ионов кальция комплексоном (EGTA), что приводит к распаду комплекса CBP-СМ.

TAP-технология имеет два основных преимущества: 1) все процедуры выполняются в близких к физиологическим условиям, 2) отмывки белковых комплексов происходят в мягких условиях, не вызывающих их разрушения.

Молекулярный фишинг на чипе оптического биосенсора

Для количественных измерений ББВ может быть использован оптический биосенсор, основанный на эффекте поверхностного плазмонного резонанса (surface plasmon resonance, SPR) [103, 104]. SPR позволяет регистрировать ББВ в реальном времени в виде сенсограмм (графики зависимости сигнала биосенсора от времени). Для этого целевой белок иммобилизуется в измерительном канале на поверхности оптического чипа, а растворы других белков пропускаются через проточную систему. При протекании рабочего буфера через измерительный канал биосенсора записывается базовая линия, далее при прохождении раствора белка-партнера происходит его связывание с целевым белком на поверхности чипа, что отражается на сенсограмме в виде подъема уровня сигнала. По окончании инъекции раствора белка в измерительный канал подается рабочий буфер и регистрируется кинетика распада белкового комплекса. Простой анализ серии сенсограмм, отвечающих инъекции белка-партнера в разных концентрациях, позволяет рассчитывать константы скоростей образования и распада белковых комплексов, а также константу диссоциации (K_d). Анализ нескольких серий таких сенсограмм, полученных при разной температуре, позволяет рассчитать термодинамические параметры ББВ (изменение энергии Гиббса, энтропии и энтальпии при образовании белкового комплекса).

Если на поверхности оптического чипа закреплен аффинный реагент для связывания целевого белка (например, антитела к целевому белку или к специфическому эпитопу; NTA для хелатного связывания белков с 6xHis-меткой в присутствии ионов никеля; глутатион для связывания белков с меткой GST), то данный белок может быть «пойман» непосредственно на оптическом чипе при пропускании клеточного лизата через измерительный канал биосенсора. Данный процесс похож на рыбную ловлю, когда в водном потоке на специфическую наживку (bait) ловится рыба (prey – добыча). Поэтому в научной литературе данный подход часто называется "молекулярной рыбалкой" (molecular fishing).

Этот подход может быть также использован для прямой копреципитации или аффинного выделения комплекса целевого белка с белками-партнерами непосредственно на оптическом чипе при пропускании клеточного лизата через измерительный канал биосенсора [105–107]. При этом процесс связывания можно контролировать в реальном времени с помощью SPR (регистрируется сколько аналита и как быстро связывается с иммобилизованным аффинным лигандом). Однако оптический биосенсор не осуществляет идентификацию выделенных белков и поэтому для интерактивного анализа он используется совместно с масс-спектрометрической идентификацией [108–113]. Такой комбинированный подход состоит из трех этапов: 1) выделение целевого белка в комплексе с белками-партнерами на оптическом чипе биосенсора при инъекции клеточного лизата (или его хроматографической фракции) ("молекулярная рыбалка"); 2) элюция выделенных белков с оптического чипа; 3) масс-спектрометрический анализ и идентификация белков.

SPR-биосенсоры также успешно используются для валидации и количественной оценки

белок-белковых и белок-лигандных взаимодействий, выявленных (предсказанных) различными методами [114–117].

Масс-спектрометрическая идентификация белков

Ключевым этапом протеомного анализа является идентификация белковых последовательностей при помощи комбинации высокопроизводительных методов разделения белков и их масс-спектрометрического анализа [118].

Несмотря на разнообразие методов, используемых в современных протеомных исследованиях, все они основаны на трех технологических платформах:

1) Использование двумерного электрофореза в комбинации с идентификацией белков методом MALDI-TOF (Matrix Assisted Laser Desorption/Ionization Time-of-Flight) масс-спектрометрии.

2) Использование одномерного электрофореза в полиакриламидном геле в комбинации с обращенно-фазовой жидкостной хроматографией (revers phase liquid chromatography, RP-LC), совмещенной с тандемной масс-спектрометрической детекцией (RP-LC-MS/MS).

3) Использование безгелевой технологии MudPiT (Multidimensional Protein Identification Technology) при помощи многомерного хроматографического разделения белков с последующим масс-спектрометрическим анализом (2DLC-MS/MS и 3DLC-MS/MS).

Двумерный электрофорез (two-dimensional electrophoresis, 2DE), разработанный О'Фареллом [119] в 1975 г., представляет собой метод разделения, основанный на последовательном использовании двух свойств белков: заряда и массы. Показано, что в одном эксперименте методом двумерного гель-электрофореза можно разделить более 2000 белков и их модифицированных изоформ, а теоретическая разрешающая способность 2DE может достигать 10 000 белков [120].

Для визуализации белковых молекул можно использовать различные методы их окрашивания. Наиболее часто используются методы серебрения, окраски Кумасси голубым и рядом флуоресцентных красителей (см. обзор [121]). Наиболее чувствительные методики позволяют зарегистрировать белки в концентрации до 10^{-9} М [122].

Для изучения белковых комплексов часто используется нативный (неденатурирующий) 2DE или голубой нативный 2DE. Разница между этими двумя протоколами незначительна и заключается добавлением во втором случае Кумасси голубого в состав буфера. Мягкие, неденатурирующие условия в первом направлении нативного 2DE позволяют проводить разделение белковых комплексов и суперкомплексов, не нарушая белок-белковых взаимодействий [123].

Для осуществления идентификации представленные на электрофореграмме в виде пятен белки вырезают, и затем проводят ферментативное расщепление белка в пятне (чаще всего с использованием трипсина) [124]. Полученный гидролизат анализируют при помощи масс-

спектрометрии. Идентификация белков по полученным масс-спектрам пептидных фрагментов проводится с использованием баз данных белковых и нуклеотидных последовательностей [125].

Другой метод протеомного анализа – одномерный электрофорез (1DE) в сочетании с тандемными масс-спектрометрическими методами идентификации белков. Так же, как и в случае 2DE, для изучения белковых комплексов используют либо нативный электрофорез, либо 1DE в присутствии SDS [126, 127]

Разделенные методом одномерного электрофореза белки также подвергают триптическому гидролизу в геле, и затем экстрагированные пептиды анализируют с использованием масс-спектрометрических подходов.

Так как разрешающая способность 1DE невелика и в одной полосе могут присутствовать десятки, а возможно и сотни разных белков, то для успешного разделения триптических пептидов используют обращенно-фазовую жидкостную хроматографию, совмещенную с масс-спектрометрическим детектором (RP-LC-MS), способным проводить анализ первичной структуры пептидов – MS/MS-анализ (ion trap, q-tof FT-ICR и т.д.) [128].

Третий метод, широко используемый в протеомных исследованиях, относится к классу безгелевых (gel-free) технологий и представляет собой многомерное разделение пептидов, полученных при гидролизе гомогенатов тканей или клеток [129].

Данный метод получил широкое распространение вследствие простоты проведения пробоподготовки и высокой степени автоматизации процесса. В некоторых статьях он обозначается как «shotgun» или MudPit (Multidimensional Protein Identification Technology) [130].

В общем случае первым этапом данной технологии является гидролиз белков в смеси, за которым следует первое направление разделения пептидов.

При разделении пептидов возможно использование нескольких методов, которые основаны на различных физико-химических свойствах пептидов. Например, последовательное разделение пептидов сначала на колонке с сильным катионо-обменником (SCX) и последующее разделение фракций пептидов методом RP-LC-MS/MS [131]. Или использование микрокапиллярного изоэлектрофореза в качестве первого направления разделения пептидов согласно их изоэлектрической точке [132]. Последним этапом технологической цепочки обычно является обращенно-фазовая хроматография с регистрацией спектров пептидов и с одновременной фрагментацией выбранных родительских ионов методом столкновительной диссоциации (collision ion dissociation, CID) и/или диссоциации ионов с переносом электронов (electron transfer dissociation, ETD) [133].

ЗАКЛЮЧЕНИЕ

В настоящее время большое разнообразие подходов в белковой интерактомике позволяет осуществлять активное изучение ББВ, дополняя наши знания в области протеомики и системной биологии. Однако надо иметь в виду, что большая часть сведений о ББВ, хранящихся в специализированных базах данных, была получена с помощью искусственных геномных и протеомных подходов и фактически может рассматриваться только в качестве гипотез. Биоинформационный анализ этих данных приводит к генерации новых гипотез, которые без экспериментальной проверки представляют собой гипотезы второго порядка.

Наиболее перспективными направлениями в последние годы стали новые экспериментальные подходы, интегрирующие различные методы аффинной очистки целевых комплексов белков с их последующей идентификацией с помощью масс-спектрометрии и валидации найденных белковых пар методом SPR.

БЛАГОДАРНОСТИ

Данная работа была поддержана Российской академией медицинских наук в рамках программы «Протеомика для медицины», грантами РФФИ № 09-04-12127-офи_м, 10-04-13304-РТ_оми и государственным контрактом № 02.740.11.5006.

СПИСОК ЛИТЕРАТУРЫ

1. Alm E., Arkin A.P. // *Curr. Opin. Struct. Biol.* 2003. V. 13. P. 193–202.
2. Claverie J.M. // *Science*. 2001. V. 291. P. 1255–1257.
3. Cesareni G., Ceol A., Gavrilu C., Palazzi L.M., Persico M., Schneider M.V. // *FEBS Lett.* 2005. V. 579. P. 1828–1833.
4. Pieroni E., de la Fuente van Bentem S., Mancosu G., Capobianco E., Hirt H., de la Fuente A. // *Proteomics*. 2008. V. 8. P. 799–816.
5. Li S., Armstrong C.M., Bertin N., Ge H., Milstein S., Boxem M., Vidalain P.O., Han J.D., et al. // *Science*. 2004. V. 303. P. 540–543.
6. Ghavidel A., Cagney G., Emili A. // *Cell*. 2005. V. 122. P. 830–832.
7. Rual J.F., Venkatesan K., Hao T., Hirozane-Kishikawa T., Dricot A., Li N., Berriz G.F., et al. // *Nature*. 2005. V. 437. P. 1173–1178.
8. Stelzl U., Worm U., Lalowski M., Haenig C., Brembeck F.H., Goehler H., Stroedicke M., et al. // *Cell*. 2005. V. 122. P. 957–968.
9. Kandpal R., Saviola B., Felton J. // *Biotechniques*. 2009. V. 46(5). P. 351–355.
10. Zhang W., Li F., Nie L. // *Microbiology*. 2010. V. 156. P. 287–301.
11. Vidal M., Endoh H. // *Trends Biotechnol.* 1999. V. 17. P. 374–381.
12. Veselovsky A.V., Ivanov Y.D., Ivanov A.S., Archakov A.I., Lewi P., Janssen P. // *J. Mol. Recognit.* 2002. V. 15. P. 405–422.
13. Ivanov A.S., Gnedenko O.V., Molnar A.A., Mezentsev Y.V., Lisitsa A.V., Archakov A.I. // *J. Bioinform. Comput. Biol.* 2007. V. 5. P. 579–592.
14. Fletcher S., Hamilton A.D. // *Curr. Top. Med. Chem.* 2007. V. 7(10). P. 922–927.
15. Veselovsky A.V., Archakov A.I. // *Curr. Comp. Aided Drug Design*. 2007. V. 3. P. 51–58.
16. Ershov P., Gnedenko O., Molnar A., Lisitsa A., Ivanov A., Archakov A. // *Biochemistry (Moscow). Supplemental Series B: Biomedical Chemistry*. 2009. V. 3. P. 272–288.
17. Robinson J.A. // *ChemBioChem*. 2009. V. 10. P. 971–973.
18. Wiles A.M., Doderer M., Ruan J., Gu T.T., Ravi D., Blackman B., Bishop A.J. // *BMC Syst. Biol.* 2010. V. 4. P. 36–52.
19. Walker-Taylor A., Jones D.T. // *Proteomics and Protein–Protein Interactions: Biology, Chemistry, Bioinformatics, and Drug Design* / Ed. G. Waksman. N.Y.: Springer, 2005. P. 89–114.
20. Pitre S., Alamgir M.D., Green J.R., Dumontier M., Dehne F., Golshani A. // *Adv. Biochem. Engin. / Biotechnol.* 2008. V. 110. P. 247–267.
21. Marcotte E.M. // *Curr. Opin. Struct. Biol.* 2000. V. 10. P. 359–365.
22. Eisen M.B., Spellman P.T., Brown P.O., Botstein D. // *Proc. Natl. Acad. Sci. USA*. 1998. V. 95. P. 14863–14868.
23. Dandekar T., Snel B., Huynen M., Bork P. // *Trends Biochem. Sci.* 1998. V. 23. P. 324–328.
24. Lyamouri M., Enerly E., Kress H., Lambertsson A. // *Gene*. 2002. V. 282. P. 199–206.
25. Marcotte E.M., Pellegrini M., Ng H.L., Rice D.W., Yeates T.O., Eisenberg D. // *Science*. 1999. V. 285. P. 751–753.
26. Marcotte E.M., Pellegrini M., Thompson M.J., Yeates T.O., Eisenberg D. // *Nature*. 1999. V. 402. P. 83–86.
27. Enright A.J., Iliopoulos I., Kyrpides N.C., Ouzounis C.A. // *Nature*. 1999. V. 402. P. 83–86.
28. Garcia-Jimenez B., Juan D., Ezkurdia I., Andrés-Leyn E., Valencia A. // *PLoS One*. 2010. V. 5. P. e9969.
29. Pellegrini M., Marcotte E.M., Thompson M.J., Eisenberg D., Yeates T.O. // *Proc. Natl. Acad. Sci. USA*. 1999. V. 96. P. 4285–4288.
30. Fang G., Bhardwaj N., Robilotto R., Gerstein M.B. // *PLoS Comput. Biol.* 2010. V. 6. P. e1000703.

31. Vert J.P. // *Bioinformatics*. 2002. V. 18. P. S276–S284.
32. Pazos F., Valencia A. // *Protein Eng.* 2001. V. 14. P. 609–614.
33. Lu L., Lu H., Skolnick J. // *Proteins*. 2002. V. 49. P. 350–364.
34. Wodak S.J., Mendez R. // *Curr. Opin. Struct. Biol.* 2004. V. 14. P. 242–249.
35. Ogmen U., Keskin O., Aytuna A.S., Nussinov R., Gursoy A. // *Nucleic Acids Res.* 2005. V. 33. P. W331–W336.
36. Aytuna A.S., Gursoy A., Keskin O. // *Bioinformatics*. 2005. V. 21. P. 2850–2855.
37. Aloy P., Russell R.B. // *Proc. Natl. Acad. Sci. USA.* 2002. V. 99. P. 5896–5901.
38. Ma B., Elkayam T., Wolfson H., Nussinov R. // *Proc. Natl. Acad. Sci. USA.* 2003. V. 100. P. 5772–5777.
39. Aloy P., Russell R.B. // *Bioinformatics*. 2003. V. 19. P. 161–162.
40. Berman H., Henrick K., Nakamura H., Markley J.L. // *Nucleic Acids Res.* 2007. V. 35. P. 301–303.
41. Sonnhammer E.L., Eddy S.R., Durbin R. // *Proteins*. 1997. V. 28. P. 405–420.
42. Finn R., Griffiths-Jones S., Bateman A. // *Curr. Protoc. Bioinformatics*. 2003. Chapter 2. Unit 2.5.
43. Espadaler J., Romero-Isart O., Jackson R.M., Oliva B. // *Bioinformatics*. 2005. V. 21. P. 3360–3368.
44. Espadaler J., Eswar N., Querol E., Avilüs F.X., Sali A., Marti-Renom M.A., Oliva B. // *BMC Bioinformatics*. 2008. V. 9. P. 249–257.
45. Deng M., Mehta S., Sun F., Chen T. // *Genome Res.* 2002. V. 12. P. 1540–1548.
46. Han D.S., Kim H.S., Jang W.H., Lee S.D., Suh J.K. // *Nucleic Acids Res.* 2004. V. 32. P. 6312–6320.
47. Han D.S., Kim H.S., Jang W.H., Lee S.D., Suh J.K. // *Genome Inform.* 2004. V. 15. P. 171–180.
48. Pitre S., Dehne F., Chan A., Cheetham J., Duong A., Emili A., Gebbia M., Greenblatt J., Jessulat M., Krogan N., Luo X., Golshani A. // *BMC Bioinformatics*. 2006. V. 7. P. 365–380.
49. Ta H.X., Holm L. // *Biochem. Biophys. Res. Commun.* 2009. V. 390. P. 357–362.
50. Kim W.K., Park J., Suh J.K. // *Genome Inform.* 2002. V. 13. P. 42–50.
51. Pattin K.A., Moore J.H. // *Expert Review of Proteomics*. 2009. V. 6. P. 647–659.
52. Sprinzak E., Margalit H. // *J. Mol. Biol.* 2001. V. 311. P. 681–692.
53. Tan S.H., Hugo W., Sung W.K., Ng S.K. // *BMC Bioinformatics*. 2006. V. 7. P. 502.
54. Mulder N.J., Kersey P., Pruess M., Apweiler R. // *Mol. Biotechnol.* 2008. V. 38. P. 165–177.
55. Bock J.R., Gough D.A. // *Bioinformatics*. 2001. V. 17. P. 455–460.
56. Englbrecht C.C., Facius A. // *Comb. Chem. High Throughput Screen.* 2005. V. 8. P. 705–715.
57. Tang Z.Q., Lin H.H., Zhang H.L., Han L.Y., Chen X., Chen Y.Z. // *Bioinform. Biol. Insights*. 2009. V. 24. P. 19–47.
58. Martin S., Roe D., Faulon J.L. // *Bioinformatics*. 2005. V. 21. P. 218–226.
59. Ben-Hur A., Noble W.S. // *Bioinformatics*. 2005. V. 21. P. i38–46.
60. Qiu J., Noble W.S. // *PLoS Comput. Biol.* 2008. V. 4. P. e1000054.
61. Fields S., Song O. // *Nature*. 1989. V. 340. P. 245–246.
62. Xenarios I., Fernandez E., Salwinski L., Duan X.J., Thompson M.J., Marcotte E.M., Eisenberg D. // *Nucleic Acids Res.* 2001. V. 29. P. 239–241.
63. Mukherjee S., Bal S., Saha P. // *Cur. Sci.* 2001. V. 81. P. 458–464.
64. Hermjakob H., Montecchi-Palazzi L., Lewington C., Mudali S., Kerrien S., Orchard S., Vingron M., Roechert B., Roepstorff P., Valencia A., Margalit H., Armstrong J., Bairoch A., Cesareni G., Sherman D., Apweiler R. // *Nucleic Acids Res.* 2004. V. 32. Database issue. P. D452–D455.
65. Xenarios I., Rice D.W., Salwinski L., Baron M.K., Marcotte E.M., Eisenberg D. // *Nucleic Acids Res.* 2000. V. 28. P. 289–291.
66. Xenarios I., Salwinski L., Duan X.J., Higney P., Kim S.M., Eisenberg D. // *Nucleic Acids Res.* 2002. V. 30. P. 303–305.

67. Salwinski L., Miller C.S., Smith A.J., Pettit F.K., Bowie J.U., Eisenberg D. // *Nucleic Acids Res.* 2004. V. 32. Database issue. P. D449–D451.
68. Grigoriev A. // *Nucleic Acids Res.* 2003. V. 31. P. 4157–4161.
69. Uetz P., Giot L., Cagney G., Mansfield T.A., Judson R.S., Knight J.R., Lockshon D., Narayan V., Srinivasan M., Pochart P., Qureshi-Emili A., Li Y., Godwin B., Conover D., Kalbfleisch T., Vijayadamodar G., Yang M., Johnston M., Fields S., Rothberg J.M. // *Nature.* 2000. V. 403. P. 623–627.
70. Ito T., Chiba T., Ozawa R., Yoshida M., Hattori M., Sakaki Y. // *Proc. Natl. Acad. Sci. USA.* 2001. V. 98. P. 4569–4574.
71. Wiles A.M., Doderer M., Ruan J., Gu T.T., Ravi D., Blackman B., Bishop A.J. // *BMC Syst. Biol.* 2010. V. 4. P. 36–51.
72. Uetz P. // *Curr. Opin. Chem. Biol.* 2002. V. 6. P. 57–62.
73. Rajagopala S.V., Uetz P. // *Meth. Mol. Biol.* 2009. V. 548. P. 223–245.
74. von Mering C., Krause R., Snel B., Cornell M., Oliver S.G., Fields S., Bork P. // *Nature.* 2002. V. 417. P. 399–403.
75. Gandhi T.K., Zhong J., Mathivanan S., Karthick L., Chandrika K.N., Mohan S.S., Sharma S., Pinkert S., Nagaraju S., Periaswamy B., Mishra G., Nandakumar K., Shen B., Deshpande N., Nayak R., Sarker M., Boeke J.D., Parmigiani G., Schultz J., Bader J.S., Pandey A. // *Nat. Genet.* 2006. V. 38. P. 285–293.
76. Suzuki H., Fukunishi Y., Kagawa I., Saito R., Oda H., Endo T., Kondo S., Bono H., Okazaki Y., Hayashizaki Y. // *Genome Res.* 2001. V. 11. P. 1758–1765.
77. Tong A.H., Evangelista M., Parsons A.B., Xu H., Bader G.D., Page N., Robinson M., Raghibizadeh S., Hogue C.W., Bussey H., Andrews B., Tyers M., Boone C. // *Science.* 2001. V. 294. P. 2364–2368.
78. Tong A.H., Lesage G., Bader G.D., Ding H., Xu H., Xin X., Young J., et al. // *Science.* 2004. V. 303. P. 808–813.
79. Tong A.H., Boone C. // *Meth. Mol. Biol.* 2006. V. 313. P. 171–191.
80. Ge H., Liu Z., Church G.M., Vidal M. // *Nat. Genet.* 2001. V. 29. P. 482–486.
81. Park H.R., Cockrell L.M., Du Y., Kasinski A., Havel J., Zhao J., Reyes-Turcu F., Wilkinson K.D., Fu H. // *Protein–Protein Interactions. Molecular Biomethods Handbook*, 2nd Ed. / Eds J.M. Walker, R. Rapley. N. J. Totowa: Humana Press, 2008. P. 463–494.
82. Phizicky E.M., Fields S. // *Microbiol. Rev.* 1995. V. 59. P. 94–123.
83. Masters S.C. // *Methods Mol. Biol.* 2004. V. 261. P. 337–350.
84. Yaciuk P. // *Methods. Mol. Med.* 2007. V. 131. P. 103–111.
85. Free R.B., Hazelwood L.A., Sibley D.R. // *Curr. Protoc. Neurosci.* 2009. Chapter 5. Unit 5.28.
86. Chatton B., Bahr A., Acker J., Kedinger C. // *Biotechniques.* 1995. V. 18. P. 142–145.
87. Forde G.M. // *Methods Mol. Biol.* 2008. V. 421. P. 125–136.
88. Viljanen J., Larsson J., Broo K.S. // *Protein Expr. Purif.* 2008. V. 57. P. 17–26.
89. Jung J.W., Jung S.H., Kim H.S., Yuk J.S., Park J.B., Kim Y.M., Han J.A., Kim P.H., Ha K.S. // *Proteomics.* 2006. V. 6. P. 1110–1120.
90. Schmitt J., Hess H., Stunnenberg H.G. // *Molecular Biology Reports.* 1993. V. 18. P. 223–230.
91. Crowe J., Dübali H., Gentz R., Hochuli E., Støber D., Henco K. // *Methods Mol. Biol.* 1994. V. 31. P. 371–387.
92. Scheich C., Sievert V., Bøssow K. // *BMC Biotechnology.* 2003. V. 3. P. 12.
93. Suter B., Fetchko M.J., Imhof R., Graham C.I., Stoffel-Studer I., Zbinden C., Raghavan M., Lopez L., Beneti L., Hort J., Fillingham J., Greenblatt J.F., Giaever G., Nislow C., Stagljar I. // *Genome Res.* 2007. V. 17. P. 1774–1782.
94. Shen A., Lupardus P.J., Morell M., Ponder E.L., Sadaghiani A.M., Garcia K.C., Bogoy M. // *PLoS*

ONE. 2009. V. 4. P. e8119.

95. *Debeljak N., Feldman L., Davis K.L., Komel R., Sytkowski A.J.* // *Anal. Biochem.* 2006. V. 359. P. 216–223.
96. *Brizzard B.* // *BioTechniques.* 2008. V. 44. P. 693–695.
97. *Rigaut G., Chevchenko A., Rutz B., Wilm M., Mann M., Seraphin B.* // *Nat. Biotechnol.* 1999. V. 17. P. 1030–1032.
98. *Puig O., Caspary F., Rigaut G., Rutz B., Bouveret E., Bragado-Nilsson E., Wilm M., Seraphin B.* // *Methods.* 2001. V. 24. P. 218–229.
99. *Gould K.L., Ren L., Feoktistova A.S., Jennings J.L., Link A.J.* // *Methods.* 2004. V. 33. P. 239–244.
100. *Burckstummer T., Bennett K., Preradovic A., Schutze G., Hantschel O., Superti-Furga G., Bauch A.* // *Nat. Methods.* 2006. V. 3. P. 1013–1019.
101. *Knuesel M., Wan Y., Xiao Z., Holinger E., Lowe N., Wang W., Liu X.* // *Mol. Cell. Proteomics.* 2003. V. 2. P. 1225–1233.
102. *Koch K.V., Reinders Y., Ho T.-H., Sickmann A., Graf R.* // *Eur. J. Cell Biol.* 2006. V. 85. P. 1079–1090.
103. *Rich R.L., Myszka D.G.* // *Curr. Opin. Biotechnol.* 2000. V. 11. P. 54–61.
104. *Surface Plasmon Resonance. Methods in Molecular Biology.* 2010. V. 627. Springer Science+Business Media.
105. *Nelson R.W., Jarvik J.W., Taillon B.E., Tubbs K.A.* // *Anal. Chem.* 1999. V. 71. P. 2858–2865.
106. *Nedelkov D., Nelson R.W.* // *Proteomics.* 2001. V. 1. P. 1441–1446.
107. *Nedelkov D., Nelson R.W.* // *J. Mol. Recognit.* 2003. V. 16. P. 15–19.
108. *Zhukov A., Schørenberg M., Jansson I., Areskoug D., Buijsa J.* // *J. Biomol. Techniques.* 2004. V. 15. P. 112–119.
109. *Grote J., Dankbar N., Gedig E., Koenig S.* // *Anal. Chem.* 2005. V. 77. P. 1157–1162.
110. *Larsericdotter H., Jansson O., Zhukov A., Areskoug D., Oscarsson S., Buijs J.* // *Proteomics.* 2006. V. 6. P. 2355–2364.
111. *Bouffartigues E., Leh H., Anger-Leroy M., Rimsky S., Buckle M.* // *Nucleic Acids Res.* 2007. V. 35. P. e39.
112. *Visser N.F.C., Heck A.J.R.* // *Expert Rev. Proteomics.* 2008. V. 5. P. 425–433.
113. *Borch J., Roepstorff P.* // *Mol. Biotechnol.* 2006. V. 33. P. 179–190.
114. *Nedelkov D., Nelson R.W.* // *J. Mol. Recognit.* 2003. V. 16. P. 9–14.
115. *Berggerd T., Linse S., James P.* // *Proteomics.* 2007. V. 7. P. 2833–2842.
116. *Jecklin M.C., Schauer S., Dumelin C.E., Zenobi R.* // *J. Mol. Recognit.* 2009. V. 22. P. 319–329.
117. *Buneeva O., Gnedenko O., Zgoda V., Kopylov A., Glover V., Ivanov A., Medvedev A., Archakov A.* // *Proteomics.* 2010. V. 10. P. 23–37.
118. *Humphery-Smith I., Cordwell S.J., Blackstock W.P.* // *Electrophoresis.* 1997. V. 18. P. 1217–1242.
119. *O'Farrell P.H.* // *J. Biol. Chem.* 1975. V. 250. P. 4007–4021.
120. *Klose J., Kobalz U.* // *Electrophoresis.* 1995. V. 16. P. 1034–1059.
121. *Miller I., Crawford J., Gianazza E.* // *Proteomics.* 2006. V. 6. P. 5385–5408.
122. *Archakov A., Ivanov Y., Lisitsa A., Zgoda V.* // *Proteomics.* 2009. V. 9. P. 1326–1343.
123. *Wittig I., Schagger H.* // *Proteomics.* 2009. V. 9. P. 5214–5223.
124. *Shevchenko A., Wilm M., Vorm O., Mann M.* // *Anal. Chem.* 1996. V. 68. P. 850–858.
125. *Fenyo D.* // *Curr. Op. Biotech.* 2000. V. 11. P. 391–395.
126. *Wittig I., Schøgger H.* // *Proteomics.* 2008. V. 8. P. 3974–3990.
127. *Filatov A.V., Krotov G.I., Zgoda V.G., Volkov Y.* // *J. Immun. Meth.* 2007. V. 319. P. 21–33.
128. *Washburn M.P., Wolters D., Yates J.R., III.* // *Nat. Biotechnol.* 2001. V. 19. P. 242–247.

129. Zgoda V.G., Moshkovskii S.A., Ponomarenko E.A., Andreewski T.V., Kopylov A.T., Tikhonova O.V., Melnik S.A., Lisitsa A.V., Archakov A.I. // *Proteomics*. 2009. V. 9. P. 4102–4105.
130. Gonzalez-Begne M., Lu B., Han X., Hagen F.K., Hand A.R., Melvin J.E., Yates J.R. // *J. Proteome Res.* 2009. V. 8. P. 1304–1314.
131. Washburn M.P., Wolters D., Yates J.R., III. // *Nat. Biotechnol.* 2001. V. 19. P. 242–247.
132. Wang W., Guo T., Song T., Lee C.S., Balgley B.M. // *Proteomics*. 2007. V. 7. P. 1178–1187.
133. Boersema P.J., Mohammed S., Heck A.J. // *J. Mass. Spectrom.* 2009. V. 44. P. 861–878.

Таблица. Подходы в современной интерактомике

1. Биоинформационные подходы:
<i>1.1. Подходы, основанные на геномной информации</i>
<ul style="list-style-type: none"> • Анализ консерватизма расположения генов в геномах • Сравнительный анализ белковых пар и гомологичных белков слияния в геномах
<i>1.2. Подходы, основанные на эволюционных взаимосвязях белков</i>
<ul style="list-style-type: none"> • Корреляционная эволюция функционально связанных белков • Оценка подобия филогенетических профилей
<i>1.3. Подходы, основанные на трехмерной структуре белка</i>
<ul style="list-style-type: none"> • Оценка пространственного совмещения двух взаимодействующих белков на основе известной трехмерной структуры комплекса белков-гомологов • Алгоритм многомерного тридинга (MULTIPROSPECTOR) для распознавания партнеров ББВ • Широкомасштабный эксперимент CAPRI по процедурам белок-белкового докинга • PRISM – интернет-сервис для анализа контактных интерфейсов белковых комплексов и предсказания возможных ББВ на основе структурных соответствий
<i>1.4. Подходы, основанные на белковых доменах</i>
<ul style="list-style-type: none"> • Комбинирование анализа подобия участков последовательностей, вовлеченных во взаимодействие, с подобием доменов взаимодействующих партнеров • Определение вероятности взаимодействия между эволюционно консервативными белковыми доменами (в базе данных белковых доменов Pfam) • Предсказание вероятности взаимодействия белков • База данных потенциально взаимодействующих пар доменов (PID)
<i>1.5. Подходы, основанные на первичной структуре белков</i>
<ul style="list-style-type: none"> • ББВ, обусловленные короткими специфическими последовательностями • Автоматическое предсказание ББВ на основе анализа первичных структур и физико-химических свойств методом опорных векторов (support vector machine) • Комбинирование данных о первичных структурах, экспериментальных данных и обрезания последовательности для использования в методе опорных векторов • Ядерные методы (kernel methods) предсказания ББВ • PIPE: сервер предсказания ББВ, использующий данные о первичной структуре белков из баз данных MIPS и DIP
2. Экспериментальные подходы
<i>2.1. Геномные подходы</i>
<ul style="list-style-type: none"> • Дрожжевая двугибридная система (yeast two hybrid system) • Синтетические генетические матрицы (synthetic genetic array (SGA) analysis) • Корреляционный профиль экспрессии мРНК (correlated mRNA expression profile)
<i>2.2. Биохимические подходы</i>
<ul style="list-style-type: none"> • Методы на основе твердофазной аффинной хроматографии • Молекулярный фишинг на чипе оптического биосенсора • Масс-спектрометрическая идентификация белков

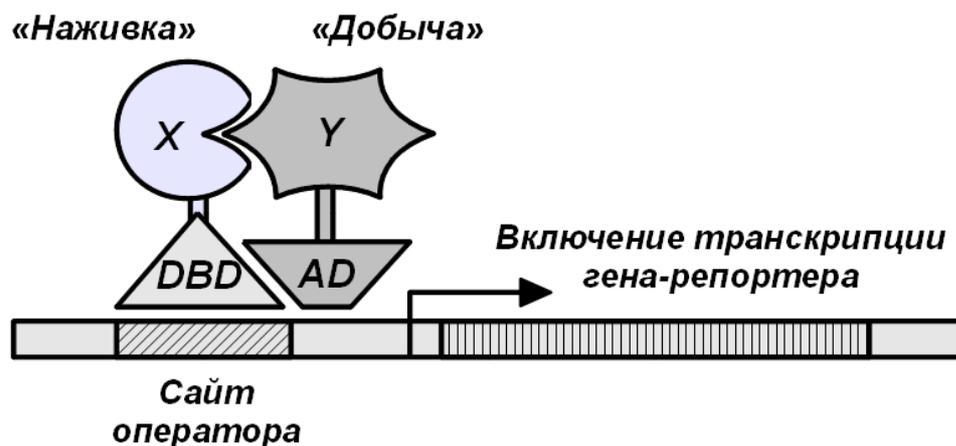


Рис. 1. Принцип дрожжевой двугибридной системы [61, 62]. Целевой белок X (белок-«наживка») экспрессируется в виде белка слияния с DBD-доменом (ДНК-связывающего домена фактора транскрипции). Этот белок слияния DBD-X связывается с сайтом оператора в области промотора, но не активирует транскрипцию гена-репортера из-за отсутствия домена AD (активационного домена того же фактора транскрипции). Взаимодействие белка X с белком-партнером Y (белок-«добыча»), который экспрессируется также в виде белка слияния с доменом AD, приводит к формированию полного фактора транскрипции и включению транскрипции гена-репортера.

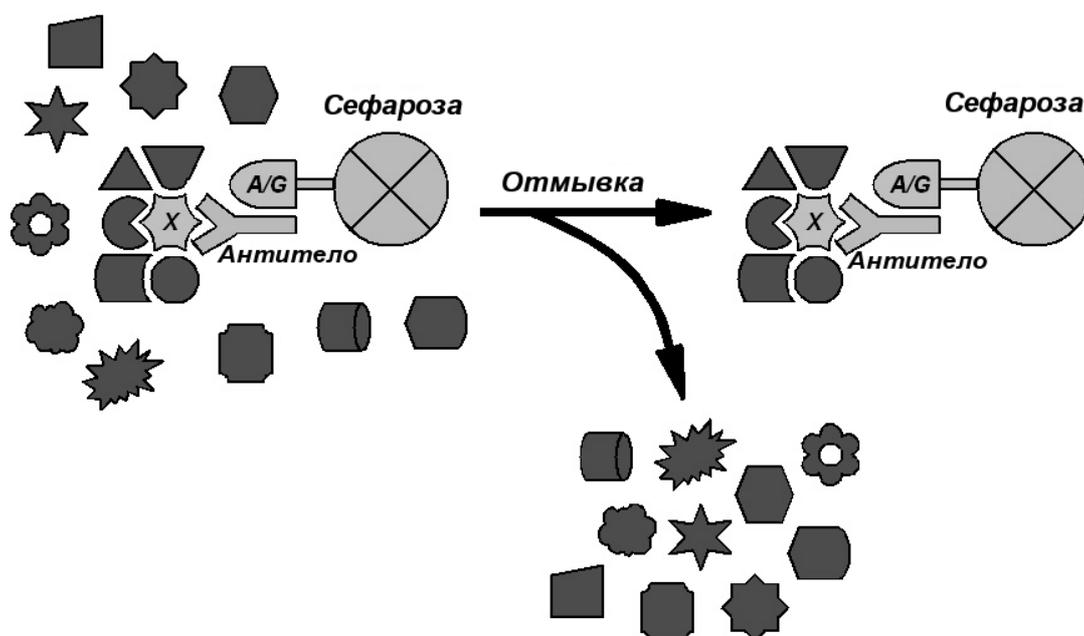


Рис. 2. Принцип выделения белковых комплексов методом коиммунопреципитации. X – целевой белок в составе комплекса с белками-партнерами. В клеточный лизат добавляются антитела к белку X. Образовавшийся комплекс антитело/целевой белок/белки-партнеры выделяется из смеси с помощью аффинного реагента A/G (белок А или белок G), конъюгированного с сефарозой.

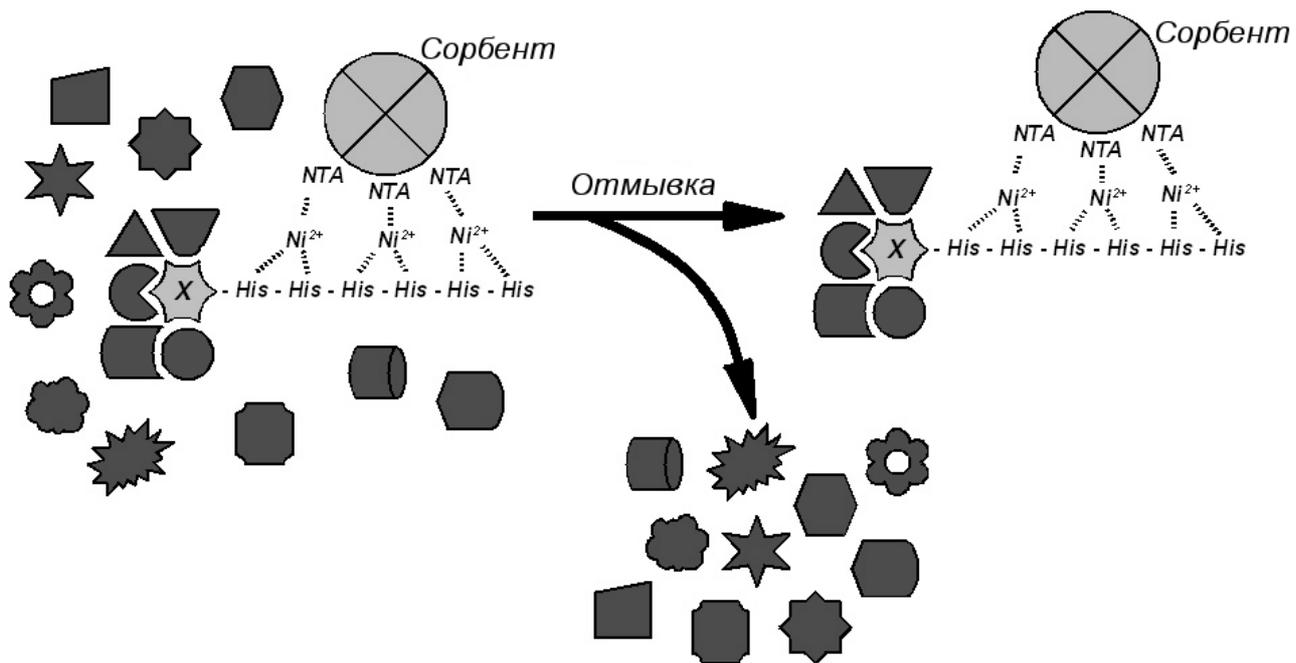


Рис. 3. Принцип аффинного выделения 6xHis-меченого белка в комплексе с белками-партнерами. X – целевой белок, который находится в комплексе с белками-партнерами; NTA – нитрилотриуксусная кислота. His – остаток гистидина.

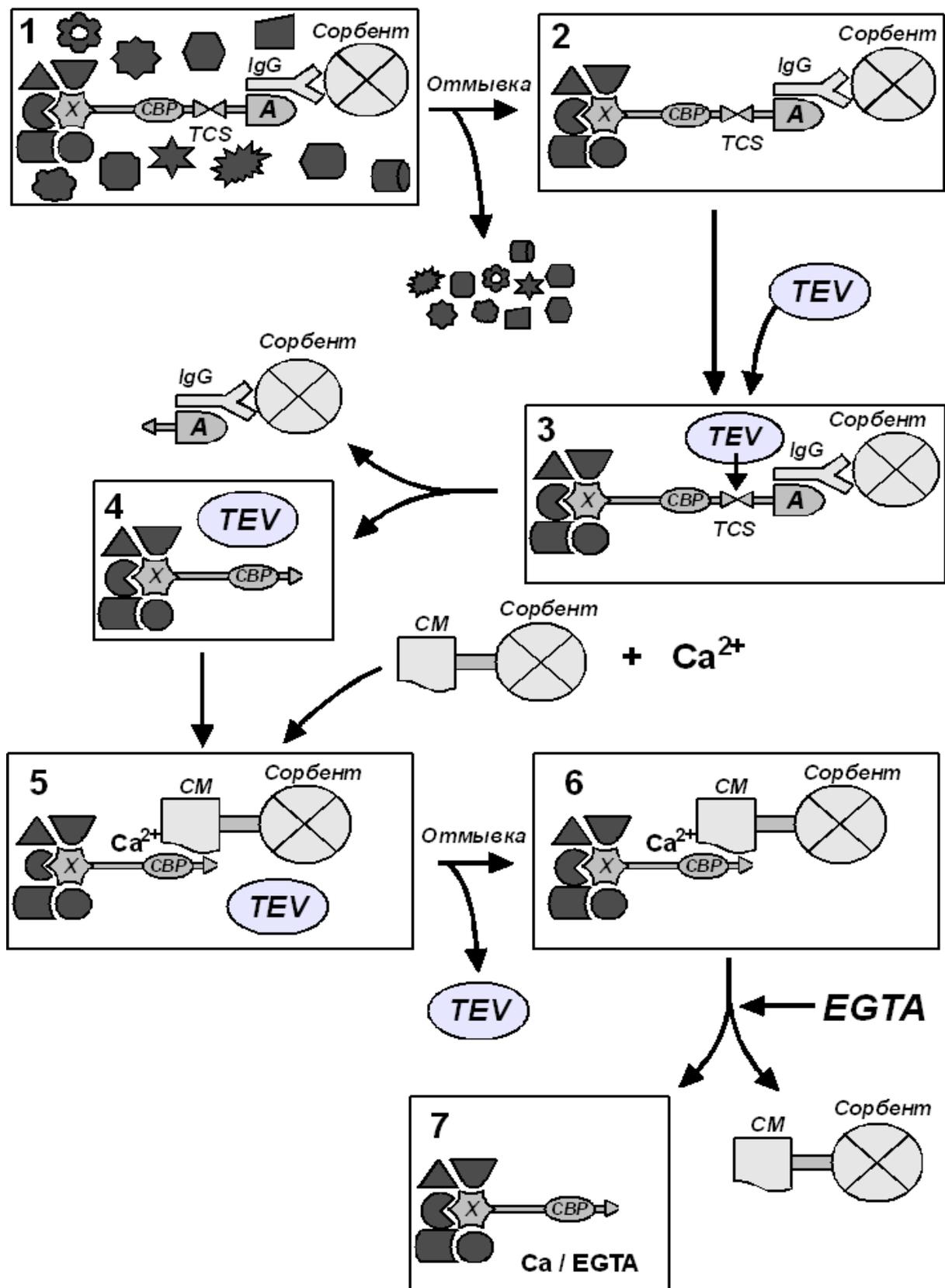


Рис. 4. Принцип тандемной аффинной очистки целевого белка в комплексе с белками-партнерами (Tandem Affinity Purification, TAP). X-CBP-TCS-A – белок слияния, состоящий из целевого белка X, кальмодулинсвязывающего пептида (CBP), специфического сайта (TCS) разрезания белковой цепи протеазой TEV (вируса гравировки табака) и белка А; IgG – иммуноглобулин G; СМ – кальмодулин; EGTA – комплексон (этиленгликоль тетраацетат).

Protein Interactomics Technologies

A. S. Ivanov[#], V. G. Zgoda, A. I. Archakov

[#] *Phone: +7 (499) 246-36-93; e-mail: alexei.ivanov@ibmc.msk.ru*

Institute of Biomedical Chemistry RAMS, Pogodinskaya str., 10, 119121, Moscow

The analysis of protein-protein interactions has the growing importance in modern proteomics. It is caused by necessity of in-depth functional study of proteins which, as a rule, operate in living systems due to each other interaction in the stable or dynamic protein complexes. In this review, the brief description and a comparative evaluation of most frequently used approaches in existing variety of protein interactomics methods are given.

Key words: *interactomics; protein-protein interactions; bioinformatics; proteomics; genome technologies; molecular fishing; optical biosensor; protein mass-spectrometry.*