

ДИНАМИЧЕСКАЯ ПРОТЕОМИКА В МОДЕЛИРОВАНИИ ЖИВОЙ КЛЕТКИ. БЕЛОК-БЕЛКОВЫЕ ВЗАИМОДЕЙСТВИЯ

© 2009 г. А. А. ТЕРЕНТЬЕВ¹,
Н. Т. МОЛДОГАЗИЕВА¹, К. В. ШАЙТАН²

¹Российский государственный медицинский университет, Москва

²Биологический факультет Московского государственного
университета им. М.В.Ломоносова

I. Введение. II. Картирование белок-белковых взаимодействий.
III. Белковые сети. IV. Структурно-функциональные модули бел-
ков, ответственные за их взаимодействия. V. Методы динамичес-
кого моделирования *in silico*. VI. Заключение.

I. ВВЕДЕНИЕ

Успехи современной протеомтики, разработка новых математических и компьютерных подходов и потребности фундаментальной медицины поставили на повестку дня весьма сложную и интригующую задачу – моделирование живой клетки, решение которой в настоящее

Принятые сокращения: МД – молекулярная динамика; Y2H (yeast two-hybrid assay) – дрожжевой двугибридный анализ; TAP-MS (tandem affinity purification-mass spectrometry) – аффинная хроматография в сочетании с масс-спектрометрией; MALDI-TOF (matrix-assisted laser desorption/ionization–time of flight) – времяпролетная матрично-активированная лазерная десорбция/ионизация; SELDI (surface enhanced laser desorption/ionization) – активированная поверхностью лазерная десорбция/ионизация; FRET (fluorescence resonance energy transfer) – резонансный перенос энергии электронного возбуждения; FLIM (fluorescence lifetime imaging) – изображение по времени жизни флуоресценции; МАПК – митогенаактивируемая протеинкиназа; TNF- α (tumor necrosis factor- α) – фактор некроза опухолей- α ; NF- κ B (nuclear factor- κ B) – ядерный фактор- κ B; CDK (cyclin-dependent kinase) – циклинзависимая киназа; CKI (cyclin-dependent kinase inhibitor) – ингибитор циклинзависимой киназы; АФП – альфа-фетопrotein; ЭФР – эпидермальный фактор роста; РТВ (phosphotyrosine-binding) – фосфотирозинсвязывающий; САТ – серинацетилтрансфераза; ОАСС – О-ацетилсеринсульфогидролаза; SH (Src homology) – гомолог продукта онкогена вируса куриной саркомы.

Адреса для корреспонденции: aaterent@mtu-net.ru, nmoldogazieva@mail.ru, shaitan@moldyn.org

Работа выполнена при финансовой поддержке Российского гуманитарного научного фонда (грант № 09-06-00241а).

время становится возможным с помощью достижений современных теоретических и экспериментальных методов. Некоторыми авторами уже предприняты попытки создания статической модели клетки, на основе знаний не только о строении клетки и внутриклеточных органелл, их биохимического состава, но и о локализации и приблизительном количестве молекул всех, существующих в клетке, низко- и высокомолекулярных соединений [1]. Однако моделирование живой, «работающей», клетки, т.е. создание её динамической модели, учитывающей изменения во времени и в пространстве химического состава клетки, а также особенности протекания всех внутри- и межклеточных биохимических процессов, является значительно более сложной задачей [2]. Еще более сложным представляется создание динамических моделей разных типов клеток, а также клеток, находящихся в различных физиологических и патофизиологических состояниях, в зависимости от микроокружения или внешних сигналов. Особый интерес представляет моделирование клетки, находящейся на разных стадиях жизненного цикла, включая её деление, дифференцировку и гибель.

Решение таких сложных задач требует сбора большого количества экспериментальных и теоретических данных и учета двух важных и взаимосвязанных, на наш взгляд, аспектов. С одной стороны, необходим системный подход, интегрированный взгляд на процессы, происходящие в клетке и/или в её отдельных компартментах. Рассмотрение клетки как единого целого является предметом исследования относительно новой междисциплинарной области науки – системной биологии [3–6]. Она обеспечивает интеграцию знаний, полученных на различных уровнях, от молекулярного до тканевого или организменного, и с помощью разных экспериментальных и теоретических методов. Целью системной биологии является выяснение того, каким образом совместное функционирование различных составляющих клетки или ткани обуславливает нормальное протекание биологических и физиологических процессов в организме.

Изучение клетки с точки зрения системной биологии предполагает наличие у клеточных составляющих приобретенных, так называемых производных, свойств или функций (emerging properties). Это означает, что те или иные функции становятся возможными только при достижении определенного уровня сложности организации системы. При этом каждая из составляющих в отдельности может не обладать свойствами (и функциями), которые приобретает система из двух составляющих. А система из двух составляющих может не обладать свойствами и функциями более сложно устроенных систем. Такая

интеграция предполагает рассмотрение клетки в широком диапазоне временных и пространственных масштабов. А это требует знаний о детальных качественных и количественных параметрах изменений на всех уровнях, включая межмолекулярные взаимодействия, что, в свою очередь, дает представление о целостных процессах, проистекающих на уровне всей клетки.

Современные экспериментальные методы предоставляют возможности осуществлять исследования либо на уровне целой клетки, как, например, методы микроскопии, либо на уровне отдельных молекул. Рассмотрение всех клеточных процессов в совокупности, с одновременным изучением детальных молекулярных механизмов того или иного процесса, пока не представляется возможным. С применением математических и компьютерных методов моделирования становится возможным изучения процессов и явлений, исследование которых с помощью даже высокоэффективных экспериментальных методов оказывается затруднительным [7, 8]. Если математические методы основаны на описании и анализе внутри- и межклеточных процессов и явлений с помощью системы математических уравнений, то компьютерные методы, в основном, имеют целью создание алгоритмов для имитации биологических процессов, их конструирования и визуализации. Для динамического моделирования детальных механизмов межмолекулярных взаимодействий и биохимических процессов, происходящих в клетке, все большие возможности предоставляют различные варианты метода молекулярной динамики (МД) [9].

Для моделирования, описания и анализа многих реально существующих процессов, происходящих как в биологических, так и в небиологических системах, используют различные типы сетей, образованных группами взаимодействующих компонентов. Среди биологических сетей особое место для изучения функционирования клетки занимают молекулярные сети, к которым относятся геновые, белковые, метаболические и сигнальные сети. По своим свойствам и принципу организации все эти типы молекулярных сетей относятся к комплексным сетям. Отражая сложность организации биологических систем, они являются объектом исследования системной биологии [10]. Анализ молекулярных сетей позволяет выявлять функциональные модули в них и выяснять роль каждого из компонентов сети в функционировании клетки. Группы физически взаимодействующих белков, которые функционируют в клетке совместно и скоординированно, контролируя взаимосвязанные процессы, происходящие в организме, образуют *белковые сети*. Белки, как

ключевые биомакромолекулы, являются основными участниками почти всех клеточных процессов. Поэтому моделирование живой клетки невозможно без всей совокупности данных динамической протеомики, которая изучает изменения концентраций и локализации белков и их взаимодействия друг с другом [11–13].

Нарушение белок-белковых взаимодействий может приводить к возникновению различных заболеваний, включая опухолевые, нейро-дегенеративные, сердечно-сосудистые, аутоиммунные и др. Поэтому изучение взаимодействующих партнеров и анализ белковых сетей, образованных белок-белковыми взаимодействиями, представляет собой важный инструмент в диагностике заболеваний, выяснении механизма их возникновения и развития, а также эффективности тех или иных терапевтических подходов [14, 15].

В настоящее время становится признанным, что большинство белков эукариот являются мультимодульными и полифункциональными [16, 17]. Каждый из модулей может выполнять самостоятельную функцию, в результате чего белок приобретает способность выполнять целый комплекс различных функций. В связи с мультимодульностью и полифункциональностью большинства белков эукариот их сложные белковые сети могут переплетаться друг с другом. Потому важным инструментом в моделировании живой клетки является структурно-функциональное картирование белков, позволяющее локализовать их функционально важные участки, в том числе и те, благодаря которым осуществляются белок-белковые взаимодействия.

Настоящий обзор посвящен описанию, обобщению и анализу данных по белок-белковым взаимодействиям и белковым сетям, полученных с помощью современных экспериментальных и теоретических методов. Описаны принципы, лежащие в основе современных методов изучения белок-белковых взаимодействий. При этом проведен анализ преимуществ и недостатков тех или иных экспериментальных методов и описаны современные подходы, предпринимаемые разными группами авторов для решения возникающих проблем. Показана роль анализа белковых сетей в решении задач фундаментальной медицины. Рассмотрены методы моделирования *in silico* и проведен анализ возможностей МД методов в создании динамической модели клетки.

II. КАРТИРОВАНИЕ БЕЛОК-БЕЛКОВЫХ ВЗАИМОДЕЙСТВИЙ

ИНТЕРАКТОМЫ

Всё множество белков в клетке находится в динамическом взаимодействии друг с другом, и эти взаимодействия обуславливают функционирование и поведение живой клетки. Обратимые белок-белковые взаимодействия представляет собой один из видов динамических процессов, происходящих в клетке и позволяющих судить о её функционировании. Интенсивные исследования, осуществляемые в этой области в течение последних лет, способствовали постепенному накоплению данных о взаимодействующих парах белков и об образованных ими белковых комплексах. В течение последних двух десятилетий был разработан ряд высокоэффективных и высокопроизводительных экспериментальных методов для изучения белок-белковых взаимодействий, которые, однако, не лишены недостатков. Поэтому экспериментальные методы дополняются различными теоретическими подходами, включая математические и компьютерные методы исследований. При этом белок-белковые взаимодействия изучаются отдельно для разных биологических видов, а для высших эукариотических организмов – также отдельно для разных тканей и типов клеток.

Вся совокупность белок-белковых взаимодействий, характерная для данного организма, получила название *интерактома*. Структурная организация интерактомов и общее количество взаимодействий в них являются одними из важных факторов, определяющих сложность биологических систем. Количество копий тех или иных белков, приходящихся на одну клетку, может колебаться от нескольких десятков (менее 50) до миллионов [18]. Поэтому интерактомы даже простых организмов могут быть образованы довольно большим количеством взаимодействий. Так, например, размер интерактома у дрожжей *S. cerevisiae* может составлять, по одним данным, до 10–17 тысяч, а по другим – до 25–35 тысяч взаимодействий [19]. Статистическая оценка возможного размера интерактома человека показала, что он может быть образован примерно 650 тысячами взаимодействий. При этом размер интерактома человека оказался приблизительно в 10 раз больше, чем у *D. melanogaster*, и может в 3 раза превышать размер интерактома *C. elegans* [20]. Эти данные позволили сделать предположение о том, что размер интерактома коррелирует с уровнем сложности организации того или иного биологического вида.

Определение физически взаимодействующих пар белков позволяет составлять интерактомные карты, которые представляют собой

графы, состоящие из узлов, в которых расположен тот или иной белок, и связей между ними, показывающих парные взаимодействия (рис. 1). Интерактомные карты рассматриваются в качестве ключей к получению знаний о функционировании белков в клетках [21]. На основании данных, полученных методами *in vitro*, составляются статические интерактомные карты, анализ которых, как будет показано далее, позволяет описывать динамические белок-белковые взаимодействия, существующие в условиях *in vivo*. Составление интерактомных карт полезно также и для фундаментальной медицины, а именно для определения роли отдельных белков и их взаимодействий в возникновении и развитии заболеваний, их диагностики, а также возможных мишеней для действия различных лекарств и контролирования эффективности лечения.

Известные в настоящее время интерактомные карты довольно неполны даже для простейших организмов. Кроме того, данные, полученные разными группами авторов, часто противоречивы. Взаимодействующие пары и белковые комплексы, идентифицированные одной группой исследователей, не обнаруживаются другими авторами. Тем не менее, разработка новых экспериментальных и теоретических подходов, которые будут обсуждаться ниже, приводит к постепенному накоплению данных, необходимых для анализа белок-белковых взаимодействий, существующих в клетке.

Экспериментальными методами в настоящее время определяются взаимодействующие пары белков и белковые комплексы, в основном, у прокариот или простых эукариотических организмов. Для выявления белок-белковых взаимодействий у более сложных организмов, таких как млекопитающие, дополнительно используется метод предсказания их на основании гомологии с белками, взаимодействующие партнеры для которых были выявлены у более простых организмов [22]. Данный подход основан на существовании гомологии между родственными белками и сравнении степени консервативности первичной и пространственной структуры одного и того же белка у разных биологических видов. Например, если экспериментально показано, что два каких-либо белка у дрожжей взаимодействуют друг с другом, то предполагается, что у человека эти белки также взаимодействуют. Две пары белков у разных организмов, эволюционно сохранивших способность взаимодействовать друг с другом, получили название интерологов. Предсказание взаимодействующих пар белков дает более или менее надежные результаты лишь при условии высокой степени сходства их первичных последовательностей. Кроме того, как показывают исследования последних лет, данные, полученные с помощью самых высокоэффективных экспериментальных методов,

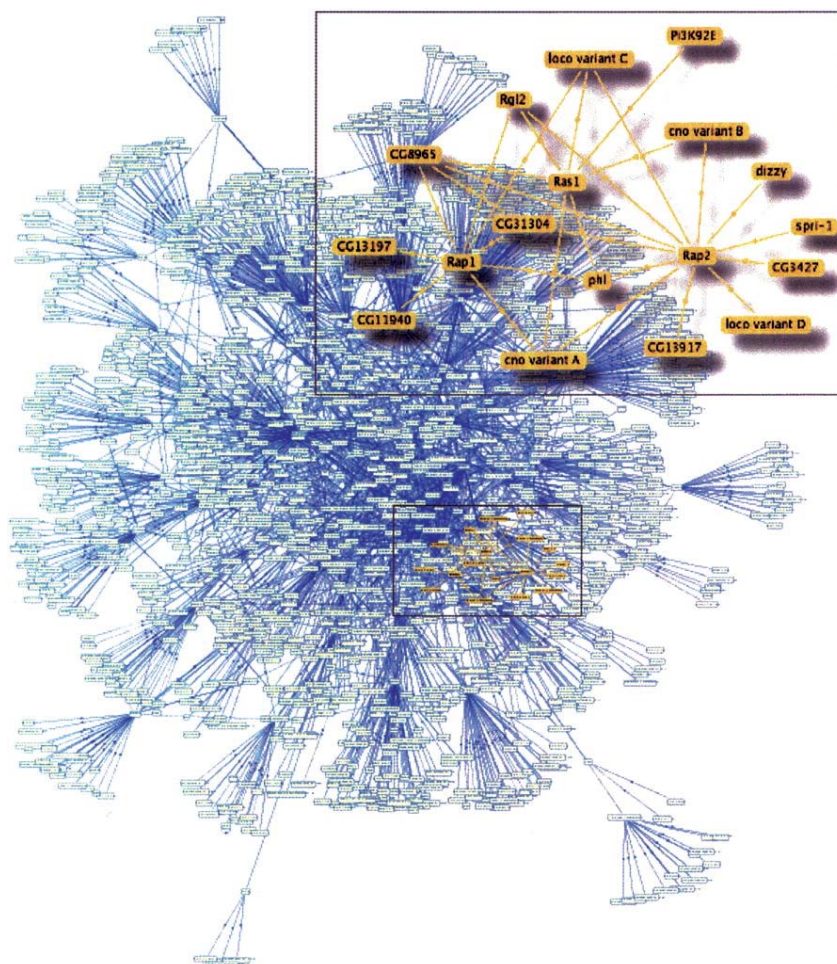


Рис. 1. Карта белок-белковых взаимодействий *Drosophila melanogaster*.

В отдельном окне в увеличенном виде показана подсеть, включающая Ras и другие малые ГТФ-азы. Рисунок представлен с разрешения Camonis, J. и Daviet, L. [96].

могут быть неточными (надежность даже таких методов, как дрожжевой двугибридный анализ, не превышает 50% [23]). Поэтому в настоящее время целый ряд работ посвящен уточнению и корректировке уже имеющихся данных с использованием сочетания различных экспериментальных и теоретических методов для построения более точных карт физического взаимодействия белков [24–26].

МЕТОДЫ ИЗУЧЕНИЯ БЕЛОК-БЕЛКОВЫХ ВЗАИМОДЕЙСТВИЙ

Первоначально для изучения белок-белковых взаимодействий использовались биохимические методы, такие как химическое сшивание, совместное фракционирование в процессе хроматографии, коиммунопреципитация. Однако в последующем для определения интерактомов того или иного организма были разработаны такие высокоэффективные и высокопроизводительные экспериментальные методы, как дрожжевой двугибридный анализ (Y2H – yeast two-hybrid assays), фаговый дисплей, аффинная очистка в тандеме с масс-спектрометрией (TAP-MS – tandem affinity purification-mass spectrometry) [27–30]. Широкие возможности для целей динамической протеомики предоставляют также различные методы микроскопии, а также математические и компьютерные методы. В ряде обзоров эти методы подробно описаны [27, 31, 32]. Поэтому здесь мы остановимся лишь на основных принципах этих методов.

Дрожжевая двугибридная система. Дрожжевой двугибридный анализ позволяет с высокой точностью определять белок-белковые взаимодействия в условиях *in vivo*. Метод основан на использовании факторов транскрипции, которые характеризуются модульностью строения и состоят из физически и функционально разделимых доменов: ДНК-связывающего (BD – binding domain) и домена активации транскрипции (AD – activation domain). Физическое разделение BD и AD доменов приводит к инактивации фактора транскрипции. Активация соответствующих генов становится возможной при реконструкции фактора транскрипции путем слияния этих двух доменов с двумя другими белками – X (называемым «наживкой») или Y (называемым «добычей»), которые взаимодействуют друг с другом [27]. Для создания двугибридной системы обычно используют ДНК-связывающие домены фактора транскрипции GAL4 у *S. cerevisiae* или репрессора lexA у *E. coli*. В качестве активирующих доменов наиболее часто используют домен активации транскрипции белка GAL4 и белок B42, у *S. cerevisiae* и *E. coli*, соответственно [33].

Дрожжевые клетки трансфицируют двумя плазмидами, первая из которых содержит нуклеотидную последовательность, кодирующую белок X, сцепленный с BD доменом, а вторая кодирует белок Y, сцепленный с AD доменом. В результате ДНК-связывающий домен совместно с белком X соединяется с определенной последовательностью репортерного гена, а с другим участком ДНК этого гена связывается AD домен совместно с белком Y. Так как участки ДНК репортерного гена, с которым связываются регуляторные белки, находятся достаточно близко друг к другу, без физического

взаимодействия между белками X и Y активации репортерного гена не произойдет. По наличию продуктов экспрессии репортерного гена в клетках дрожжей можно судить о существовании взаимодействия между анализируемыми белками.

Метод дрожжевого двугибридного анализа был впервые применен С. Филдсом и О. Сонгом в 1989 г. при изучении фактора транскрипции GAL4 у дрожжей *S. cerevisiae*, регулирующего экспрессию гена, кодирующего синтез фермента β -галактозидазы [34], который расщепляет лактозу до глюкозы и галактозы. В составе GAL4 выделяют два функционально важных домена: N-концевой ДНК-связывающий домен, способный взаимодействовать с операторной последовательностью UASg, и C-концевой домен активации транскрипции, богатый кислыми аминокислотами. Об уровне экспрессии β -галактозидазы судят по степени окраски клеточных колоний, продуцирующих фермент, после их инкубации с субстратом. Высокий уровень активации транскрипции наблюдается только в случае, если оба гибридных белка присутствуют в клетке дрожжей. Если белки X и Y взаимодействуют друг с другом, то из двух гибридных белков реконструируется функционально активный белок GAL4, и имеет место активация транскрипции.

При использовании системы LexA точность определения обеспечивается использованием дрожжевых штаммов, содержащих репортерные гены, несущие различное число LexA операторных элементов в промоторах репортерных генов (обычно, lacZ и LEU2). Более чувствительные штаммы дрожжей обладают до шести LexA-связывающими элементами, в то время как менее чувствительные штаммы содержат только 2 связывающих элемента [27].

Дрожжевая двугибридная система была предложена в качестве метода для скрининга библиотек белков, способных взаимодействовать с неким известным белком, используемым в качестве «наживки». Возможность определения физического взаимодействия между разными белками позволяет использовать данную систему для определения специфических аминокислотных остатков, ответственных за взаимодействие [35]. Однако Y2H метод не позволяет оценивать взаимодействие с участием трех и более белков, за исключением таковых у дрожжей. Кроме того, данный метод не всегда позволяет судить о функциональной значимости наблюдаемых физических белок-белковых взаимодействий.

Метод аффинной очистки в тандеме с масс-спектрометрией был предложен в 1999 г. Ригаутом с соавт. как оригинальный способ очистки белков, экспрессируемых в естественных условиях в физиологичес-

ких концентрациях [29]. Метод основан на использовании аффинной метки, которая присоединяется к белку-мишени. Конструкция, состоящая из генов, кодирующих компоненты метки и белок-мишень, внедряется с использованием ретровируса в клетку-хозяин, способную поддерживать экспрессию белка-мишени на уровне, близком к физиологическому. Стандартная метка, используемая у дрожжей, состоит из двух иммуноглобулин G-связывающих фрагментов белка А у *Staphylococcus aureus*, участков для воздействия протеазы вируса табачной мозаики и кальмодулинсвязывающего пептида. Комплекс белка-мишени с меткой выделяют из экстракта клеток путем последовательной двухэтапной процедуры аффинной очистки. Первый этап очистки основан на связывании IgG-сефарозой части метки, представленной белком А, после чего комплекс подвергается воздействию вышеуказанной протеазы. Второй этап основан на связывании части метки, представленной кальмодулинсвязывающим пептидом с кальмодулин-сефарозой в присутствии кальция, и комплекс элюируется с помощью ЭДТА. Использование аффинной метки позволяет осуществлять достаточно быструю очистку белковых комплексов из небольшого числа клеток без необходимости предварительного выяснения белкового состава комплексов и функций отдельных белков в них. В комбинации с масс-спектрометрией данный метод обеспечивает осуществление идентификации исследуемых белков и их взаимодействий [36].

Вначале метод аффинной очистки в тандеме с масс-спектрометрией был применен для изучения белок-белковых взаимодействий и функциональной организации протеомов у простых организмов. Так, в результате работы, проведенной большой группой немецких ученых, в клетках дрожжей *S. cerevisiae* были экспрессированы сотни белков с аффинной меткой, а также изучена их способность формировать комплексы с другими белками [37]. В результате этой работы было очищено 589 белковых комплексов и предсказаны функции для 344 различных белков.

Впоследствии метод TAP-MS был адаптирован для изучения белок-белковых взаимодействий у других организмов, в том числе млекопитающих [38]. Было предложено много разновидностей аффинных меток, а также разработаны метки, легко удаляемые с использованием специфических пептидаз [39, 40]. Для увеличения эффективности метода при изучении белок-белковых взаимодействий в клетках млекопитающих были разработаны также метки с использованием G-белка, обладающего большей аффинностью к иммуноглобулину G, чем белок А. Вместо кальмодулинсвязывающего пептида можно

использовать стрептавидинсвязывающий пептид (GS-TAP – G-protein streptavidin-binding peptide TAP), а для элюции – биотин. Это приводит к десятикратному увеличению количества обнаруживаемых белковых комплексов и повышению специфичности метода. Данный подход позволяет осуществлять с использованием небольшого количества клеток очистку белковых комплексов, которые ранее не удавалось очистить стандартным методом TAP-MS [38].

Масс-спектрометрия основана на определении масс молекул, таких как пептиды или белки, путем предварительной их ионизации и распределения полученных ионов в электрическом поле в зависимости от соотношения массы к заряду (m/z) [30]. Различают два типа мягкой ионизации, наиболее часто используемые в масс-спектрометрии: матрично-активированная лазерная десорбция/ионизация (MALDI – matrix-assisted laser desorption/ionization) и метод электронапыления (ESI – electrospray ionization). Наибольшую популярность в протеомных исследованиях получил метод MALDI, при котором образец, содержащий пептиды, смешивается с молекулами специфической матрицы, на которую воздействует ионизирующий пучок лазера [41, 42].

Впервые возможность применения матрицы для подавления фрагментации при анализе нелетучих органических соединений, таких как белки и пептиды, была продемонстрирована в 1987 году немецким ученым М. Карасом с соавт. [43]. Матрица представляет собой материал, свойства которого обуславливают ионизацию анализируемых молекул, а также понижение деструктивной способности лазерного излучения. Выбывающиеся при этом ионы проходят через масс-анализатор и попадают на детектор, снимающий масс-спектры ионов в соответствии с их значением соотношения массы к заряду m/z . Полученные спектры анализируются путем их сопоставления со спектрами, содержащимися в соответствующих библиотеках, для чего используются специальные компьютерные программы. Одной из разновидностей MALDI является времяпролетная масс-спектрометрия MALDI-TOF (TOF – time of flight), при которой время пролета ионов через масс-анализатор зависит от массы и заряда изучаемых веществ [44, 45].

Другой универсальной разновидностью масс-спектрометрии, активно применяемой в протеомных исследованиях, является активированная поверхностью лазерная десорбция/ионизация (SELDI – surface enhanced laser desorption/ionization). В этом методе образец, содержащий пептиды, не смешивается с матрицей, а наносится на поверхность специального чипа, который в дальнейшем помещается в

вакуумную ячейку, где происходит ионизация исследуемых пептидов или небольших белков [46, 47]. Полученные ионы разгоняются в направлении детектора, в зависимости от их массы. Недостатком данного способа является невозможность непосредственной идентификации белков, представленных в масс-спектрах. Для идентификации белков используют дополнительные методы, такие как, например, фракционирование с помощью ионообменной хроматографии и электрофореза в полиакриламидном геле [48].

Использование масс-спектрометрии, в том числе и в сочетании с другими методами анализа, является в настоящее время универсальным подходом для идентификации белковых маркеров различных заболеваний, включая опухолевые, сердечно-сосудистые и др. [45, 46]. Масс-спектрометрия, наряду с другими методами протеомного анализа, такими как двумерный электрофорез, жидкостная хроматография, белковые биочипы, широко используется также для оценки эффективности тех или иных терапевтических подходов [49].

Метод фагового дисплея применяется для изучения взаимодействия молекул, в том числе белок-белковых взаимодействий, и выявления участков, ответственных за эти взаимодействия. Он основан на использовании бактериофагов для соотнесения генов и кодируемых ими белков. При этом в оболочке фага оказывается представленной (displayed) белковая молекула, информация о которой содержится в рекомбинантной ДНК вируса [50, 51]. Для этих целей используются нитевидные фаги M13, fd и f1, которые наилучшим образом подходят для конструирования рекомбинантной ДНК. Наличие в фаговом геноме участка, который несущественен для его жизнедеятельности, является важнейшим условием для создания гибридных молекул ДНК. В фаговый геном встраивают чужеродный ген, кодирующий синтез определенного белка или пептида (селективного маркера), который после синтеза оказывается представленным в оболочке фага. Фаг, несущий рекомбинантную ДНК, внедряется в бактериальную молекулу (*E.coli*), где происходит его размножение (амплификация). Таким образом, создаются библиотеки, содержащие миллионы фагов, каждый из которых содержит в своем капсиде уникальный белок (или пептид). Далее осуществляют процесс, называемый *in vitro* селекцией, в котором производят скрининг этих библиотек путем взаимодействия экспонированных на поверхности фагов белков с определенным иммобилизованным лигандом.

Метод фагового дисплея обеспечивает прослеживание прямой зависимости между фенотипом и генотипом, так как в оболочке фага оказывается экспрессированной молекула белка, информация о

структуре которой содержится в ДНК вируса. Благодаря простоте и высокой скорости анализа последовательности ДНК, метод фагового дисплея позволяет осуществлять быструю идентификацию исследуемых белков. Используя подобную технологию рекомбинантных ДНК, можно также создавать библиотеки белков или пептидов.

Одним из подходов для оценки белок-белковых взаимодействий является использование фаговых клонов, способных специфически взаимодействовать с полистироловой поверхностью и несущих гены, кодирующие аффинные метки, существенно увеличивающие сродство к этой поверхности [52]. В качестве модельных систем используются мультиферментные комплексы или комплекс антиген-антитело. Например, цистеинсинтазный мультиферментный комплекс у *E.coli* содержит два фермента, а именно серинацетилтрансферазу (САТ) и О-ацетилсеринсульфгидролазу (ОАСС), которые при условии достаточной концентрации серы взаимодействуют друг с другом. При этом активность САТ возрастает, а активность ОАСС падает, что приводит к образованию О-ацетилсерина. Иммуобилизация на полистироловой поверхности гибридного фермента ОАСС, полученного либо путем генетического слияния или химического сшивания с пептидной меткой, увеличивает интенсивность сигнала, оцениваемого с помощью иммуноферментного анализа, по сравнению с сигналом, полученным при иммуобилизации только фермента. Более того, когда меченный пептидами фермент предварительно взаимодействует с САТ в растворе с последующей иммуобилизацией на полистироловой поверхности, интенсивность сигнала еще больше увеличивается благодаря взаимодействию этих ферментов без каких-либо стерических препятствий.

Методы микроскопии в настоящее время широко используются как для количественной оценки изменений концентраций и внутриклеточной локализации различных белков, так и для качественного изучения белок-белковых взаимодействий. Белковые комплексы, образованные благодаря белок-белковым взаимодействиям, могут быть исследованы путем выявления в клетках областей белковых скоплений [53]. Современные методы микроскопии, такие как флуоресцентная микроскопия и криоэлектронная томография, позволяют визуализировать внутриклеточные структуры с разрешением до 4–5 нм [54, 55]. Диаметр белковой глобулы составляет, в среднем, 3–5 нм, а макромолекулярных комплексов – 10–100 нм. Поэтому сочетание двух вышеуказанных методов позволяет реконструировать макромолекулы, их комплексы и отдельные внутриклеточные структуры в их нативном состоянии [56, 57].

Сочетание методов микроскопии с различными экспериментальными подходами, а также с вычислительными методами даёт возможность не только в целом представить внутриклеточную архитектуру, но и создать полный и всеобъемлющий пространственный молекулярный атлас интактной клетки [58]. Так, создаваемый таким способом молекулярный атлас человека содержит информацию о наборах генов и профили экспрессии белков в различных нормальных и патологических тканях. В данном атласе белки классифицированы в соответствии с выполняемыми функциями, а также по тканям и биологическим жидкостям, в которых они обнаруживаются [59].

Для изучения белок-белковых взаимодействий применяют такие методы флуоресцентной микроскопии, как Ферстеровский индуктивный резонансный перенос энергии электронного возбуждения (FRET – fluorescence resonance energy transfer) и изображение по времени жизни флуоресценции (FLIM – fluorescence lifetime imaging). Они позволяют осуществлять качественный анализ белок-белковых взаимодействий с изучением как динамики конформационных изменений, происходящих в белках, в пространстве и во времени, так и аминокислотных последовательностей, участвующие во взаимодействии [60]. Флуоресцентная микроскопия основана на измерении различных характеристик флуоресценции, таких как интенсивность, время затухания, поляризация и длина волны. Метод FRET основан на измерении количества энергии, испускаемой возбужденной молекулой флюорофора и переносимой на молекулу акцептора. Перенос энергии проявляется по увеличению флуоресценции акцептора, которое сопровождается тушением флуоресценции молекулы – донора энергии [61, 62]. При этом перекрывание спектра флуоресценции донора со спектром поглощения акцептора является необходимым условием для переноса энергии. В то же время это условие создает помехи при измерении спектров, и это является недостатком метода. Устранение этого недостатка обеспечивает использование метода FRET для количественной оценки расстояния между взаимодействующими парами молекул в условиях *in vitro* и *in vivo*, что нашло применение в лазерно-сканирующей конфокальной микроскопии. FLIM микроскопия основана на измерении времени жизни флуоресценции в каждой точке пространственного изображения [63]. Она позволяет не только оценивать взаимодействия между белками, но и анализировать локальное окружение флюорофоров, например, pH среды или концентрации различных ионов, кислорода и т.д.

В настоящее время активно разрабатываются методы микроскопии с использованием флюоресцентных белков в качестве молекулярных маркеров, позволяющих наблюдать в реальном времени динамические изменения локализации и концентраций тысяч белков в различных частях одной, отдельно взятой, клетки. Для этих целей создается библиотека клеточных клонов, каждый из которых флюоресцентно мечен по какому-либо определенному белку. Данный подход по получению меченых белков с сохранением их естественной локализации и функций в условиях живой клетки был разработан Ярвиком и соавт. во второй половине 90-х годов и успешно апробирован у *C. reinhardtii* и *D. melanogaster* [64, 65]. Впоследствии метод был применен для создания библиотек клеток, меченных с помощью флюоресцентных белков у млекопитающих, в том числе и у человека [66].

Ряд изящных экспериментов по изучению в режиме реального времени динамических изменений локализации и концентраций различных белков в процессе пролиферации клеток был осуществлен группой Ури Алона [67, 68]. Клетки карциномы легких H1299 человека инфицировались ретровирусом, несущим ген, кодирующий желтый флюоресцентный белок (YFP – yellow fluorescent protein). Была создана библиотека более 1200 клеточных клонов, каждый из которых мог экспрессировать свой флюоресцентно меченый белок. Клетки, содержащие тот или иной меченый белок, отбирались с использованием проточной цитометрии, и затем производилась идентификация меченых белков. С помощью флюоресцентной микроскопии в режиме реального времени наблюдалась динамика изменений локализации и концентраций 20 ядерных белков в течение клеточного цикла. Было установлено, что различные белки характеризуются разной динамикой накопления в ядре клетки. Например, динамика накопления топоизомеразы TOP1 имела синусоидальный характер с максимальным накоплением в ядре в S фазе клеточного цикла, в то время как другие белки характеризовались максимальной ядерной аккумуляцией в G1 или G2 фазах. С помощью данного метода удалось продемонстрировать существование различий в локализации разных белков в ходе клеточного цикла.

Использование таких технологий позволяет изучать также действие лекарств на динамику белков в опухолевых клетках, механизм развития резистентности клеток к лекарствам и роль различных белков в выживаемости клеток. Так, изучение динамики около 1200 различных белков в клетках карциномы легких H1299 под воздействием противоопухолевого лекарства кампто-

тецина, блокирующего топоизомеразу-1 в комплексе с ДНК, что сопровождается разрывами ДНК и подавлением транскрипции генов, позволило выявить изменения концентраций и локализации различных белков под влиянием данного лекарства [68]. Клетки интенсивно делились в течение 24 часов с продолжительностью клеточного цикла около 20 часов. Однако уже через 10 часов после добавления лекарства наблюдалось уменьшение подвижности клеток и торможение их деления; начинались морфологические изменения, свидетельствующие о гибели клеток. Спустя 36 часов описанные изменения охватывали 15% всех клеток.

При этом с течением времени почти у 76% белков наблюдались изменения в интенсивности флюоресценции. Группы функционально родственных белков демонстрировали сходную динамику изменений их внутриклеточной локализации и концентраций. Так, например, было показано, что рибосомальные белки подвергались быстрой деградации, в то время как белки цитоскелета и ферменты разрушались относительно медленно. При этом, наиболее медленную деградацию под воздействием лекарства демонстрировали хеликазы и белки, регулирующие апоптоз, такие как Bcl2-ассоциированные белки BAG2 и BAG3, а также PDCD5 [68]. Топоизомераза-1 подвергалась быстрой деградации, наиболее существенно изменялась локализация фермента. Концентрация двух белков (а именно, РНК-хеликазы и белка DDX5) значительно увеличивалась в тех клетках, у которых наблюдалась тенденция к выживанию, и уменьшалась в тех, которые претерпевали морфологические изменения, приводящие к их гибели.

Таким образом, было показано, что различия в реакции клеток на воздействие лекарства определяются различиями в изменениях концентрации и локализации тех или иных белков. Огромным достоинством подобных методов микроскопии является возможность изучать процессы, происходящие в условиях живой клетки с сохранением естественных функций внутриклеточных макромолекул. Кроме того, они позволяют наблюдать процессы, происходящие в одной, отдельно взятой, клетке в режиме реального времени.

Компьютерные методы. Одним из способов преодоления трудностей, возникающих при изучении белок-белковых взаимодействий, является использование комбинации различных экспериментальных и теоретических методов [69–73]. При этом необходимо отделение истинных результатов от ложноположительных, что требует разработки систем оценки достоверности данных. В связи с трудоемкостью и дороговизной высокопроизводительных экспериментальных методов, важнейшая роль отводится различным компьютерным мето-

дам предсказания белок-белковых взаимодействий [74]. Для этого используют сведения о структуре генов и белков, кодируемых ими, известные данные о функциях белков и возможных функциональных взаимосвязях между ними.

В последние годы разрабатываются и применяются различные компьютерные методы кластеризации, позволяющие оценивать степень функционального сходства между белками, а также выявлять белковые комплексы. Полученные при этом кластеры представляют собой пространственные и функциональные объединения белков. Они сопоставляются с экспериментально подтвержденными и описанными в специальных аннотированных базах данных белковыми комплексами [75]. Некоторые из этих подходов позволяют выявлять также функционально важные модули в интерактомных картах.

Компьютерное исследование белковых комплексов требует применения высокоэффективных вычислительных методов и разработки новых алгоритмов, таких как MCL (Markov Clustering), RNSC (Restricted Neighborhood Search Clustering), SPC (Super Paramagnetic Clustering), MCODE (Molecular Complex Detection) [76, 77]. В последнее время достигнут значительный прогресс в широкомасштабном картировании интерактомов различных организмов, а также в создании баз данных и специальных инструментов для анализа информации, хранящейся в них.

Современные базы данных содержат информацию о сотнях тысяч взаимодействий, образованных несколькими тысячами белков у десятков биологических видов [78–80]. Так, например, база данных BioGRID (Biological General Repository for Interaction Datasets) в настоящее время содержит информацию примерно о 198 тысячах взаимодействий для шести биологических видов [81]. Такие базы данных содержат информацию о взаимодействующих парах белков, полученных либо с помощью экспериментальных методов, либо определенных на основании гомологии между аминокислотными последовательностями, либо предсказанных с использованием компьютерных методов. Целью таких баз данных, как DIP (Database of Interacting Proteins), BIND (Biomolecular Interaction Network Database) и INTERACT является интегрирование огромного количества экспериментальных данных, обеспечение легкого доступа к ним и возможности их визуализации. Базы данных также снабжены инструментами для оценки достоверности экспериментально полученных результатов. Они широко используются для конструирования и анализа белковых сетей, составляющих основу функционирования живой клетки.

Создаваемые в настоящее время базы данных не только содержат информацию о взаимодействующих партнерах, но также позволяют производить детальный структурный анализ участков, ответственных за взаимодействие [82–84]. Например, база данных SCOWLP (Structural Characterization Of Water, Ligands and Proteins) содержит информацию об аминокислотных остатках и группах атомов, участвующих во взаимодействиях. Благодаря этому она обеспечивает осуществление детального анализа взаимодействий между белками, доменами и пептидными мотивами разных белков, а также их взаимодействий с растворителем [82, 83]. Другим примером является глобальная интерактомная карта PSIMAP (Protein Structural Interactome Map), построенная с использованием данных о домен-доменных взаимодействиях с участием всех белков, для которых экспериментально установлена пространственная (3D) структура, представленная в базе данных PDB (Protein Data Bank). С помощью алгоритма PSIMAP можно рассчитать эвклидовы расстояния между аминокислотными остатками двух взаимодействующих доменов в составе разных белков [85]. Два домена считаются взаимодействующими, если, по крайней мере, пять аминокислотных остатков находятся на расстоянии менее 5 Å (правило 5–5). Данный алгоритм позволяет осуществлять предсказание взаимодействующих партнеров на основании гомологии первичных аминокислотных последовательностей белков и их структурных доменов. Информация о взаимодействующих партнерах содержится в базе данных PSIbase.

КАРТИРОВАНИЕ ИНТЕРАКТОВ РАЗЛИЧНЫХ БИОЛОГИЧЕСКИХ ВИДОВ

Наиболее полной картой, составленной для прокариотических организмов, является карта белок-белковых взаимодействий патогенного микроорганизма *Campylobacter jejuni* [86]. С использованием Y2H метода к настоящему времени определены и воспроизведены около 12 тысяч взаимодействующих пар, включая белки, участвующие в регуляции различных биологических явлений, таких как, например, хемотаксис. Другим интенсивно изучаемым представителем прокариот является кишечная палочка *E. coli*, которая рассматривается в качестве модельного организма для изучения прокариотических интерактомов [87]. Однако данные, полученные разными группами авторов для этого организма, противоречивы. Размер интерактома *E. coli*, по разным данным, может составлять от нескольких тысяч до нескольких десятков тысяч взаимодействий.

Так, например, использование подходов функциональной и сравнительной геномики, позволило предсказать существование более 78 тысяч парных взаимодействий у *E.coli* [88]. Кроме того, было показано, что белки, вовлеченные в процессы репликации, транскрипции, трансляции, репарации ДНК и синтеза клеточной стенки, характеризуются большой плотностью связей друг с другом. Особенно интенсивно изучается интерактом клеточной оболочки *E. coli* [89]. На основании данных, полученных как с помощью экспериментальных протеомных методов, так и методов сравнительной и функциональной геномики, была создана база данных Bacteriome.org, содержащая информацию об интерактомах данного организма [90]. Она обеспечивает интегрированный взгляд на интерактом *E.coli* и позволяет пользователям выявлять и анализировать структурные, функциональные и эволюционные взаимоотношения между группами взаимодействующих белков. К настоящему времени эта база данных содержит информацию о более чем 5 тысячах экспериментально подтвержденных взаимодействий с участием более тысячи белков.

Классическим объектом в протеомных исследованиях являются дрожжи *S. cerevisiae*, для которых получены наиболее полные для одноклеточных эукариотических организмов интерактомные данные. Однако, результаты, полученные с использованием разных экспериментальных методов разными группами авторов для *S. cerevisiae*, противоречивы [91, 92]. При этом наиболее точные сведения, включая данные о количестве копий каждого белка и их внутриклеточной локализации, получены при сочетании различных методов [93, 94]. Так, с использованием метода TAP-MS выявлено 7123 взаимодействий с участием 2708 белков [95]. Кластеризация данных с использованием алгоритма Маркова позволила выявить 547 белковых комплексов, каждый из которых содержал в среднем 4,9 белка.

Для получения более точных и надежных данных большой группой авторов была разработана новая «эмпирически контролируемая» система картирования [94]. С её помощью из литературных данных был выбран пул парных взаимодействий, которые были в дальнейшем протестированы с использованием методов Y2H и TAP-MS. Это позволило создать «второе поколение» низкопроизводительных, но высококачественных данных. Такой подход способствовал получению высококачественных результатов, охватывающих около 20% интерактома *S. cerevisiae*.

Картирование белок-белковых взаимодействий у *D. melanogaster* рассматривается в качестве модельной системы для изучения

биологии, развития и механизмов возникновения заболеваний у человека. С применением Y2H системы был произведен скрининг библиотек кДНК *D. melanogaster* для выявления взаимодействующих партнеров для 102 белков, использованных в качестве «наживки» [96]. Большинство из этих белков являлись ортологами опухолеассоциированных или сигнальных белков человека. Были выявлены 2300 парных взаимодействий, из которых 710 были оценены как вызывающие высокое доверие. Подсчет степени достоверности полученных результатов с выявлением взаимодействующих доменов способствовали уточнению данных об уже известных белковых комплексах и предсказанию новых. Картирование взаимодействующих пар для белков – регуляторов клеточного цикла у *D. melanogaster* позволило выявить 1814 взаимодействий для 488 белков [97]. Для данного биологического вида также создаются специальные аннотированные базы данных, содержащие информацию об экспериментально полученных или предсказанных с помощью компьютерных методов данных по физическим белок-белковым взаимодействиям [98–100].

Картирование интерактома человека находится лишь на начальном этапе исследований. Статистическая оценка размера интерактома человека позволила предположить, что он может достигать 650 тысяч взаимодействий [20]. Однако по данным Венкатесана и соавт. интерактом человека представлен 130 тысячами взаимодействий, для подавляющего большинства которых взаимодействующие партнеры еще не выявлены [101]. Эти данные были получены с применением нового «эмпирически контролируемого» подхода, описанного выше. Такой подход позволил произвести оценку качественных параметров методов, используемых для изучения белок-белковых взаимодействий, и результатов, полученных с их помощью. Этими параметрами являлись: чувствительность метода, полнота скрининга, количество выявленных взаимодействий и точность метода (количество артефактов). Исследования этой группы авторов показали, что для оценки белок-белковых взаимодействий у человека наиболее подходит метод двугибридного анализа (Y2H). Интерактомные карты, построенные таким образом, оказались более точны, чем те, которые были получены путем анализа литературных данных. Это обусловлено тем, что последний способ основан на результатах отдельно взятой публикации.

С использованием Y2H метода и метода аффинной очистки в сочетании с биоинформатическими подходами в настоящее время выявлены взаимодействующие партнеры для белков некоторых тканей, включая ткань мозга, почек, эритроциты и др. [102–104]. Постепенно

накапливаемые экспериментальные данные заносятся в специальные базы данных, такие как HPID (Human Protein Interaction Database) и OPHID (Online Predicted Human Interaction Database), которые содержат информацию о белок-белковых взаимодействиях, характерных для человека. Эти базы данных составлены с использованием не только экспериментально полученных результатов, но и предсказанных на основании гомологии с взаимодействующими парами, выявленными у более простых модельных организмов [105, 106].

Изучение белок-белковых взаимодействий и выявление взаимодействующих партнеров, специфичных для той или иной патологии, представляет собой важный инструмент для выяснения механизмов возникновения и развития заболеваний. Так, нарушение синтеза тех или иных компонентов цепи передачи сигнала или мутация в генах, кодирующих их синтез, зачастую является причиной возникновения заболеваний, в том числе и опухолевых [107]. Использование метода аффинной очистки в сочетании с масс-спектрометрией позволило выявить 221 молекулярный комплекс, образованный фактором некроза опухолей TNF- α (tumor necrosis factor- α), его рецептором и внутриклеточными эффекторами [108]. Фактор некроза опухолей TNF- α запускает каскадный механизм передачи сигнала, приводящий, в конечном итоге, к активации ядерного фактора NF- κ B, который выполняет функцию фактора транскрипции и регулирует экспрессию ряда генов, ответственных за пролиферацию и выживание клеток [109]. Нарушение этой регуляции лежит в основе развития многих патологических процессов в организме, таких как опухолевый рост, воспалительные и аутоиммунные заболевания и др. Так, например, ядерный фактор NF- κ B индуцирует экспрессию генов антиапоптотических белков TRAF1 и TRAF2, регулируя тем самым активность ферментов семейства каспаз. Мутация в гене, кодирующем NF- κ B, или в генах, регулирующих его активность, наблюдается при ряде опухолей [110].

Начаты работы по выявлению взаимодействующих пар белков, ассоциированных с нейродегенеративными заболеваниями, серповидноклеточной анемией, шизофренией и др. [111–113]. Картирование интерактомов нейродегенеративных заболеваний, таких как болезнь Альцгеймера, Паркинсона и Хантингтона, амиотрофный латеральный склероз, а также прионовые болезни, позволило выявить, что белки, ассоциированные с этими заболеваниями, характеризуются наличием общих взаимодействующих партнеров [112]. Было выявлено 19 белков, общих для всех этих патологий, причем большинство из них оказалось регуляторами апоптоза или участниками передачи сигнала,

опосредуемого митогенактивируемой протеинкиназой (МАПК). Также в составе этих белков были выявлены характерные для всех них домены, такие как SH2 (Src homology 2) и фосфотирозинсвязывающий РТВ (phosphotyrosine-binding) домены.

III. БЕЛКОВЫЕ СЕТИ

Группа физически взаимодействующих друг с другом белков образует *белковую сеть*. Белковые сети являются одной из разновидностей молекулярных сетей, к которым относятся также: генные сети, включающие гены и контролирующие их функционирование регуляторные РНК и факторы транскрипции; метаболические сети, состоящие из субстратов и продуктов биохимических реакций; и сети сигнальных молекул, включающие рецепторы, их лиганды и внутриклеточные эффекторы [114–118]. Приведенная здесь классификация молекулярных сетей достаточно условна, так как факторы транскрипции и ферменты, катализирующие биохимические реакции, являются по своей природе белками. Кроме того, между компонентами разных типов молекулярных сетей существует функциональная связь. Так, с одной стороны, экспрессия того или иного гена контролируется внешними сигналами, опосредуемыми белковыми рецепторами и их внутриклеточными эффекторами. При этом лигандами рецепторов могут выступать как белки, так и низкомолекулярные промежуточные продукты метаболических путей. Часть внутриклеточных эффекторов – компонентов сигнальных сетей – может проникать в ядро и выполнять функцию факторов транскрипции, контролируя экспрессию генов. С другой стороны, скорость протекания биохимических реакций зависит от активности ферментов, являющихся продуктами структурных генов. В свою очередь, активность ферментов может регулироваться низкомолекулярными субстратами или продуктами биохимических реакций.

Из всех перечисленных типов молекулярных сетей наиболее сложно в отношении функциональных взаимосвязей между компонентами устроены сигнальные сети. Компоненты сигнальных сетей могут как физически взаимодействовать друг с другом (например, лиганд-рецепторное взаимодействие), так и участвовать в химических модификациях других компонентов (например, протеинкиназы) или в регуляции экспрессии генов (внутриклеточные эффекторы). Как видно, белки могут входить в состав разных типов молекулярных сетей в качестве их структурно-функциональных компонентов. Однако

для моделирования физических белок-белковых взаимодействий используют особый тип молекулярных сетей – белковые сети.

ПРИНЦИПЫ ОРГАНИЗАЦИИ И СВОЙСТВА БЕЛКОВЫХ СЕТЕЙ

Молекулярные сети относятся к биологическим сетям, наряду с сетями нервных волокон, кровеносных и лимфатических сосудов и др. Как биологические, так и небιологические сети (такие, как социальные и технологические) по своим характеристикам принадлежат к комплексным сетям, для описания которых используют методы математического анализа и теорию графов [119, 120]. Как показали исследования последних лет, все виды комплексных сетей, как биологических, так и небιологических, устроены по одному и тому же принципу. При использовании метода графов, комплексные сети могут быть представлены в виде совокупности узлов, связанных друг с другом ориентированными или неориентированными ребрами. В узлах располагаются компоненты сети, а ребра показывают связи между ними. Регуляторные геновые и метаболические сети могут быть представлены в виде графов с ориентированными ребрами, в которых направление связи либо указывает на ген, который испытывает регуляторное воздействие со стороны фактора транскрипции, или показывает направление реакции от субстрата к продукту [119, 120]. Сети сигнальных молекул могут быть представлены в виде графов, как с ориентированными, так и с неориентированными ребрами. Поскольку в белок-белковых взаимодействиях оба партнера принимают одинаковое участие, то белковые сети представляют в виде графов, в которых соседние узлы, образуемые отдельными белками, связаны друг с другом неориентированными ребрами (рис. 2).

Глобальная структура комплексных сетей может быть представлена большими графами, состоящими из тысяч взаимосвязанных узлов. При описании *локальной* структуры рассматривают отдельные части сетей (подсети), которые могут быть представлены всего несколькими узлами и связями между ними [121]. Среди глобальных параметров, используемых для описания принципов организации комплексных сетей, выделяют топологические и динамические характеристики. Полученные с использованием этих параметров знания об архитектуре и динамике белковых сетей позволяют выявить основные принципы функционирования внутриклеточных структур [122]. К *топологическим характеристикам* относят: количество узлов в составе сети и количество ребер у каждого узла (т.е. связанных с ним соседних узлов), среднюю длину пути или диаметр сети, её плотность и гетерогенность, коэффициент кластеризации. Среди

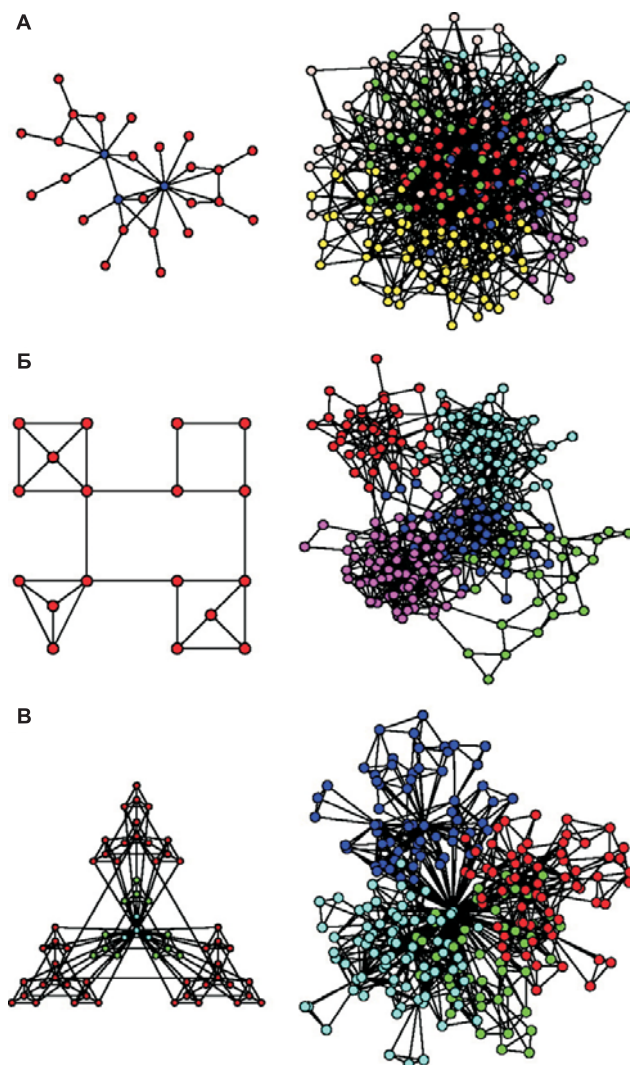


Рис. 2. Типы комплексных сетей (по [131]).

А – схематическое представление (слева) и конфигурация (справа) безмасштабной сети. Синие кружочки (слева) соответствуют хабам.

Б – схематическое представление (слева) и конфигурация (справа) модульно организованной сети. Здесь все узлы имеют одинаковое число связей с соседними узлами. Такая сеть не содержит хабов.

В – схематическое представление (слева) и конфигурация (справа) иерархично и модульно организованной безмасштабной сети. Рисунок представлен с разрешения Barabasi, A.-L. [131].

динамических параметров можно выделить устойчивость сети к тем или иным внешним воздействиям, частоту и амплитуду колебаний (осцилляций), возникающих в сети.

Одним из важных показателей комплексных сетей является *длина пути* или расстояние между любыми двумя узлами в сети, характеризуемое некоторым количеством других узлов между ними. Наименьшее расстояние между двумя узлами представляет собой путь наименьшей длины. С помощью подсчета среднего значения длин всех таких путей между всеми парами узлов определяют среднюю длину пути в сети или её диаметр. Было показано, что белковые сети проявляют свойства «маленького мира» (small world) с диаметром (т.е. средним значением длины пути), равным 4–5 узлам. Сети с характеристиками «маленького мира» были впервые описаны Уотсом и Строгатсом [123], установившими, что расстояния между узлами во многих биологических и небιологических сетях характеризуются небольшой протяженностью. Как следствие, процессы, происходящие в большинстве комплексных сетей, характеризуются быстрой динамикой, эффектом усиления и синхронизации сигнала. Сети со свойствами «маленького мира» занимают промежуточное положение между правильными графами, имеющими тенденцию к минимизации связей, и графами со случайной архитектурой, характеризующимися большим количеством связей [124]. Такие сети демонстрируют наличие *хабов* (hub – центр, средоточие), т.е. небольшого числа узлов, имеющих большое количество связей.

Коэффициент кластеризации представляет собой меру способности узлов образовывать области с высокой плотностью связей, т.е. кластеров [124]. Среднее значение коэффициента кластеризации для всех узлов показывает коэффициент кластеризации всей сети. В реальных сетях для узла со степенью k коэффициент кластеризации $C(k)$ пропорционален k^{-1} , где k показывает количество соседних, связанных с ним узлов ($k = 1, 2, 3 \dots$). Коэффициент кластеризации является мерой гетерогенности (неоднородности) и иерархичности организации сети. *Иерархичность* сети означает существование многоуровневой формы организации со строгой соподчиненностью нижних уровней верхним, т.е. одних групп узлов другим. Каждая из этих групп (кластеров), характеризующаяся большой плотностью связей между узлами, представляет собой структурно-функциональный модуль в составе сети.

Для характеристики гетерогенности комплексных сетей обычно используются два основных типа моделей – графы со случайной

геометрией (random geometry) и «безмасштабные» (scale-free) [125]. Данные, полученные разными группами авторов и характеризующие архитектуру белковых сетей, неоднозначны и противоречивы. Так, глобальный и локальный анализ белковых сетей у *S. cerevisiae* и *D. melanogaster*, осуществленный Прзулж с соавт. [126] с использованием разных моделей, показал, что для описания физических белок-белковых взаимодействий лучше подходят графы со случайной геометрией. Сети со случайной геометрией описываются с помощью графов $G(n, r)$, состоящих из n числа узлов, представленных с помощью n независимых, одинаковым и случайным образом распределенных в метрическом пространстве точек, расстояние между которыми равно r . Такие сети характеризуются достаточной однородностью, и количественная оценка вероятности связей для любого узла характеризуется биномиальным распределением [127]. В случае больших сетей плотность вероятности того, что некий узел имеет k связей, характеризуется распределением Пуассона (рис. 3), которое имеет следующий вид:

$$K_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda},$$

где $k = 0, 1, 2, \dots$ – количество связей, приходящихся на узел, λ – среднее арифметическое.

Здесь каждый узел имеет примерно одинаковую вероятность быть связанным с любым другим узлом.

Однако, системный анализ топологических характеристик белковых сетей, осуществленный у *S. cerevisiae*, *C. elegans* и *D. melanogaster* другими группами исследователей, выявил «безмасштабный» (scale-free) характер большинства сетей и высокую степень их кластеризации [128, 129]. В данном случае «безмасштабность» имеет отношение только к характеру распределения и означает существование степенной зависимости в ней [128]. Функция распределения степеней для «безмасштабных» сетей может быть задана как $P(k) = Ak^{-\gamma}$, где $P(k)$ – плотность вероятности образования связей между соседними узлами, A – постоянная, показатель γ обычно $2 < \gamma < 3$ (чаще всего 2,2) для всех организмов [122]. Такая функция распределения показывает высокую гетерогенность сети, т.е. практическую невозможность обнаружения в ней типичного узла, с помощью которого можно было бы охарактеризовать все остальные узлы в сети.

Ключевым свойством «безмасштабной» архитектуры комплексных сетей является наличие хабов, т.е. узлов с большой плотностью

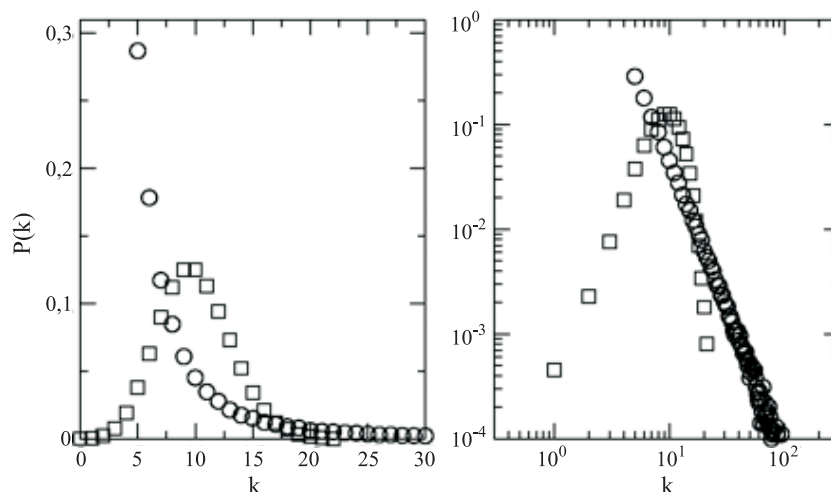


Рис. 3. Сравнительные графики функций распределения степеней для белковых сетей со случайной геометрией (квадратики) и «безмасштабных» белковых сетей (кружочки). Слева даны линейные графики, справа – в логарифмической шкале (по [85]). Колоколообразный график функции распределения для белковых сетей со случайной геометрией указывает на статическую гомогенность таких сетей. График функции распределения для «безмасштабных» белковых сетей подчиняется формуле $P(k) = Ak^{-3}$ и характеризует гетерогенность сети. Характер падения функции распределения указывает на существование большого числа узлов с малым количеством связей и небольшого количества хабов, обладающих большим количеством связей.

связей, в то время как большинство узлов характеризуется небольшим количеством связей (рис. 2). Однако это небольшое число хабов обеспечивают устойчивость всей сети, связывая всю сеть в единое целое. О роли хабов свидетельствуют эксперименты по их удалению из белковых и метаболических сетей у *D. melanogaster* [129]. Оказалось, что сети с «безмасштабной» архитектурой устойчивы к случайному удалению узлов. Даже при удалении большого количества случайно выбранных узлов связи между остальными узлами в сети не нарушаются и топология сети не изменяется. Удаление же только хабов приводит к увеличению диаметра сети в два-три раза.

Биологическое значение удаления хабов показано в экспериментах на *S. cerevisiae*, продемонстрировавших, что лишение этих организмов (нокаутирование) генов, кодирующих белки, расположенные в хабах, сопровождается увеличением летальности. В то время как удаление генов, кодирующих белки, расположенные в других узлах,

не приводит к такому эффекту [130]. Подобное свойство сетей было названо летальность–центральность (lethality–centrality). Эти данные свидетельствуют о том, что белки, расположенные в хабах, необходимы для выживания организма и, в целом, могут быть функционально важнее, чем белки, расположенные в других узлах.

Иерархичность служит фундаментальной характеристикой многих комплексных сетей и показывает, что большие группы узлов в таких сетях состоят из более мелких групп (модулей), организованных в иерархическом порядке [131]. Модули могут быть определены как пространственно и структурно независимые единицы, состоящие из нескольких компонентов и способные относительно самостоятельно функционировать. При этом связи между узлами, принадлежащими разным модулям, характеризуются меньшей плотностью, чем связи между узлами, принадлежащими одному и тому же модулю.

Предположение о модульности организации белковых сетей было сделано на основе системного анализа с использованием биоинформатических ресурсов данных по экспрессии, внутриклеточной локализации, эволюции, структуре и функциям белков и их взаимодействующих партнеров [132–135]. Несмотря на то, что некоторыми авторами обосновывается гипотеза об отсутствии биологической значимости модулей белковых сетей [135], целый ряд данных свидетельствует об их функциональности. Сравнение экспериментальных методов и функциональной аннотации генов показало возможность существования двух типов модулей в белковых сетях: а) белковые комплексы и б) динамические функциональные модули, объединяющие белки, участвующие, например, в регуляции клеточного цикла [136].

Существование строгой корреляции между строением белковой сети и функциональной ролью и внутриклеточной локализацией белков, входящих в их состав, было продемонстрировано в ряде работ [137, 138]. Так, на основании информации о взаимодействующих парах белков, извлеченных из базы данных DIP, были построены белковые сети для *S. cerevisiae*. С использованием алгоритма Girvan-Newman (G-N) и программы MoNet было установлено, что эти сети организованы в 86 простых модулей, каждый из которых состоял из более, чем трех белков [133]. При этом каждый модуль был представлен, в основном, функционально взаимосвязанными белками. Используя интегрированный подход с привлечением, наряду с протеомными данными, результатов, полученных с помощью анализа экспрессии генов с применением олигонуклеотидных чипов, другими авторами было выявлено уже 266 функциональных модулей

в белковых сетях дрожжей [139]. Вероятность взаимодействия между белками, принадлежащими функционально различным модулям, оказалась низкой [140].

С использованием метода компьютерного моделирования было показано, что модульность организации молекулярных сетей возникает, вероятнее всего, как следствие дупликации генов. Выявление того обстоятельства, что эволюция белковых сетей происходит на уровне модулей, обуславливает необходимость изучения степени консервативности между взаимодействующими парами белков. Это достигается путем сравнения первичных структур белков, входящих в состав тех или иных сетей с помощью метода выравнивания, а также сравнением архитектуры различных сетей [141, 142]. В связи с этим в ряде работ предлагаются специальные алгоритмы, которые предполагают поиск сходных сетей среди набора молекулярных сетей у разных биологических видов или внутри одного вида.

Для характеристики топологии белковых сетей используется построение пространственных (3D) моделей и их визуализация [143, 144]. Такие модели создаются с использованием протеомных данных, таких как внутриклеточная локализация белков, приблизительное количество копий каждого белка в клетке, их физико-химические характеристики, данные о пост-трансляционных модификациях и ортологах белков. Специальные базы данных содержат информацию об органеллоспецифических белок-белковых взаимодействиях, которая также используется для трехмерного моделирования [145–147].

Реально существующие сети, в отличие от статических графов, характеризуются наличием *динамических* свойств, т.е. изменений в пространстве и во времени. Временные характеристики белок-белковых взаимодействий можно изучать с помощью анализа экспрессии генов с применением олигонуклеотидных чипов. Подсчет корреляции экспрессии белков, расположенных в хабах, с экспрессией их взаимодействующих партнеров позволил выделить два типа белков, расположенных в хабах [148]. Для первого типа оказалась характерна высокая, а для другого – низкая степень ко-экспрессии белков. Хабы первого типа являются статическими (party hubs), а второго типа – динамическими (date hubs). Было высказано предположение, что белки, расположенные в статических хабах, характеризуются постоянным набором взаимодействующих партнеров, а белки в динамических хабах взаимодействуют с разными партнерами в разное время. Первые, видимо, играют локальную роль в сетях, характеризуясь прочными связями внутри функционального модуля. Вторые имеют глобальное значение, связывая между собой разные

функциональные модули. Было показано, что удаление динамических хабов приводит к более тяжелым последствиям (увеличению диаметра сети и её дезинтеграции), чем удаление статических хабов. Однако вскоре выяснилось, что увеличение летальности связано с удалением обоих типов хабов [149, 150].

Среди важных динамических характеристик комплексных сетей необходимо выделить их устойчивость (robustness) к воздействию того или иного фактора и существование периодических колебаний (осцилляций) в функционировании компонентов сети. Эти колебания отражают цикличность процессов, происходящих в клетке. Циклы клеточной жизнедеятельности контролируются путем каскадной регуляции по принципу прямой и обратной связей [151, 152]. Жизнедеятельность клетки зависит от скоординированного функционирования генов, их белковых продуктов, а также низкомолекулярных метаболитов, принимающих участие в подобных регуляторных путях. Динамические характеристики белковых сетей и способы их моделирования будут рассмотрены в разделе V.

МЕДИКО-БИОЛОГИЧЕСКОЕ ЗНАЧЕНИЕ АНАЛИЗА БЕЛКОВЫХ СЕТЕЙ

Анализ белковых сетей позволяет решать ряд задач фундаментальной медицины, среди которых необходимо выделить выявление и объяснение механизмов возникновения и развития заболеваний, включая опухолевые, нейродегенеративные, сердечно-сосудистые, аутоиммунные, а также поиск молекулярных мишеней для действия лекарств.

В ряде работ с использованием базы данных OPHID, содержащей информацию о белок-белковых взаимодействиях, экспериментально подтвержденных или предсказанных для человека, были построены графы, иллюстрирующие белковые сети, включающие продукты генов, экспрессируемых при опухолях. Было показано, что сети, образованные опухолеассоциированными белками – продуктами генов, отличающихся по способу регуляции, больше по размеру, чем сети, образованные случайным набором белков [153]. Это указывает на существование функциональной взаимосвязи между белками. На основе анализа белковых сетей была предложена модель канцерогенеза, которая рассматривает это явление как специфически организованный на молекулярном уровне процесс, характеризующийся снижением экспрессии топологически и функционально связанных белков, синхронизированных с повышением экспрессии других белков [154].

Имеющиеся в настоящее время данные по анализу белковых сетей, ассоциированных с заболеваниями, неполны и неоднозначны. Пока-

зано, что белки, ассоциированные со сходными заболеваниями, характеризуются большей вероятностью физически взаимодействовать друг с другом. Было высказано предположение о существовании в белковых сетях функциональных модулей, специфичных для тех или иных заболеваний. При этом белки, необходимые для эмбрионального развития и нормального функционирования клеток, синтезируются в различных органах и располагаются в хабах белковых сетей. В то время как подавляющее большинство белков, ассоциированных с заболеваниями, как правило, располагаются на периферии сетей [155].

Однако в ряде работ показано, что белки, ассоциированные с опухолями, характеризуются большой плотностью связей, располагаются, в противовес нормальным белкам, в центральных хабах и содержат большое количество структурных доменов, участвующих в белок-белковых взаимодействиях [156–158]. Опухлеассоциированные белки содержат в два раза больше взаимодействующих партнеров, чем нормальные белки. Наличие большего количества взаимодействующих партнеров обуславливает центральную роль этих белков в сети и означает большую вовлеченность в патофизиологические процессы, происходящие в клетке. Вочи с соавт. картировали в интерактоме человека белковые продукты 360 генов, экспрессия которых повышена, и 270 генов, экспрессия которых снижена при раке легких [159]. Выявлено, что белки, демонстрирующие гиперэкспрессию, характеризуются большим количеством связей, по сравнению с белками, экспрессия которых снижена. Таким образом, была показана высокая степень централизации белков, синтез которых повышен в опухолевой ткани, по сравнению с нормальной.

Анализ модулей белковых сетей показывает, что они содержат продукты совместно регулируемых и функционально взаимосвязанных генов и могут быть ассоциированы, например, с полиморфизмом генов или механизмом возникновения заболевания [160]. Более того, белки, проявляющие активность только в составе определенного функционального модуля, могут рассматриваться в качестве маркеров данного модуля или в качестве потенциальных мишеней для лекарственных средств. Выявление генов, ассоциированных с опухолевыми заболеваниями, и их белковых продуктов, взаимодействующих с известными белками – биомаркерами опухолей, может способствовать разработке новой стратегии в диагностике заболеваний [161].

Использование специальных компьютерных программ типа Cytoscape позволяет визуально сопоставлять экспериментальные

данные и использовать их совместно с информацией, содержащейся в аннотированных базах данных по молекулярным сетям [162]. Так, с применением данного подхода было показано, что белки, участвующие в регуляции эпителиально-мезенхимного перехода, инициированного в клетках почек под воздействием TGF- β 1, образуют единую сеть [163]. Анализ онтологии генов показал гиперэкспрессию белков, входящих в состав сигнальных путей, контролирующих морфогенез и эмбриональное развитие.

Анализ молекулярных сетей способствует также пониманию механизмов, лежащих в основе возникновения комплексных заболеваний, обусловленных как генетическими, так и негенетическими факторами, такими как факторы окружающей среды, тип питания и др. [164]. На этой основе предлагается новый подход к диагностике и классификации заболеваний. Анализ молекулярных сетей позволяет также выявлять новые потенциальные мишени для лекарств и определять резистентность клеток к лекарствам. Фактически, это обеспечивает новые подходы к лечению заболеваний. Так, сравнительный анализ сетей, образованных белками, вовлеченными в апоптоз клеток HeLa и нормальных фибробластов у человека, способствовал не только выяснению механизма развития апоптоза, но и поиску потенциальных мишеней для действия лекарств [165]. Было установлено существование большого числа взаимодействий (841) в опухолевых (HeLa) и нормальных клетках. Из них 18,7% присутствовали в опухолевых клетках и отсутствовали в нормальных. И примерно столько же взаимодействий, наоборот, было выявлено в нормальных, но не обнаруживалось в опухолевых клетках. Суммарно эти взаимодействия были определены как потенциальные мишени для действия лекарств. Было высказано предположение, что белки Bcl2, PT53 и каспаза-3 могут являться мишенями для действия лекарств. Интересным результатом, полученным в этой работе, представляется также выявление белков, расположенных в статических и динамических хабах белковых сетей. Было показано, что каспаза-3 располагается в динамических хабах сетей, образованных белками, ответственными за апоптоз как нормальных, так и опухолевых клеток. Каспазы-2 и -9 оказались ответственными за топологические различия между сетями. Более подробно анализ молекулярных сетей в выяснении механизмов, диагностике и классификации заболеваний изложен в обзорах [155, 156, 166].

БЕЛКОВЫЕ КОМПЛЕКСЫ

Белок-белковые взаимодействия обуславливают образование стабильных белковых комплексов, имеющих важное функциональное значение. Методами микроскопии было выявлено, что в цитоплазме живых клеток белковые молекулы распределены не равномерно, а существуют в виде отдельных скоплений [167–169]. В таких скоплениях содержание макромолекул может колебаться от 50 до 400 г/л, облегчаются белок-лигандные и белок-белковые взаимодействия, конформационные переходы макромолекул и образование самоорганизующихся надмолекулярных структур [170].

Методы современной протеомики позволяют достаточно детально охарактеризовать состав и способы организации этих внутриклеточных структур. Белковые комплексы являются ключевыми надмолекулярными структурами, в которых интегрированы продукты нескольких генов и которые, в основном, предназначены для выполнения каких-либо взаимосвязанных функций. Это может быть мультиферментный комплекс, предназначенный для катализа последовательности биохимических реакций, или комплекс белков – участников пути передачи сигнала.

Белковые комплексы образуются благодаря тому, что каждая белковая молекула может обладать сразу несколькими белок-связывающими участками. Так, например, изучение линкерного белка для активации Т-лимфоцитов, LAT (linker for activation of T cells), показало, что в его взаимодействии с SH2-доменами адаптерных белков сигнальных путей участвуют четыре участка, содержащих фосфорилированные остатки тирозина [171]. Это способствует образованию белковых комплексов, которые облегчают передачу сигнала с мембраны внутрь клеток, что лежит в основе нормального созревания и дифференцировки иммунокомпетентных клеток. Мутации, приводящие к замене остатков тирозина, участвующих в связывании различных адаптерных белков, приводят к нарушениям дифференцировки Т-лимфоцитов и созревания В-лимфоцитов.

В белковых комплексах можно выделить ядро, которое образовано постоянным набором белков, и периферийную часть, состав которой может изменяться. Показано, что существует кооперативность взаимодействия между различными белками в составе такого комплекса, что определяется различной аффинностью и специфичностью их связывания друг с другом [172]. Характер и механизм ассоциации и диссоциации белковых комплексов, в основном, зависит от их размеров. Исследования с *S. cerevisiae* показали экспоненциальное падение распределения размеров белковых комплексов. Однако изу-

чение динамики белковых комплексов выявило, что их ассоциация может происходить независимо от размера комплекса [173].

Часть белков может участвовать в образовании сразу нескольких комплексов, что может объясняться полифункциональностью и мультимодульностью этих белков. Поскольку белковые комплексы предназначены, в основном, для выполнения какой-либо определенной функции, полифункциональные белки могут реализовывать свои возможности в составе разных комплексов с участием разных функциональных модулей.

IV. СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЕ МОДУЛИ БЕЛКОВ, ОТВЕТСТВЕННЫЕ ЗА ИХ ВЗАИМОДЕЙСТВИЯ

Обнаружение одних и тех же белков в составе различных белковых комплексов может быть объяснено наличием в их составе нескольких функционально важных участков, включая те, которые ответственны за белок-белковые взаимодействия. Это означает мультимодульность строения белков. Исследования последних лет показывают, что модульность является универсальным свойством живого и проявляется на всех уровнях его организации, от молекулярного до надорганизменного. Модульность характерна, как уже было упомянуто, также и для белковых сетей, что обуславливает сложный, иерархичный характер их организации.

В свою очередь, модульность организации может быть присуща и отдельным белкам, что означает, что белковая молекула может состоять из нескольких (и даже множества) структурно и функционально независимых элементов (доменов и мотивов). При этом каждый из модулей ответственен за ту или иную функцию белка и может функционировать независимо от других. В результате этого белковая молекула приобретает способность выполнять целый комплекс различных функций; такие белки называются мультимодульными и полифункциональными [17, 18]. Вероятно, эти функции взаимосвязаны, т.е. набор функциональных модулей того или иного конкретного белка, видимо, образован не случайным образом. Какой из модулей белка может быть задействован, зависит от типа клетки, её микроокружения, физиологического и патофизиологического состояния клетки, а также микроокружения самой белковой молекулы. Мозаичность строения, мультимодульность и полифункциональность, по-видимому, характерны для большинства белков эукариот.

Структурно сходные модули могут встречаться в составе различных белков, вероятно, обуславливая сходство некоторых их

функций. Возможно, что молекулы различных белков построены путем различной комбинации ограниченного набора структурно и функционально независимых модулей, что, в свою очередь, детерминировано физиологической (биологической) ролью того или иного белка [174]. Данное предположение подтверждается экспериментами по созданию искусственных полифункциональных белков, образованных разным сочетанием пептидных мотивов с уже известной функцией. Эти эксперименты показали: 1) компактная упаковка белковой молекулы не является необходимым условием для реализации её функций; 2) функции того или иного мотива зависят от их набора и взаимного расположения [175]. Более того, оказалось, что путем перегруппировок различных мотивов можно получить белки с совершенно разными функциями.

Мультимодульность белков может приводить к значительному усложнению характера белковых сетей [176]. В случае, если в узлах белковых сетей располагаются мультимодульные, полифункциональные белки, каждый узел может быть представлен не отдельным белком, а его функциональным модулем, и связи между узлами становятся взаимопереплетенными. В связи с этим представляется важным составление структурно-функциональных карт для мультимодульных и полифункциональных белков с целью выявления участков, ответственных за ту или иную функцию, в том числе за белок-белковые взаимодействия.

Экспериментально показано, что белок-белковые взаимодействия осуществляются с помощью доменов одних белков и соответствующих им коротких линейных пептидных мотивов других белков. Если белок содержит несколько доменов для взаимодействия или несколько связывающих мотивов, то он может одновременно взаимодействовать с несколькими белками, что приводит к образованию белковых комплексов. Примерами доменов, участвующих в белок-белковых взаимодействиях, являются SH2 и РТВ домены, связывающиеся с фосфорилированными остатками тирозина в составе рецепторов, или SH3 и WW домены, реагирующие с пролин-богатыми мотивами белков [177–179].

Одной из модельных систем для изучения белок-белковых взаимодействий является взаимодействие SH3-доменов с белками, содержащими домены, богатые пролином [180, 181]. Белковые сети, образованные взаимодействиями различных белков с SH3-доменами, были выявлены у *S. cerevisia*, у которых определено 28 белков, содержащих такие домены. С использованием Y2H метода было выявлено

233 взаимодействия с участием 145 белков, метод фагового дисплея выявил 394 взаимодействия между 206 белками.

Изучение строения SH3-доменов показало, что их полипептидная цепь содержит около 50–70 аминокислотных остатков и организована в пять β -складчатых структур. В настоящее время известно более 1500 различных SH3-доменов, классическим участком связывания которых является PXXP мотив (где Р – пролин, а Х – любая аминокислота) [182–184]. SH3-домены содержатся в составе ферментов, таких как киназы, липаза или ГТФ-аза. Наиболее изучены функции SH3-доменов адаптерных белков, таких как с-Src или Grb2, принимающих участие в передаче сигнала внутрь клеток, от мембранных рецепторов к их цитоплазматическим эффекторам. Так, с участием PXXP мотива богатая пролином тирозинфосфатаза (PEP – proline-enriched tyrosine phosphatase) связывается с SH3-доменом цитоплазматической тирозинкиназы Csk (C-terminal Src kinase) [184]. Показано, что в этом взаимодействии задействован участок, включающий A40, T42 и L43 в составе SH3-домена, образующий гидрофобные связи с PEP. Другим примером является взаимодействие PXXP мотивов белков динаминов-1 и -2 с очищенными SH3-доменами таких белков, как с-Src, Grb2 и интерсектин. Исследование кинетики подобных взаимодействий показало, что различные SH3-домены могут связываться с одним и тем же пролин-богатым доменом. Видимо, в условиях *in vivo* несколько SH3-доменов могут конкурировать за связывание с PXXP мотивами [185].

Короткие линейные мотивы или сегменты белков – это состоящие, в основном, из 3–10 аминокислотных остатков последовательности, ответственные за ту или иную функцию белка [186]. Они участвуют в белок-белковых взаимодействиях, взаимодействиях по типу белок-лиганд, белок-нуклеиновая кислота, являются участками посттрансляционных модификаций белков, таких как фосфорилирование, гликозилирование и др. Первыми линейными мотивами, обнаруженными в составе белков, были последовательности KDEL и HDEL, являющиеся участками белков цистерн эндоплазматического ретикулума и ответственные за предотвращение секреции этих белков [187]. Было также показано, что мотив KKXX, обнаруженный в цитоплазматическом домене трансмембранных белков выполняет функцию сигнального участка, ответственного за возвращение белков из аппарата Гольджи в цистерны эндоплазматического ретикулума [188, 189].

Короткие линейные мотивы, состоящие из небольшого количества аминокислотных остатков, трудно выявить путем сравнения первичных последовательностей белков. И поэтому для их выявления часто используются довольно трудоемкие и многоступенчатые экспериментальные методы. Для выявления коротких линейных мотивов и роли тех или иных аминокислотных остатков в их функционировании обычно используются такие экспериментальные методы, как точечный мутагенез или фаговый дисплей. В настоящее время появляются новые биоинформатические ресурсы для обнаружения коротких линейных мотивов в составе белков [190, 191]. Компьютерные методы для обнаружения связывающих мотивов основаны на использовании баз данных по белок-белковым взаимодействиям и извлечении мотивов, общих для группы взаимодействующих друг с другом белков. Для этого могут быть использованы специальные алгоритмы такие, как D-MOTIF, D-STAR, MEME, Gibbs Sampler, PRATT и TEIRESIAS [192–195]. Данный подход основан на допущении, что белки, имеющие общих взаимодействующих партнеров, должны характеризоваться наличием сходных мотивов. Появляющиеся в настоящее время базы данных содержат информацию обо всех известных функционально важных участках белков. Так, например, база данных SCOWLP содержит информацию о более чем 9 тысячах белок-связывающих участков белков, принадлежащих более чем к 2,5 тысячам семейств. При этом, оказалось, что представители 65% семейств содержат более одного связывающего участка, а 22% участков задействованы в формировании комплексов с несколькими белками, принадлежащими разным семействам [82, 83].

Методы сравнительной геномики и кластеризации генов были использованы для выявления функционально важных олигопептидов в белках семи биологических видов [196]. Данный подход позволил выявить три- и тетрапептиды с уже известной функцией, такие как сигнальные мотивы SKL, KDEL/HDEL и KKXX, а также предсказать новые мотивы, возможно, имеющие функциональное значение. Специальные инструменты дают возможность определять потенциальные контакты между аминокислотными остатками сразу нескольких полипептидных цепей [197]. Так, алгоритм Con-Struct Map позволяет изучать изменения, происходящие в пространственной структуре белков в результате замены аминокислотных остатков, или, наоборот, определять консервативные остатки, важные для взаимодействия, исследовать структурно-функциональные взаимоотношения между неродственными, негомологичными белками.

V. МЕТОДЫ ДИНАМИЧЕСКОГО МОДЕЛИРОВАНИЯ *IN SILICO*

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

Используемые для описания топологии и свойств белковых сетей статические графы не отражают конформационно-динамических характеристик макромолекулярных комплексов и множественности функций белков. Поэтому в настоящее время актуальным является развитие методов описания конформационно-динамических свойств белковых комплексов и создание динамических моделей процессов, происходящих в клетке. В системной биологии используются методы: а) математическое моделирование с помощью системы уравнений, или б) компьютерное моделирование, основанное на специальных алгоритмах для конструирования, дизайна и визуализации внутри- и межклеточных процессов и явлений [198–200].

Математическое моделирование становится важным инструментом, используемым для интегрированного подхода к пониманию сложных внутри- и межклеточных процессов [201]. Оно представляет собой универсальный способ описания явлений и процессов, происходящих в живой клетке, и служит естественным языком для интерпретации экспериментальных данных, предсказания свойств и поведения биомолекул в различных условиях. Наиболее изученными объектами с точки зрения математического моделирования являются биомолекулярные системы, включающие лиганды, рецепторы, адаптерные белки и внутриклеточные эффекторы сигнальных путей [202]. Так, с использованием данного подхода была предпринята попытка выяснения механизмов пролиферации фибробластов под воздействием эпидермального фактора роста (ЭФР) [203]. Используя параметры кинетики связывания фактора роста с его рецептором, а также динамики интернализации, деградации и рециклизации лиганд-рецепторного комплекса и синтеза ДНК, удалось объяснить многие известные на тот момент экспериментальные данные и предсказать новые свойства этой сигнальной системы. Аналогично обстоит дело и с изучением динамических свойств белковых сетей, включающих регуляторы клеточного деления [204, 205]. Математическое моделирование используется здесь для оценки различных динамических параметров клеточного цикла, таких как скорость биохимических реакций, динамика накопления белков и их деградации, продолжительность различных стадий клеточного цикла или его ареста в интерфазе [206]. Математические модели могут быть использованы также и для предсказания фенотипических

последствий мутаций, например, как мутации в генах, кодирующих белки – регуляторы клеточного цикла, влияют на его протекание.

Методы компьютерного моделирования чаще всего используются для изучения динамических характеристик внутриклеточной передачи сигнала, например, каскадного механизма передачи сигнала, опосредуемого митоген-активируемой протеинкиназой (МАПК). Активность компонентов каскадной передачи сигнала регулируется по принципу положительной или отрицательной обратной связи. Функциональная организация подобных каскадных механизмов обеспечивает существование феномена сверхчувствительности клеток к внешним сигналам. Причем чувствительность клеток резко увеличивается при каждом увеличении количества компонентов такого каскадного механизма. Эти свойства обуславливают наличие осцилляций, или периодических колебаний уровня фосфорилирования МАПК [207]. С использованием кинетических данных было предсказано, что эти осцилляции могут длиться от нескольких минут до нескольких часов. Уменьшение и увеличение уровня фосфорилирования МАПК приводят к возникновению волн при передаче сигнала от клеточной мембраны к ядру клетки, т.е. периодическое затухание и усиление сигнала [208]. Последнее обстоятельство обеспечивает возможность передачи сигнала на достаточно дальние расстояния. Изучение динамики активации ERK (extracellular-signal-regulated kinase), разновидностью которого является МАПК, показало, что ее кратковременная активация зависит от исходного уровня лигандов – ЭФР и фактора роста нервов. В то же время долговременная активация определяется конечной концентрацией факторов роста [209]. При этом динамика активации ERK зависит от динамики активации малых ГТФ-аз – Ras и Rap1, которые определяют временные и концентрационные параметры активации внутриклеточных эффекторов.

Моделирование является одним из способов оценки устойчивости системы и колебаний в ней при определенных заданных параметрах. Если модель «работает» лишь в узком диапазоне заданных параметров, то система, для описания которой она используется, слишком чувствительна и вряд ли может иметь биологическое значение. Изучение периодических колебаний в динамике сети передачи сигнала, опосредуемого ядерным фактором NF- κ B, показало, что количество, амплитуда и частота осцилляций заметно меняются с изменением параметров модели [210]. При этом между различными параметрами наблюдается синергизм действия, т.е. эффекты одного параметра прямо зависят от значения другого параметра и наоборот. Это демонстрирует сложность организации комплексных сетей и функциональную взаимозависимость их компонентов.

МЕТОДЫ МОЛЕКУЛЯРНОЙ ДИНАМИКИ

Белок-белковые взаимодействия часто сопровождаются конформационными изменениями белков, участвующих в них [211, 212]. Одним из способов изучения конформационной подвижности белков являются методы молекулярной динамики (МД), которые позволяют с максимально возможной на данный момент степенью детализации моделировать конформационные изменения и внутри- и межмолекулярные взаимодействия, возникающие как результат действия силовых полей на отдельные атомы в молекуле [213–217].

Методы молекулярной динамики основаны на расчетах траекторий движения атомов в молекулах путем решения системы классических уравнений движения с использованием законов Ньютона. В настоящее время использование МД методов доступно для систем, содержащих до 10^6 атомов. В мире имеется тенденция быстрого роста производительности и доступности суперкомпьютеров, что увеличивает интерес к МД методам [218, 219]. Моделирование с использованием современных силовых полей дает результаты, которые хорошо согласуются с данными физико-химических экспериментов. Использование в численных экспериментах явно заданного растворителя также не является более существенной проблемой [220]. Использование распределенных вычислительных сетей (гридов), состоящих из десятков тысяч параллельно работающих компьютеров, также увеличивает возможности МД методов [221]. Преимуществом такого подхода является возможность управления вычислениями с разных, порой географически удаленных, точек и взаимодействие между исследователями, осуществляющими расчеты с разных компьютеров. Все это способствует тому, что МД методы становятся мощным инструментом для изучения структуры и свойств биомакромолекул, таких как белки и нуклеиновые кислоты, а также их взаимодействий и механизмов функционирования.

Методы МД позволяют изучать детальные механизмы, лежащие в основе стабильности биомакромолекул, их разворачивания и сворачивания, переноса ионов через мембрану, конформационно-динамических изменений белков и пептидов, их внутреннюю динамику [222–225]. МД методы представляют собой также важный этап при компьютерном моделировании трехмерной (3D) структуры белков и их комплексов с лигандами на основе гомологии с другими белками с известной трехмерной структурой, полученной экспериментальным способом с использованием рентгеноструктурного анализа [226, 227]. Они позволяют произвести оценку правильности модели и обнаружить несоответствия в ней. Методы МД обеспечивают нагляд-

ное представление внутри- и межклеточных процессов и зачастую становятся хорошей основой для возникновения гипотез о функционировании клетки.

В настоящее время имеется множество примеров успешного применения МД для изучения функционирования биомакромолекул, включая белок-белковые взаимодействия. Примером использования МД методов может служить изучение механизмов взаимодействия SH3 доменов с различными пептидами, содержащими RXXR мотив. Методом МД были подтверждены экспериментально полученные данные о взаимодействующих партнерах. Релаксационный анализ показал, что SH3 домен обладает динамической подвижностью, оказывающей большое влияние на белок-белковое взаимодействие [181]. Конформационно-динамические изменения в белках Hck и c-Src, наблюдаемые в результате фосфорилирования С-концевых остатков тирозина, также были изучены методами МД. В этом исследовании было показано, что дефосфорилирование приводит к конформационным изменениям в молекулах белков и к нарушению взаимодействия между SH2 и SH3 доменами этих протеинкиназ с их последующей активацией [228]. Точечная замена аминокислоты, необходимой для взаимодействия между доменами, на глицин, также приводит к активации этих белков.

Еще одним примером использования методов МД может служить изучение влияния водного растворителя на белок-белковые взаимодействия [229]. В этих целях был осуществлен анализ 17 белковых комплексов, включающих белки, относящиеся к двум белковым семействам. Сравнение энергетических и динамических свойств производилось для аминокислотных остатков, взаимодействующих непосредственно или через одну или две молекулы воды. Эти исследования показали, что присутствие одной молекулы воды между взаимодействующими аминокислотными остатками (wet spots) приводит к заметному уменьшению их подвижности. Такие молекулы воды вносят значительный вклад в изменение свободной энергии белковых комплексов.

В настоящее время МД методы широко используются также для моделирования процессов, происходящих в мембране клеток [222, 223]. Моделирование взаимодействия двух сигнальных пептидов белка NS2 вируса гепатита С и клеточной мембраны показало, что один из пептидов состоит из двух жестких спиральных структур, соединенных шарнирным участком. Этот участок обеспечивает гибкость полипептидной цепи, благодаря чему пептид может относительно легко проникать в мембрану. Вычислительные эксперименты пока-

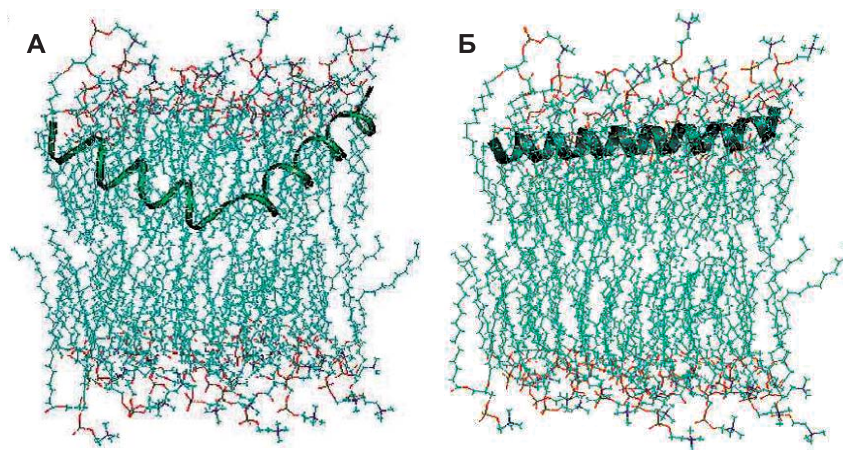


Рис. 4. Визуализация молекулярной динамики проникновения сигнального пептида вируса С: А – исходный пептид, Б – модифицированный пептид.

зали, что точечная замена аминокислотных остатков в шарнирном участке приводит к потере гибкости структуры. Пептид превращается в жесткую спираль, что резко ограничивает возможность его проникновения через мембрану клетки (рис. 4). Таким образом, с помощью компьютерного моделирования можно предсказывать изменения патогенных свойств вируса и использовать полученную информацию при создании противовирусной вакцины.

Методы МД в настоящее время также успешно применяются для моделирования биомолекул с потенциальным терапевтическим действием и компьютерного дизайна наноконтейнеров для направленной доставки биологически активных веществ [230]. Таким образом, методы МД представляют также мощный инструмент для разработки и создания лекарственных средств, которые могут быть использованы для лечения целого ряда заболеваний, включая опухолевые, инфекционно-аллергические и др.

VI. ЗАКЛЮЧЕНИЕ

Изучение клетки с точки зрения системной биологии предполагает интеграцию всех ее составляющих на разных уровнях организации – от атомного до клеточного и тканевого. Такая интеграция подразумевает взаимосвязь, взаимозависимость и взаимодействие этих составляющих, что лежит в основе их совместного, скоординированного функционирования. Одними из объектов исследования

системной биологии служат молекулярные сети, отражающие сложность организации биологических систем. Разные типы молекулярных сетей, включая генные, белковые, метаболические и сигнальные, используются для моделирования реальных процессов, происходящих в клетке.

В связи с тем, что белки, как ключевые биомакромолекулы, являются участниками почти всех внутри- и межклеточных процессов, моделирование живой клетки требует анализа всей совокупности данных динамической протеомики. Физические белок-белковые взаимодействия, существующие в клетке, образуют белковые сети. Для белковых сетей характерна «безмасштабность», модульность и иерархичность организации, наличие свойства «маленького мира». Последнее обстоятельство определяет быструю динамику процессов, описываемых с помощью белковых сетей.

Современные высокопроизводительные экспериментальные методы, используемые для изучения белок-белковых взаимодействий, не лишены недостатков. Результаты, получаемые разными группами авторов весьма противоречивы. В связи с этим одной из актуальных задач на современном этапе развития этой области представляется разработка подходов для получения более надежных и достоверных данных по белок-белковым взаимодействиям. Эти данные способствуют решению в долгосрочной перспективе двух фундаментальных задач системной биологии: а) выявлению динамических структурно-функциональных взаимоотношений на разных уровнях организации живого; б) созданию на этой основе динамической модели клетки (виртуальной клетки) и изучению влияния различных факторов на её функционирование. Все это, в свою очередь, способствует развитию новых подходов к изучению механизмов возникновения, диагностики и лечения заболеваний.

ЛИТЕРАТУРА

1. Betts, M.J., Russell, R.B. (2007) FEBS Lett., **581**, 2870–2876.
2. Priami, C., Quaglia, P. (2004) Brief. Bioinform., **5**, 259–269.
3. Ideker, T., Galitski, T., Hood, L. (2001) Annu. Rev. Genomics Hum. Genet., **2**, 343–372.
4. Liu, E.T. (2005) Cell, **121**, 505–506.
5. Friboulet, A., Thomas, D. (2005) Biosens. Bioelectron., **20**, 2404–2407.
6. Coveney, P. V., Fowler, P.W. (2005) J. R. Soc. Interface, **2**, 267–280.
7. You, L. (2004) Cell Biochem. Biophys. **40**, 167–184.
8. Fisher, J., Henzinger, T.A. (2007) Nat. Biotechnol., **25**, 1239–1249.
9. Шайтан К.В., Турлей Е.В., Голик Д.Н., Терешкина К.В., Левцова О.В., Федик И.В., Шайтан А.К., Лу А., Кирпичников М.П. (2006) Росс. хим. журн., **50**, 53–65.

10. Ferrel, J.E., Jr. (2009) *J. Biol.*, **8**:2.
11. Shaitan, K.V. (2003) In: *Stochastic Dynamics of Reacting Biomolecules* (ed. Ebeling, W. et al.) World Scientific. 283–308.
12. Milo, R. (2007) *Mol. Biosyst.*, **3**, 542–546.
13. Trinkle-Mulcany, L., Lamond, A.I. (2007) *Science*, **318**, 1402–1407.
14. Mayer, B.J. (1999) *Mol. Biotechnol.*, **13**, 201–213.
15. Houtman, J.C.D., Barda-Saad, M., Samelson, L.E. (2005) *FEBS J.*, **272**, 5426–5435.
16. Терентьев А.А., Молдогазиева Н.Т. (2006) *Биохимия*, **71**, 157–172.
17. Zaretsky, J.Z., Wreschner, D.H. (2008) *Translat. Oncogenomics*, **3**, 99–136.
18. Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., Weissman, J. S. (2003) *Nature*, **425**, 737–741.
19. Parrish, J.R., Gulyas, K.D., Finley, R.L., Jr. (2006) *Curr. Opin. Biotechnol.*, **17**, 387–93.
20. Stumpf, M.P.H., Thorne, T., de Silva, E., Stewart, R., An, H.J., Lapper, M., Wiuf, C. (2008) *Proc. Natl. Acad. Sci. (USA)*, **105**, 6959–6964.
21. Cusick, M.E., Klitgord, M.E., Vidal, M., Hill, D.E. (2005) *Hum. Mol. Genet.*, **14**, Spec No. 2, R171–81.
22. Mika, S., Rost, B. (2006) *PLoS Comput. Biol.*, **2**, e79.
23. Sprinzak, E., Sattath, S., Margalit, H. (2003) *J. Mol. Biol.*, **327**, 919–923.
24. Collins, S.R., Kemmeren, P., Zhao, X.C., Greenblatt, J.F., Spencer, F., Holstege, F.C., Weissman, J.S., Krogan, N.J. (2007) *Mol. Cell Proteomics*, **6**, 439–450.
25. Singhal, M., Resat, H. (2007) *BMC Bioinformatics*, **8**, 199.
26. Lee, H., Deng, H., Sun, F., Chen, T. (2006) *BMC Bioinformatics*, **7**, 269.
27. Causier, B. (2004) *Mass Spectrom. Rev.*, **23**, 350–367.
28. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., Seraphin, B. (2001) *Methods*, **24**, 218–229.
29. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., Seraphin, B. (1999) *Nat. Biotechnol.*, **17**, 1030–1032.
30. Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreau, M., Musk, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willem, A.R., Sassi, H., Nielsen, P.A., Rasmussen, K.J., Andersen, J.R., Johansen, L.E., Hansen, L.H., Jespersen, H., Podtelevnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B.D., Matthiesen, J., Hendrickson, R.C., Gleeson, F., Pawson, T., Moran, M.F., Durocher, D., Mann, M., Hogue, C.W., Figeys, D., Tyers, M. (2002) *Nature*, **415**, 180–183.
31. Zhou, M., Veestra, T.D. (2007) *Proteomics*, **7**, 2688–2697.
32. Domon, B., Aebersold, R. (2006) *Science*, **312**, 212–217.
33. Bartel, P., Fields, P. (1995) In *Methods: A Companion of methods in Enzymology*, **254**, 241–263.
34. Fields, S., Song, O. (1989) *Nature*, **340**, 245–246.
35. Ito, T., Ota, K., Kubota H., Yamaguchi, Y., Chiba, T., Sakuraba K., Yoshida, M. (2002) *Mol. Cell Proteomics*, **1**(8), 561–566.
36. Kaiser, P., Meierhofer, D., Wang, X., Huang, L. (2008) *Methods Mol. Biol.*, **439**, 309–326.
37. Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Shultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic,

- M., Ruffner, H., Merino, A., Klein, K., Hudak, Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelman, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Boymeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., Superti-Furga, G. (2002) *Nature*, **415**(6858), 141–147.
38. Burckstummer, T., Bennet, T., Preradovic, A., Schultze, G., Hantschel, O., Superti-Furga, G., Bauch, A. (2006) *Nat. Methods*, **3**(12), 1013–1019.
39. Tagwerker, C., Flick, K., Cui, M., Guerrero, C., Dou, Y., Auer, B., Baldi, P., Huang, L., Kaiser, P. (2006) *Mol. Cell. Proteomics*, **5**(4), 737–748.
40. Arnau, J., Lauritzen, C., Petersen, G.E., Pedersen, J. (2006) *Protein Expr. Purif.*, **48**, 1–13.
41. Domon, B., Aebersold, R. (2006) *Science*, **312**, 212–217.
42. Zhou, M., Veenstra, T.D. (2008) *Bio-techniques*, **44**, 667–670.
43. Abu-Farsha, M., Elisma, F., Figeys, D. (2008) *Adv. Biochem. Eng. Biotechnol.*, **110**, 67–80.
44. Karas, M., Bachmann, D., Bahr, D., Hillenkamp, F. (1987) *Int. J. Mass Spectrom. Ion Proc.*, **78**, 53–68.
45. Yip, T. T., Hutchens, T. W. (1992) *FEBS Lett.*, **308**, 149–53.
46. Ahram, M., Petricoin, E.F. (2008) *Biomarker Insights*, **3**, 325–333.
47. Engwegen, J.Y., Helgason, H.H., Cats, A., Harris, N., Bonfrer, J.M., Schellens, J.H., Beijnen, J.H. (2006) *World J. Gastroenterol.*, **12**, 1536–1544.
48. Poon, T.C. (2007) *Expert Rev. Proteomics*, **4**, 51–65.
49. Sanchez, J.C., Guillaume, E., Lescuyer, P., Allard, L., Carrette, O., Scherl, A., Burgess, J., Corthals, G.L., Burkhard, P.R., Hochstrasser, D.F.J. (2004) *Proteomics*, **4**, 2229–2233.
50. Hong, M.L., Jiang, N., Gopinath, S., Chew, F.T. (2006) *Clin. Exp. Pharmacol. Physiol.*, **33**, 563–568.
51. Feng, B., Dai, Y., Wang, L., Tao, N., Huang, S., Zeng, H. (2009) *Biologicals*, **37**, 48–54.
52. Conrad, U., Scheller, J. (2005) *Comb. Chem. High Throughput Screen*, **8**, 117–126.
53. Kumada, Y., Zhao, C., Ishimura, R., Imanaka, H., Imamura, K., Nakaniishi, K. (2007) *J. Biotechnol.*, **128**, 354–361.

53. Nickell, S., Beck, F., Korinek, A., Mihalache, O., Baumeister, W., Plitzko, J.M. (2006) *Nat. Rev. Mol. Cell Biol.*, **7**, 225–230.
54. Pepperkok, R., Ellenberg, J. (2006) *Nat. Rev. Cell Biol.*, **7**, 690–696.
55. Lucic, V., Forster, F., Baumeister, W. (2005) *Annu. Rev. Biochem.*, **74**, 833–865.
56. Baumeister, W. (2005) *FEBS Lett.*, **579**, 933–937.
57. Ortiz, J.O., Forster, F., Kurner, J., Linares, A.A., Baumeister, W.J. (2006) *J. Struct. Biol.*, **156**, 334–341.
58. Hober, S., Uhlen, M. (2008) *Curr. Opin. Biotechnol.*, **19**, 30–35.
59. Ponten, F., Jirstrom, K., Uhlen, M. (2008) *J. Pathol.*, **216**, 387–393.
60. Wallrabe, H., Periasamy, A. (2005) *Curr. Opin. Biotechnol.*, **16**, 19–27.
61. Gordon, G.W., Berry, G., Liang, X.H., Levine, B., Herman, B. (1998) *Biophys. J.*, **74**, 2702–2713.
62. Periasamy, A., Wallrabe, H., Chen, Y., Barroso, M. (2008) *Methods Cell Biol.*, **89**, 569–598.
63. Suhling, K., French, P.M., Philips, D. (2005) *Photochem. Photobiol. Sci.*, **4**, 13–22.
64. Jarvik, J. W., Adler, S. A., Telmer, C. A., Subramaniam, V., Lopez, A. J. (1996) *BioTechniques*, **20**, 896–904.
65. Jarvik, J. W., Telmer, C. A. (1998) *Annu. Rev. Genet.*, **32**, 601–618.
66. Sigal, A., Danon, T., Cohen, A., Milo, R., Geva-Zatorsky, N., Lustig, G., Liron, Y., Alon, U., Perzov, N. (2007) *Nat. Protoc.*, **2**, 1515–1527.
67. Sigal, A., Milo, R., Cohen, A., Geva-Zatorsky, N., Klein, Y., Alaluf, I., Swerdlin, N., Perzov, N., Danon, T., Liron, Y., Raveh, T., Carpenter, A.E., Lavah, G., Alon, U. (2006) *Nat. Methods*, **3**, 525–531.
68. Cohen, A.A., Geva-Zatorsky, N., Eden, E., Frenkel-Morgenstern, M., Issaeva, I., Sigal, A., Milo, R., Cohen-Saidon, C., Liron, Y., Kam, Z., Cohen, L., Danon, T., Perzov, N., Alon, U. (2008) *Science*, **322**, 1511–1516.
69. Shoemaker, B.A., Panchenko, A.R. (2007) *PLoS Comput. Biol.*, **3**, e43.
70. Ramirez, F., Schlicker, A., Assenov, Y., Lengauer, T., Albrecht, M. (2007) *Proteomics*, **7**, 2541–2552.
71. Kiemer, L., Costa, S., Ueffing, M., Cesarini, G. (2007) *Proteomics*, **7**, 932–943.
72. Li, D., Liu, W., Liu, Z., Wang, J., Liu, Q., Zhu, Y., He, F. (2008) *Mol. Cell Proteomics*, **7**, 1043–1052.
73. Braun, P., Tasan, M., Dreze, M., Barrios-Rodiles, M., Lemmens, I., Yu, H., Sahalie, J.M., Murray, R.R., Roncari, L., de Smet, A.S., Venkatesan, K., Rual, J.F., Vandenhaute, J., Cusick, M.E., Pawson, T., Hill, D.E., Tavernier, J., Wrana, J.L., Roth, F.P., Vidal, M. (2009) *Nat. Methods*, **6**, 91–97.
74. Skrabanek, L., Saini, H.K., Bader, G.D., Enright, A.J. (2008) *Mol. Biotechnol.*, **38**, 1–17.
75. Bader, S., Kuhner, S., Gavin, A.C. (2008) *FEBS Lett.*, **582**, 1220–1224.
76. Altaf-Ul-Amin, M., Shinbo, Y., Mihaara, K., Kurokawa, K., Kanaya, S. (2006) *BMC Bioinformatics*, **7**, 207.
77. Brohee, S., van Helden, J. (2006) *BMC Bioinformatics*, **7**, 488.
78. Eilbeck, K., Brass, A., Paton, N., Hodgman, C. (1999) In: «Intelligent Systems for Molecular Biology», AAAI Press, Palo Alto, **7**, 87–94.
79. Xenarios, I. E. E. F., Salwinski, L., Duan, X. J., Hegney, P., Kim, S.M., Eisenberg, D. (2002) *Nucleic Acids Res.*, **30**, 303–305.
80. Bader, G. D., Betel, D., Vogue, C. W. V. (2003) *Nucleic Acids Res.*, **31**, 248–250.
81. Breitkreutz, B.J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D.H.,

- Bachler, J., Wood, V., Dolinski, K., Tyers, M. (2008) *Nucleic Acids Res.*, **36** (Database issue), D637–640.
82. Teyra, J., Doms, A., Schroeder, M., Pisabarro, M.T. (2006) *BMC Bioinformatics*, **7**, 104.
83. Teyra, J., Paszkowski-Rogacz, M., Anders, G., Pisabarro, M.T. (2008) *BMC Bioinformatics*, **8**, 9.
84. Spirin, S., Titov, M., Karayagina, A., Alexeevski, A. (2007) *BMC Bioinformatics*, **23**, 3247–3248.
85. Gong, S., Yoon, G., Jang, I., Bolser, D., Dafas, P., Schroeder, M., Choi, H., Cho, Y., Han, K., Lee, S., Choi, H., Lappe, M., Holm, L., Kim, S., Oh, D., Bhak, J. (2005) *Bioinformatics*, **21**, 2541–2543.
86. Parrish, J.R., Yu, J., Liu, G., Hines, J.A., Chan, J.E., Mangolia, B.A., Zhang, H., Pacifico, S., Fotouhi, S., DiRita, V.J., Ideker, T., Andrews, P., Finley, R.L., Jr. (2007) *Genome Biol.*, **8**, R130.
87. Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., Emili, A. (2005) *Nature*, **433**, 531–537.
88. Yellaboina, S., Goyal, K., Mande, S.C. (2007) *Genome Res.*, **17**, 527–535.
89. Diaz-Mejia, J.J., Babu, M., Emili, A. (2009) *FEMS Microbiol. Rev.*, **33**, 66–97.
90. Su, C., Peregrin-Alvarez, J.M., Butland, G., Phanse, S., Fong, V., Emili, A., Parkinson, J. (2008) *Nucleic Acids Res.*, **36** (Database issue), D632–636.
91. Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., O'Shea, E.K. (2003) *Nature*, **425**, 688–691.
92. Dolinsky, T.J., Burgers, P.M., Karpus, K., Baker, N.A. (2004) *Bioinformatics*, **20**, 2312–2314.
93. Bader, G.D., Hogue, C.W. (2002) *Nature Biotechnol.*, **20**, 991–997.
94. Yu, H., Braun, P., Yildirim, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J.F., Dricot, A., Vazquez, A., Murray, R.R., Simon, C., Tardivo, L., Tam, S., Svrikapa, N., Fan, C., de Smet A.S., Motyl, A., Hudson, M.A., Park, J., Xin, X., Gusick, M.A., Moore, T., Boone, C., Snyder, M., Roth, E.P., Barabasi, A.L., Tavernier, J., Hill, D.E., Vidal, M. (2008) *Science*, **322**, 104–110.
95. Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., Punna, T., Peregrin-Alvarez, J.M., Shales, M., Zhang, X., Davey, M., Robinson, M.D., Paccanaro, A., Bray, J.E., Sheung, A., Beattie, B., Richards, D.P., Canadien, V., Lalev, A., Mena, F., Wong, F., Starostine, A., Canete, M.M., Vlasblom, J., Wu, S., Orsi, C., Collins, S.R., Chandran, S., Haw, R., Ristone, J.J., Gandi, K., Thompson, N.J., Musso, G., St Onge, P., Ghanny, S., Lam, M.H., Butland, G., Altaf-Ul, A.M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J.S., Ingles, C.J., Hughes, T.R., Parkinson, J., Gerstein, M., Wodak, S.J., Emili, A., Greenblatt, J.F. (2006) *Nature*, **440**, 637–643.
96. Formstecher, E., Aresta, S., Collura, V., Hamburger, A., Meil, A., Trehin, A., Reverdy, C., Betin, V., Maire, S., Brun, C., Jacq, B., Arpin, M., Bellaiche, Y., Bellusci, C., Benaroch, P., Bornens, M., Chanet, R., Chavrier, P., Delattre, O., Doye, V., Fehon, R., Faye, G., Galli, T., Girault, G.A., Goud, B., de Gunzburg, J., Johannes, L., Junier, M.P., Mirous, V., Mukherjee, A., Papadopoulos, D., Perez, F., Plessis, A., Rosse, C., Saule, S., Stoppa-Lyonnet, D., Vincent, A., White, M., Legrain, P., Wojcik, J., Camonis, D., Daviet, L. (2005) *Genome Res.*, **15**, 376–384.

97. Stanyon, C.A., Liu, G., Mangolia, B.A., Patel, N., Kuang, B., Zhang, H., Zhong, J., Finley, R.L., Jr. (2004) *Genome Biol.*, **5**, R96.
98. Lin, C.Y., Chen, S.H., Cho, C.S., Chen, C.L., Lin, F.K., Lin, C.H., Chen, P.Y., Lo, C.Z., Hsiung, C.A. (2006) *BMC Bioinformatics*, **7** (Suppl. 5):S18.
99. Yu, J., Pacifico, S., Liu, G., Finley, R.L., Jr. (2008) *BMC Genomics*, **9**, 461.
100. Drysdale, R., FlyBase Consortium (2008) *Methods Mol. Biol.*, **420**, 45–59.
101. Venkatesan, K., Rual, J.F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K.I., Yildirim, M.A., Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J.M., Cevik, S., Simon, C., de Smet, A.S., Dann, E., Smolyar, A., Vinayagan, A., Yu, H., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R.R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M.E., Roth, F.P., Hill, D.E., Tavernier, J., Wanker, E.E., Barabasi, A.L., Vidal, M. (2009) *Nat. Methods*, **6**, 83–90.
102. Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehpussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D.S., Zhang, L.V., Wong, S.L., Franklin, G., Li, S., Albala, J.S., Lim, J., Fraughton, C., Llamasas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R.S., Vandenhaute, J., Zoghbi, H.Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M.E., Hill, D.E., Roth, F.P., Vidal, M. (2005) *Nature*, **437**, 1173–1178.
103. Chen, J.Y., Sivachenko, A.Y., Bell, R., Kurscher, C., Ota, I., Sahasrabudhe, S. (2003) *Proc. IEEE Comput. Soc. Bioinform. Conf.*, **2**, 229–234.
104. Goodman, S.R., Kurdia, A., Ammann, L., Kakhniashvili, D., Daescu, O. (2007) *Exp. Biol. Med. (Maywood)*, **232**, 1391–1408.
105. Han, K., Park, B., Kim, H., Hong, J., Park, J. (2004) *Bioinformatics*, **20**(15), 2466–2470.
106. Brown, K.R., Jurisica, I. (2005) *Bioinformatics*, **21**, 2076–2082.
107. Cho, W.C. (2007) *Genomics Proteomics Bioinformatics*, **5**, 77–85.
108. Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P.-O., Bergamini, G., Crougton, K., Cruciat, C., Eberhard, D., Gagneur, J., Ghidelli, S., Hopf, C., Huhse, B., Mangano, R., Michon, A.-M., Schirle, M., Schlegl, J., Schwab, M., Stein, M.A., Bauer, A., Casari, G., Drewes, G., Gavin, A.-C., Jackson, D.B., Joberty, G., Neubauer, G., Rick, J., Kuster, B., Superti-Furga, G. (2004) *Nature Cell Biology*, **6**, 97–105.
109. Sheikh, M.S., Huang, Y. (2003). *Cell Cycle*, **2**, 550–552.
110. Escarcega, R.O., Fuentes-Alexandro, S., Garcia-Carrasco, M., Gatica, A., Zamora, A. (2007) *Clinical Oncology (Royal College of Radiologists (Great Britain))*, **1**, 154–161.
111. Lim, J., Hao, T., Shaw, C., Patel, A.J., Szabo, G., Rual, J.F., Fisk, C.J., Li, N., Smolyar, A., Hill, D.E., Barabasi, A.L., Vidal, M., Zoghbi, H.Y. (2006) *Cell*, **125**, 801–814.
112. Limviphuvadh, V., Tanaka, S., Goto, S., Ueda, K., Kanehisa, M. (2007) *Bioinformatics*, **23**, 2129–2138.
113. Camargo, L.M., Collura, V., Rain, J.C., Mizuguchi, K., Herjakob, H., Kerrien, S., Bonnert, T.P., Whiting, P.J., Brandon, N.J. (2007) *Mol. Psychiatry*, **12**, 74–86.
114. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabasi, A.L. (2000) *Nature*, **407**, 651–654.

115. Wagner, A., Fell, D.A. (2001) Proc. Roy Soc. London Series B, **268**, 1803–1810.
116. Galbraith, S.J., Tran, L.M., Liao, J.C. (2006) Bioinformatics, **22**, 1886–1894.
117. Li, Z., Shaw, S.M., Yedabnick, M.J., Chan, C. (2006) Bioinformatics, **22**, 747–754.
118. Гунбин К.В., Суслов В.В., Колчанов Н.А. (2008) Биохимия, **73**, 270–282.
119. Albert, R., Barabasi, A.L. (2002) Rev. Mod. Phys., **74**, 47–97.
120. Newman, M.E.F. (2003) SLAM Review, **45**, 167–256.
121. Scholtens, D., Vidal, M., Gentleman, R. (2005) Bioinformatics, **21**, 3548–3557.
122. Barabasi, A. L., Oltvai, Z. N. (2004) Nat. Rev. Genet., **5**, 101–113.
123. Watts, D. J., Strogatz, S. H. (1998) Nature, **393**, 440–442.
124. Mathias, N., Gopal, V. (2001) Phys. Rev. E. Stat. Nonlin. Soft Matter Phys., **63**(2 Pt 1):021117.
125. Albert, R. (2005) J. Cell Sci., **118**, 4947–4957.
126. Przulj, N., Corneil, D.G., Jurisica, I. (2004) Bioinformatics, **20**, 3508–3515.
127. Erdos, P., Renyi, A. (1959) Publ. Math., **6**, 290–297.
128. Li, D., Li, J., Ouyang, S., Wang, J., Wan, P., Zhu, Y., Xu, X., He, F. (2006) Proteomics, **6**, 456–461.
129. Rajarathinam, T., Lin, Y.H. (2006) Genomics Proteomics Bioinformatics, **4**, 80–89.
130. Jeong, H., Mason, S. P., Barabasi, A.L., Oltvai, Z.N. (2001) Nature, **411**, 41–42.
131. Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N., Barabasi, A.-L. (2002) Science, **297**, 1551–1561.
132. Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dimpelfeld, B., Edelmann, A., Heurtier, M.A., Hoffman, V., Hofert, C., Klein, K., Hudak, M., Michon, A.M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwineester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J.M., Kuster, B., Bork, P., Russell, R.B., Superti-Furga, G. (2006) Nature, **440**, 631–636.
133. Hart, G.T., Lee, I., Marcotte, E.R. (2007) BMC Bioinformatics, **8**, 236.
134. Luo, F., Yang, Y., Chen, C.F., Chang, R., Zhou, J., Scheuermann, R.H. (2007) Bioinformatics **23**, 207–214.
135. Wang, Z., Zhang, J. (2007) PLoS Comput. Biol., **3**, e107.
136. Spirin, V., Mirny, L.A. (2003) Proc. Natl. Acad. Sci. USA, **100**, 12123–12128.
137. Yook, S.H., Oltvai, Z.N., Barabasi, A.L. (2004) Proteomics, **4**, 928–942.
138. Jiang, T., Keating, A.E. (2005) BMC Bioinformatics, **6**, 136.
139. Chen, J., Yuan, B. (2006) Bioinformatics, **22**, 2283–2290.
140. Maslov, S., Sneppen, K. (2002) Science, **296**, 910–913.
141. Koyuturk, M., Kim, Y., Topkara, U., Subramaniam, S., Szpankowski, W., Grama, A.J. (2006) Comput. Biol., **13**, 182–199.
142. Zhenping, L., Zhang, S., Wang, Y., Zhang, X.S., Chen, L. (2007) Bioinformatics, **23**, 1631–1639.
143. Goodsell, D.S. (2005) Structure, **13**, 347–354.
144. Ho, E., Webber, R., Wilkins, M.R. (2008) J. Proteome Res., **7**, 104–112.
145. Wiwatwattana, N., Landau, C.M., Cope, G.J., Harp, G.A., Kumar, A. (2007) Nucleic Acids Res., **35**, D810–814.
146. Peirlioni, A., Martelli, P.L., Fari-selli, R., Casadio, R. (2007) Nucleic Acids Res., **35**, D208–212.

147. Kals, M., Natter, K., Thalinger, G.G., Trajanovski, Z., Kohlwein, S.D. (2005) *Yeast*, **22**, 213–218.
148. Han, J.D., Bertin, N., Hao, T., Goldberg, D.S., Berriz, G.F., Zhang, L.V., Dupuy, D., Walhout, A.J., Cusick, M.E., Roth, F.P., Vidal, M. (2004) *Nature*, **430**, 88–93.
149. Gursov, A., Keskin, O., Nussinov, R. (2008) *Biochem. Soc. Trans.*, **36**(Pt 6), 1398–1403.
150. Fraser, H.B. (2005) *Nat. Genet.* **37**, 351–352.
151. Tsai, T.Y., Choi, Y.S., Ma, W., Pomeroy, J.R., Tang, C., Ferrel, J.E. Jr. (2008) *Science*, **321**, 126–129.
152. Novak, B., Tyson, J. J. (2008) *Nat. Rev. Mol. Cell Biol.*, **9**, 981–991.
153. Platzter, A., Perco, P., Lukas, A., Mayer, B. (2007) *BMC Bioinformatics*, **8**, 224.
154. Hernandez, P., Huerta-Cepas, J., Montaner, D., Al-Shahrour, F., Valls, J., Gomez, L., Capella, G., Dopazo, J., Pujana, M.A. (2007) *BMC Genomics*, **8**, 185.
155. Goh, K.I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., Barabasi, A.L. (2007) *Proc. Natl. Acad. Sci. USA*, **104**, 8685–8690.
156. Wang, E., Lenferink, A., O'Connor-McCourt, M. (2007) *Cell Mol. Life Sci.*, **64**, 1752–1762.
157. Jonsson, P.F., Bates, P.A. (2006) *Bioinformatics*, **22**, 2291–2297.
158. Jonsson, P., Cavanna, T., Zicha, D., Bates, P. (2006) *BMC Bioinformatics*, **7**, 2.
159. Wachi, S., Yoneda, K., Wu, R. (2005) *Bioinformatics*, **21**, 4205–4208.
160. Benson, M., Breitling, R. (2006) *Curr. Mol. Med.*, **6**, 695–701.
161. Saito, S., Ojima, H., Ichikawa, H., Hirohashi, S., Kondo, T. (2008) *Cancer Sci.*, **99**, 2402–2409.
162. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T. (2003) *Genome Res.*, **13**, 2498–2504.
163. Campanaro, S., Picelli, S., Torregrossa, R., Colluto, L., Ceol, M., Del Prete, D., D'Angelo, A., Valle, G., Anglani, F. (2007) *BMC Genomics*, **8**, 383.
164. Loscalzo, J., Kohane, I., Barabasi, A.-L. (2007) *Mol. Syst. Biol.*, **3**:124.
165. Chu, L.-H., Chen, B.-S. (2008) *BMC Syst. Biol.*, **2**:56.
166. Nickolson, J.K. (2008) *Mol. Syst. Biol.*, **2**:52.
167. Snoussi, K., Halle, B. (2005) *Biophys. J.*, **88**(4), 2855–2866.
168. Чеботарева Н.А., Курганов Б.И., Ливанова Н.Б. (2004) *Биохимия*, **69**, 1239–1251.
169. Чеботарева Н.А. (2007) *Биохимия*, **72**, 1478–1490.
170. Takahashi, K., Vel Arjunan, S.N., Tomita, M. (2005) *FEBS Lett.*, **579**, 1783–1788.
171. Lin, J., Weiss, A. (2001) *J. Biol. Chem.*, **276**, 29588–29595.
172. Houtman, J.C., Higashimoto, Y., Dimasi, N., Cho, S., Yamaguchi, H., Bowden, B., Regan, C., Malchiodi, E.L., Mariuzza, R., Schuck, P., Apella, E., Samelson L.E. (2004) *Biochemistry*, **43**, 4170–4178.
173. Beyer, A., Wilhelm, T. (2005) *Bioinformatics*, **28**, 1610–1616.
174. Cesareni, G., Gimona, M., Yaffe, M. eds. (2004). *Modular protein domains*. Wiley-VCH, Weinheim, Germany.
175. Saito, S., Kashida, S., Inoue, T., Shiba, K. (2007) *Nucleic Acids Res.*, **35**, 6357–6366.
176. Tong, A.H., Drees, B., Nardelli, G., Bader, G.D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quoandam, M., Zucconi, A., Hogue, C.W., Fields, S., Boone, C.,

- Cesareni, G. (2002) *Science*, **295**, 321–324.
177. Kay, B.K., Williamson, M.P., Sudol, M. (2000). *FASEB J.*, **14**, 231–241.
 178. Sudol, M. (1998) *Oncogene*, **17**, 1469–1474.
 179. Zhou, S.Y. (1999) *Prog. Biophys. Mol. Biol.*, **71**, 359–372.
 180. Ferraro, E., Via, A., Ausiello, G., Helmer-Citterich, M. (2005) *BMC Bioinformatics*, **6**, Suppl. 4, S13.
 181. Hou, T., Chen, K., McLaughlin, W.A., Lu, B., Wang, W. (2006) *PLoS Comput. Biol.*, **2**, e1.
 182. Dalgarno, D.C., Botfield, M.C., Rickles, R.J. (1997) *Biopolymers*, **43**, 383–400.
 183. Cesareni, G., Panni, S., Nardelli, G., Castagnoli, L. (2002) *FEBS Lett.*, **513**, 38–44.
 184. Ghose, R., Shekman, A., Goger, M.J., Ji, H., Cowburn, D. (2001) *Nat. Struct. Biol.*, **8**, 998–1004.
 185. Solomaha, E., Szeto, F.L., Youself, M.A., Palfrey, H.C. (2005) *J. Biol. Chem.*, **280**, 23147–23156.
 186. Neduva, V., Russel, R.B. (2005) *FEBS Lett.*, **579**, 3342–3345.
 187. Semenza, J.C., Hardwick, K.G., Dean, N., Pelham, H.R. (1990) *Cell*, **61**, 1349–1357.
 188. Letourneur, F., Gaynor, E.C., Hennecke, S., Demolliere, C., Duden, R., Emr, S.D., Riezman, H., Cosson, P. (1994) *Cell*, **79**, 1199–1207.
 189. Harter, C., Wieland, F.T. (2004) *Proc. Natl. Acad. Sci. USA*, **95**, 11649–11654.
 190. Puntervoll, P., Linding, R., Gemund, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D.M., Ausiello, G., Branetti, B., Costantini, A., Ferre, F., Maselli, V., Via, A., Cesarini, G., Diella, F., Superti-Furga, G., Wyrwicz, L., Ramu, C., McGuigan, C., Gudavalli, R., Letunic, I., Bork, P., Rychlewski, L., Kuster, B., Helmer-Citterich, M., Hunter, W.H., Aasland, R., Gibson, T.J. (2003) *Nucleic Acid Res.*, **31**, 3625–3630.
 191. Tan, S.-H., Hug, W., Sung, W.-K., Ng, S.-K. (2006) *BMC Bioinformatics*, **7**, 502.
 192. Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., Noble, W.S. (2009) *Nucleic Acids Res.*, **37**, 125–128.
 193. Bailey, T.L., Williams, N., Misleh, C., Li, W.W. (2006) *Nucleic Acids Res.*, **34**, W360–373.
 194. Jonassen, J., Collins, F. Higgins, D.G. (1995) *Protein Sci.*, **4**, 1587–1595.
 195. Burgard, A.P., Moore, G.L., Maranas, C.D. (2001) *Metab. Eng.*, **3**, 285–288.
 196. Austin, R.S., Provart, N.J., Cutler, S.R. (2007) *BMC Genomics*, **8**, 191.
 197. Chung, J.L., Beaver, J.E., Scheef, E.D., Bourne, P.E. (2007) *Bioinformatics*, **23**, 2491–2492.
 198. Visvanathan, M., Breit, M., Pfeifer, B., Baumgartner, C., Modre-Osprian, R., Tilg, B. (2007) *Methods Inf. Med.*, **46**, 386–391.
 199. Suresh, B.C.V., Joo, S.E., Yoo, Y.S. (2006) *Biochimie*, **88**, 277–283.
 200. Conzelmann, H., Saez-Rodriguez, J., Sauter, T., Bullinger, E., Allgower, F., Gilles, E.D. (2004) *Syst. Biol. (Stevenage)*, **1**, 159–169.
 201. Ivanov, V.V., Ivanova, N.V. (2006) *Mathematical models of the cells and cell-associated objects. Elsevier, Mathematics in science and engineering. Series ed. Chui, C.K.*, 206 p.
 202. Gilbert, D., Fuss, H., Gu, H., Orton, R., Robinson, S., Vyshemirsky, V., Kurth, M.J., Downes, C.S., Dubitzky, W. (2006) *Brief. Bioinform.*, **7**, 339–353.
 203. Starbruck, C., Lauffenburger, D.A. (1992) *Biotechnol. Prog.*, **8**, 132–143.

204. Sible, J.C., Tyson, J.J. (2007) *Methods*, **41**, 238–247.
205. Allen, N.A., Chen, K.C., Shaffer, C.A., Tyson, J.J., Watson, L.T. (2006) *Syst. Biol. (Stevenage)*, **153**, 13–21.
206. Calzone, L., Thieffry, D., Tyson, J.J., Novak, B. (2007) *Mol. Syst. Biol.*, **3**, 131.
207. Kholodenko, B.N. (2000) *Eur. J. Biochem.* **267**, 1583–1588.
208. Markevich, N.I., Tsyganov, M.A., Hoek, J.B., Kholodenko, B.N. (2006) *Mol. Syst. Biol.*, **2**, 61.
209. Sasagawa, S., Ozaki, Y., Fujita, K., Kuroda, S. (2005) *Nat. Cell Biol.*, **7**, 365–373.
210. Ihekweaba, A.E., Broomhead, D.S., Grimley, R., Benson, N., White, M.R., Kell, D.B. (2005) *Syst. Biol. (Stevenage)*, **152**, 153–160.
211. Goh, C.S., Milburn, D., Gerstein, M. (2004) *Curr. Opin. Struct. Biol.* **14**, 104–109.
212. Tobi, D., Bahar, I. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 18908–18913.
213. Cornell, W.D., Cieplak, P., Bayly, C., Gould, I.R., Merz, K.M., Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. (1995) *J. Am. Chem. Soc.*, **117**, 5179–5188.
214. Ponder, J., Case, D. (2003) *Adv. Protein Chem.*, **66**, 27–33.
215. Шайтан К.В., Михайлюк М.Г., Леонтьев К.М., Сарайкин С.С., Беляков А.А. (2003) *Биофизика*, **48**, 210–216.
216. Шайтан К.В., Михайлюк М.Г., Леонтьев К.М., Сарайкин С.С., Беляков А.А. (2002) *Биофизика*, **47**, 411–419.
217. Шайтан К.В., Васильев А.К., Сарайкин С.С., Михайлюк М.Г. (1999) *Биофизика*, **44**, 668–675.
218. Gohlke, H., Kiel, C., Case, D.A. (2003) *J. Mol. Biol.*, **330**, 891–913.
219. Case, D.A., Cheatham, III T.E., Darden, T., Gohlke, H., Lou, R., Merz, Jr. K.M., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J. (2005) *J. Comput. Chem.*, **26**, 1668–1688.
220. Антонов М.Ю., Балабаев Н.К., Шайтан К.В. (2006) Свидетельство для официальной регистрации программ для ЭВМ N 2006613516.
221. Coveney, P.V. (2005) *Philos. Transact. A Math. Phys. Eng. Sci.*, **363**, 1707–1713.
222. Турлей Е.В., Шайтан К.В., Балабаев Н.К. (2005) *Биол. мембр.*, **22**, 491–502.
223. Шайтан К.В., Ли А., Терешкина К.Б., Кирпичников М.П. (2007) *Биофизика*, **52**, 560–475.
224. Молдогазиева Н.Т., Терентьев А.А., Казимирский А.Н., Антонов М. Ю., Шайтан К.В. (2007) *Биохимия*, **72**, 655–667.
225. Молдогазиева Н.Т., Шайтан К.В., Терешкина К.Б., Антонов М.Ю., Терентьев А.А. (2007) *Биофизика*, **52**, 611–624.
226. Mordvintsev, D. Yu., Polyak, Ya.L., Levtsova, O.V., Tourleigh, Ye.V., Kasheverov, I.E., Shaitan K.V., Utkin, Yu.N., Tsetlin, V.I. (2005) *Comput. Biol. Chem.*, **29**, 398–411.
227. Mordvintsev, D. Yu., Polyak, Ya.L., Kuzmin, D.A., Levtsova, O.V., Tourleigh, Ye.V., Utkin, Yu.N., Shaitan K.V., Tsetlin, V.I. (2007) *Comput. Biol. Chem.*, **31**, 72–81.
228. Young, M.A., Gonfloni, S., Supertifurga, G., Roux, B., Kuriyan, J. (2001) *Cell*, **105**, 115–126.
229. Samsonov, S., Teyra, J., Pisabarro, M.T. (2008) *Proteins*, **73**, 515–525.
230. Shaitan K.V., Tourleigh, Ye.V., Golik, D.N., Kirpichnikov, M.P. (2006) *J. Drug Deliv. Sci. Technol.*, **16**, 253–258.