

Слайд 1

НАЗВАНИЕ

Слайд 2

Лекарственные препараты разрабатывались человеком начиная с древних времен. Можно наблюдать эволюцию материалов, используемых для получения лекарств и методов тестирования. Огромный прогресс был достигнут в конце 20 века, при создании комбинаторных библиотек химических соединений и сверхпроизводительной системы автоматического тестирования миллионов таких соединений, так называемой High Throughput Screening.

Однако, даже наличие таких систем как UltraHigh Throughput Screening не позволяет быстро протестировать более или менее большие молекулы, полученные путем всех возможных комбинаций элементарных химических соединений, число которых может насчитывать биллионы. В начале 21 века быстро развивается направление –скрининг in silico.

Слайд 3

Мишени и механизмы действия лекарств могут быть расклассифицированы следующим образом.

РЕПЛИКАЦИЯ/ТРАНСКРИПЦИЯ ДНК

ДНК – алкилирующие агенты, вещества, связывающиеся по малой бороздке ДНК, интерколирующие агенты, фальшивые субстраты.

ФУНКЦИИ РНК

РНК-интерференция (подавление трансляции и ускорение деградации РНК)

СТРУКТУРА И ФУНКЦИЯ БЕЛКОВ

Процессы фолдинга белков после трансляции и создания биологически активных модулей (units) для полисубъединичных белков - ингибиторы белкового фолдинга

Активные центры ферментов - обратимые и необратимые ингибиторы

Транспортеры - uptake inhibitors

УРОВЕНЬ ГЕННОЙ СЕТИ И ПУТЕЙ ПЕРЕДАЧИ СИГНАЛА

Рецепторы – агонисты и антагонисты

СУПРАМОЛЕКУЛЯРНЫЕ СТРУКТУРЫ

Ионные каналы - блокаторы и активаторы

Слайд 4

Рациональный дизайн лекарств включает в себя следующие этапы:

Выработка биологической концепции

Computer-aided design: белковая кристаллография, ЯМР, поиск по базам третичных структур, de novo design

Экспериментальная оценка взаимосвязи структура-активность, QSAR (количественный анализ взаимосвязи структура-активность), молекулярное моделирование

Слайд 5

РАЗДЕЛ ПО ГЕННЫМ СЕТЯМ

Слайд 6

РАЗДЕЛ ПО МИКРО РНК

Слайд 7

После того, как выработана биологическая концепция и выбрана белковая мишень для действия препарата встают задачи на уровне дизайна лекарственных препаратов, основанного на анализе структурных данных (Structure-Based Drug design)

- 1) Распознавание каталитически-активных центров в третичной структуре белков
- 2) Распознавание сайтов связывания аллостерических эффекторов, таких как ионы металлов.
- 3) Распознавание сайтов белок-белковых взаимодействий.
- 4) Распознавание сайтов белок-ДНК взаимодействий.
- 5) Распознавание сайтов белок-РНК взаимодействий.
- 6) Распознавание ядер нуклеации третичной структуры для создания ингибиторов белкового фолдинга.

В мире разработано множество программ облегчающих решение этих задач.

Слайд 8

В частности, здесь представлены средства распознавания функциональных сайтов в пространственных структурах белков.

Основной недостаток в том, что они позволяют распознавать только некоторые типы функциональных сайтов, в частности не распознаются сайты, представленные малым количеством в базе PDB. Это связано с тем, что в основе методов положено построение статистической модели сайта. **ЧТОБЫ ЕЕ ПОСТРОИТЬ НЕОБХОДИМО ИМЕТЬ БОЛЬШОЕ КОЛИЧЕСТВО ПРЕДСТАВИТЕЛЕЙ КОНКРЕТНОГО САЙТА**

Слайд 9

Программы распознавания функциональных сайтов в пространственных структурах белков и быстрого предварительного докинга, разработанные в ИЦиГ

Слайд 10

Нами были разработаны базы данных и программы распознавания функциональных сайтов белков.

Здесь показана схема получения данных для PDBSite и PDBLigand.

База PDBSite накапливает данные по функциональным сайтам белков (<http://www.mgs.bionet.nsc.ru/mgs/systems/fastprot/>)

База PDBLigand содержит данные по лигандам для этих сайтов (<http://www.mgs.bionet.nsc.ru/mgs/systems/fastprot/>)

Эти две базы связаны друг с другом перекрестными linkами.

Слайд 11

Использовались два дополняющих друг-друга подхода для получения данных по сайтам:

Данные извлекались из PDB на основе информации, содержащейся в поле SITE, базы PDB, указывающей на аминокислотные остатки активных центров, сайтов посттрансляционных модификаций белков, сайтов связывания различных лигандов;

Сайты белок-белковых, белок-РНК и белок-ДНК взаимодействий определялись на основе анализа координат атомов белок-белковых, белок-ДНК и белок-РНК гетерокомплексов. В состав сайта включались аминокислотные остатки, контактирующие с лигандом (белок, РНК или ДНК). Два остатка считались контактными если любые три атома одного остатка находились на расстоянии менее 5 Å от трех атомов другого остатка.

Окружение сайта рассчитывалось по следующему критерию. В окружение сайта включались остатки, хотя бы один атом которых находился на расстоянии ближе 5 Å от любого атома, остатков сайта.

Слайд 12

Для каждого сайта в базе PDBSITE отводится отдельная запись, содержащая следующие данные, извлеченные из базы PDB:

1. координаты атомов функционального сайта;
2. функциональное описание сайта;
3. описание белка, которому принадлежит сайт;
4. список, окружающих сайт аминокислотных остатков;
5. физико-химические и структурные характеристики сайта и его окружения;
6. ссылки с известными молекулярно-генетическими базами данных (SWISS-PROT, EMBL, PDB и др.)

Слайд 13

PDBSITE содержит более 8000 сайтов

Сюда входят

Набор каталитически активных центров различных ферментов

Сайты посттрансляционной модификации белков

Сайты связывания ионов металлов

Слайд 14

Сайты связывания неорганических соединений

Сайты связывания органических соединений

Слайд 15

Сайты связывания лекарственных препаратов

Сайты белок-белковых взаимодействий

Сайты белок-ДНК взаимодействий

Сайты белок-РНК взаимодействий

Слайд 16

База данных PDBLigand содержит информацию по более 7000 лигандов
База PDBLigand содержит данные по низкомолекулярным лигандам, белкам, ДНК и РНК, которые связываются с сайтами из базы PDBSite.

В базе содержатся координаты атомов лигандов, а также их функциональное описание, извлеченные из базы PDB.

Каждая запись базы данных PDBLigand, содержащая информацию о отдельно взятом лиганде линкована с записью базы PDBSite, содержащей информацию о сайте связывания данного лиганда.

Слайд 17

На основе этих баз данных нами разработана PDBSiteScan - программа распознавания функциональных сайтов в третичных структурах белков

Алгоритм распознавания строится на автоматическом поиске пространственных фрагментов в третичной структуре белка, имеющих структурное сходство с функциональными сайтами из базы PDBSite.

Слайд 18

PDBSiteScan доступна по Интернету.

Слайд 19

Здесь мы видим пример распознавания каталитического центра в сериновой протеазе человека (1ELV) с помощью структурного выравнивания с одним из каталитических центров сериновых протеаз, содержащихся в базе PDBSite, иллюстрирующий работу программы PDBSiteScan.

Слайд 20

Здесь приведен другой пример распознавание сайта связывания иона меди в белке plastocyanin программой PDBSiteScan

Белок plastocyanin (PDB ID 1BXU) представлен в виде фрагмента поверхности. Остатки медь-связывающего сайта показаны как Stick. Зеленым цветом изображены остатки распознанного сайта в белке plastocyanin, Синим цветом изображены остатки сайта-темплата из базы PDBSite (ID 1B3ICU). Ион меди показан оранжевым шариком

Слайд 21

Пример распознавание сайта связывания иона SO₄ в белке hydrolase программой PDBSiteScan

Hydrolase (PDB ID 1EAX) представлена в виде фрагмента поверхности. Остатки SO₄-связывающего сайта показаны как Stick. Зеленым цветом изображены остатки распознанного сайта в hydrolase, синим цветом изображены остатки сайта-темплата из базы PDBSite (ID 1QL8AC2). Ион SO₄ показан шариками.

Слайд 22

После обнаружения необходимых сайтов на поверхности белков встает задача предсказания малых молекул, способных связываться с выявленными сайтами

Существует два основных компьютерных подхода:
виртуальный скрининг
выращивание молекулы в кармане (de novo structure generation).

Виртуальный скрининг осуществляется с помощью:
Правил (например, набор критериев биоактивности Липински), нейронных сетей, анализа фармакофоров, анализа сходства, поиска структурных классов, а также докинга и скоринга.

Слайд 23

Здесь приведены некоторые из известных программ облегчающих решение задачи скрининга.

Важное место среди таких программ занимает Докинг, который позволяет напрямую решать задачу межмолекулярного распознавания и взаимодействия. Однако, для эффективной работы таких методов как докинг необходимо иметь начальное положение лиганда в комплексе с молекулой. Задачу первичного выбора лиганда, способного связываться с белком и быстрого предварительного докинга решает PDBSiteScan и база PDBLigand.

Слайд 24

КРАТКАЯ СХЕМА АЛГОРИТМА БЫСТРОГО ПРЕДВАРИТЕЛЬНОГО ДОКИНГА

В базе PDBSite содержатся сайты-матрицы, для которых известны координаты атомов их комплексов с лигандами из базы PDBLigand. Быстрый докинг осуществляется путем переноса этих лигандов совместно с сайтом-матрицы при структурном выравнивании сайта-матрицы с белком.

Слайд 25

Для каждого лиганд-связывающего сайта из базы PDBSite дан лиганд в базе PDBLigand

ЩЕЛЧОК МЫШИ

Координаты атомов сайта и лиганда в этих базах даны для их комплекса. Поэтому преобразования ПОВОРОТ

ЩЕЛЧОК МЫШИ

И СДВИГ

ЩЕЛЧОК МЫШИ

необходимые для структурного выравнивания сайта с фрагментом белка применяются и для лиганда, таким образом, мы имеем комплекс белок-лиганд.

Слайд 26

Рассмотрим пример быстрого докинга для фосфотрансферазы с АДФ. На первом шаге проведем распознавание АДФ-связывающего сайта в phosphotransferase человека программой PDBSiteScan.

Слайд 27

На втором шаге эта программа выполнит быстрый докинг путем переноса (сдвига и поворота) АДФ- в пространстве с параметрами переноса, полученными для совмещения распознанного АДФ-связывающего сайта Phosphotransferase с сайтом матрицей из базы PDBSite (ID 1E9CADP).

Слайд 28

Другой пример приведен для белок-белкового докинга.

Анализировали ингибитор сериновых протеаз.

На первом шаге PDBSiteScan распознал сайт связывания альфа-химотрипсина в ингибиторе сериновых протеаз на основании структурного сходства участка в этом белке с сайтом связывания альфа-химотрипсина в другом ингибиторе протеаз - ovomucoid ингибиторе.

Зеленом цветом показаны остатки предсказанного сайта связывания белка в ингибиторе сериновых протеаз. Синим цветом показаны остатки реального сайта-«МАТРИЦЫ» ovomucoid ингибитора (PDB ID 1HJA), использовавшегося для распознавания сайтов.

Слайд 29

А затем выполнил Быстрый докинг альфа-химотрипсина с ингибитором триптазы (сериновых протеаз)

Слайд 30

Рассмотрим пример докинга белка с ДНК.

Показана пространственная структура комплекса человеческого SREBP-1A, связанного с промотором LDL-рецептора (PDB ID 1AM9).

Как видно, этот белок представляет из себя классическую спираль-петля-спираль лейциновую застёжку. Попробуем предсказать возможное положение ДНК-связывающего сайта в этом белке и построить гипотетический комплекс белок-ДНК. А затем сравним с реальными данными.

Слайд 31

О наличии ДНК-связывающего сайта в SREBP-1A можем судить по структурной гомологии участка этого белка с реальным сайтом-матрицей из базы PDB. В данном случае это ДНК-связывающий сайт мышиноного белка MAX.

Зеленым цветом отмечены остатки ДНК-связывающего сайта белка SREBP-1A, синим цветом отмечены остатки ДНК-связывающего сайта мышиноного белка MAX (PDB ID 1AN2), являющимся сайтом-матрицей для распознавания.

Слайд 32

В результате быстрого докинга мы имеем структуру предсказанного комплекса белка SREBP-1A с фрагментом ДНК. Этот фрагмент ДНК извлечен из реального комплекса мышинный белок MAX - ДНК.

Слайд 33

С помощью наложения структуры цепи ДНК, полученной в результате предварительного докинга, с кристаллографическими данными комплекса белок SREBP-1A /ДНК видим хорошее согласие геометрии этих молекул.

Слайд 34

Теперь рассмотрим некоторые Примеры применения PDBSiteScan для поиска сайтов в белках, как потенциальных мишеней для лекарственных препаратов.

В качестве первого примера рассмотрим белок p53.

В настоящее время хорошо известна роль белка p53 в регуляции клеточного цикла. Белок p53 может останавливать деление клетки или вызывать ее апоптоз (Lee et al, 1995).

Нарушения функции этого белка в результате мутаций часто приводят к развитию опухолей. Одной из причин таких нарушений могут быть возникшие в результате мутаций новые сайты.

Цель исследования

Поиск новых сайтов в природных мутантных формах белка p53 человека, возникших в результате мутаций, ассоциированных с развитием опухолей

Краткая схема исследования

Мутации белка p53 брались из базы данных SWISS-PROT. Проанализировано 181 аминокислотная замена в p53, ассоциированная с различными опухолями человека;

Третичные структуры мутантных форм ДНК-связывающего домена белка p53 человека моделировались с использованием сервера SWISS-MODEL;

Для каждой мутации получали третичную структуру и с помощью PDBSiteScan проводили поиск новых сайтов.

Слайд 35

Функция белка p53, как транскрипционного фактора осуществляется через связывание с ДНК. На данном рисунке представлена третичная структура комплекса ДНК-связывающего домена p53 с ДНК. Белок показан в виде поверхности. Зеленым цветом изображены аминокислоты Zn-связывающего сайта. Шариком показан ион цинка. Известно, что Zn служит аллостерическим эффектором, регулирующим связывание белка с ДНК.

Слайд 36

Мы обнаружили, что в результате мутации в позиции 245, приводящей к замене Глицина.

Слайд 37

На Цистеин.

Слайд 38

Возникает новый сайт связывания цинка, в который входит мутантный остаток цистеин. Положение сайта-матрицы при распознавании нового Zn-связывающего сайта в мутантном белке p53 показано синим цветом.

На рисунке видно, что положения аминокислот белка должны претерпеть конформационные изменения для того, чтобы образовать сайт, строго соответствующий матрице. Однако эти изменения могут находиться в пределах разрешенных углов вращения. Для того чтобы выяснить работоспособность нового сайта, необходимо проводить более детальный анализ с привлечением методов молекулярной динамики.

Слайд 39

Здесь мы видим новый Zn-связывающий сайт, если за положение аминокислот сайта в пространстве взять положение аминокислот сайта-матрицы. Такой подход может быть удобен для получения начального положения координат атомов при моделировании с помощью молекулярной динамики.

Слайд 40

В состав нового сайта вошли 2 аминокислоты из старого (нормального) сайта и мутантный цистеин.

Слайд 41

На рисунке видно, что новое положение цинка, отличается от нормального. Кроме того, связывание цинка сопровождается конформационными перестройками сайта. Таким образом, может нарушаться правильное связывание мутантного белка с ДНК.

На последующих четырех слайдах проиллюстрируем возможный механизм диффузии иона цинка из нормального сайта в мутантный и обратно, сопровождаемый конформационными перестройками.

Слайд 42

Слайд 43

Слайд 44

Слайд 45

Слайд 46

Можно предположить, что когда цинк находится в новом положении, то его нормальное положение заблокировано для других ионов цинка и белок функционирует неверно, что приводит к развитию опухоли. Таким образом, новый сайт связывания цинка в мутантном белке является мишенью для специфических ингибиторов, препятствующих его взаимодействиям с цинком, на основе которых может быть создан лекарственный антиопухолевый препарат.

Слайд 47

Следующий пример это поиск потенциальных мишеней для белков вируса Гепатита-С

В Hepatitis C virus NS3 helicase обнаружили потенциальный сайт связывания SO₄. Этот сайт расположен рядом с реальным сайтом в одном кармане молекулы. Кроме того,

потенциальный SO4-binding сайт контактирует с RNA-связывающим сайтом helicase. Мы предполагаем, что ион SO4 может мигрировать внутри кармана от реального сайта к потенциальному. При связывании иона SO4 с потенциальным сайтом может оказываться аллостерическое воздействие на связывание helicase с RNA. Можно предположить, что связывание SO4 с потенциальным сайтом служит сигналом для разрыва связи helicase с молекулой RNA. Таким образом, потенциальный SO4-binding сайт может быть объектом дальнейших исследований в качестве возможной мишени для создания специфических антивирусных ингибиторов.

Слайд 48

В другом вирусном белке, а именно в трансферазе вируса гепатита-С, нами был обнаружен сайт взаимодействия трансферазы с циклином человека. На основании этих данных мы построили потенциальный комплекс трансфераза – циклин.

Слайд 49

Здесь мы видим изображение этого же комплекса в виде молекулярных поверхностей. Однако следует заметить, что эти данные, основаны на предварительном анализе и требуют более детальных дальнейших исследований с применением программ докинга и молекулярной динамики.

Слайд 50

Еще одно направление наших исследований касается анализа взаимосвязи между структурой и активностью/свойствами белков.

Слайд 51

Использование рекомбинантных белков в качестве лекарственных препаратов имеет много привлекательных сторон. В частности, белки обладают многими функциями, поэтому, ингибирование одной функции может приводить к потере белком его остальных функций, а с помощью мутаций можно получить варианты белков с избирательно сниженной активностью для заданной функции.

Кроме того, для борьбы с некоторыми заболеваниями часто стоит задача обеспечения пролонгированного действия того или иного белка. Такой эффект, например, может быть обеспечен путем мутаций, препятствующих быстрому протеолизу. Так мутация Arg31 на Gln в TNF-alpha человека не производит эффекта на цитолитическую активность этого цитокина in a test in murine fibroblasts L929, но увеличивает его устойчивость к протеолитическим ферментам (Korepanova A.V. et al, 1994).

В этом направлении возникают задачи, связанные с прогнозированием мутаций, направленно меняющих активность/свойства белков без нарушения их третичной структуры.

Слайд 52

Нами разработана программа WebProAnalyst, позволяющая проводить поиск позиций в белках, мутации в которых способны направленно изменять активность или свойства белка. Метод основан на анализе корреляций структура-активность в множественном выравнивании семейств мутантных или родственных белков.

Слайд 53

В программе WebProanalyst реализован следующий метод количественного анализа взаимосвязи структура-активность в семействах гомологичных белков.

На вход подаются множественное выравнивание белков и численные значения активностей белков. Программа движет подвижное окно, заданной ширины вдоль множественного выравнивания. Это окно определяет аминокислотные остатки сайта, для которого проверяется наличие связи между физико-химическими свойствами аминокислот и значениями активностей. Для построения зависимости применяются множественный линейный регрессионный анализ, а также нейронные сети.

Сайты, на которых наблюдаются корреляции структура-активность мы называем активностью-модулирующими сайтами.

Слайд 54

Был проведен анализ распределения активность-модулирующих сайтов в структуре белков шести семейств по отношению к известным функциональным сайтам в этих белках.

Для этого вдоль множественного выравнивания для каждого семейства строился числовой профиль с максимальным значением корреляции. Максимальное значение корреляции отбиралось среди вычисленных корреляций для различных физико-химических свойств заданного сайта. Сайты задавались в множественном выравнивании с помощью подвижной рамки.

Для шести белковых семейств показано, что максимально высокие пики на профилях корреляций в основном расположены вблизи функциональных районов, но также существуют и отдаленные от них модулирующие сайты

Слайд 55

Здесь показано, как распределены активность-модулирующие сайты и функциональные сайты в третичной структуре панкреотического ингибитора трипсина из семейства ингибиторов протеиназ.

Можно видеть, что функциональный сайт и модулирующие сайты сближены, но также существуют и удаленные от функционального центра модулирующие сайты. В этих случаях, влияние на активность может реализоваться путем аллостерических эффектов.

Таким образом, активность модулирующие сайты могут быть мишенями, как для мутаций, так и низкомолекулярных соединений, способных специфически связываться с ними и аллостерически воздействовать на активность белка.

Слайд 56

Еще одно направление наших исследований, связано с анализом сетей взаимодействующих аминокислотных остатков в белках.

Слайд 57

Для моделирования мутаций в белках, направленно изменяющих их активность, необходимо учитывать сети взаимодействующих остатков, изменения в которых могут сказываться на стабильности белка, а так же влиять на его активность за счет кооперативного эффекта.

Если структура белка известна, то ее анализ позволяет определять такие взаимодействия на уровне структуры.

Если структура неизвестна, то такие взаимодействия приходится определять на уровне последовательностей белка. Однако, анализ на уровне последовательностей может дополнить анализ на уровне структуры.

Одним из подходов является сравнительный анализ наборов последовательностей белков, выполняющих одинаковые функции.

Предполагается, что для всех белков критические взаимодействия сохраняются.

Поэтому, критические взаимодействия могут быть выявлены на основе анализа консервативных характеристик последовательностей гомологичных белков.

Слайд 58.

Существует три возможных режима критических взаимодействий в белках в ходе эволюционного процесса.

Первый заключается в инвариантности аминокислотных остатков в наиболее важных позициях белка. Такие остатки определяют структуру и функцию белка и являются облигатными.

Второй режим заключается в возможности замен аминокислотных остатков на остатки со сходными физико-химическими свойствами. Такие небольшие вариации остатков могут обеспечивать специфичность белковых взаимодействий или модулировать активность/белка.

Третий режим заключается в заменах пар или групп аминокислотных остатков координированным образом, так, что суммарное значение физико-химического свойства в них (например, заряд) в результате мутаций остается постоянным.

Такие замены модулируют активность белка, взаимодействия между ними стабилизируют белковую структуру.

Слайд 59.

На рисунке приведен иллюстративный пример координированной фиксации замен в паре остатков, формирующих солевой мостик. Белки с одноименно заряженными парами остатков менее стабильны и под давлением отбора будут элиминированы в ходе эволюции. Белки с разноименными парами будут иметь селективное преимущество. Таким образом, набор родственных последовательностей современных белков будет содержать характерные паттерны замен аминокислот, свидетельствующих об их статистической зависимости.

Слайд 60

Для анализа таких координированных замен нами была разработана программа CRASP, Доступная в Интернет.

Слайд 61.

В нашей работе мы используем представление аминокислотных остатков в виде наборов чисел – величин физико-химических свойств. Предполагается, что эти свойства отражают основные типы взаимодействия между остатками в белке. В частности гидрофобность отражает антипатичность боковых групп к растворителю, индексы вторичных структур – к локальной упорядоченности полипептидной цепи, размера боковых групп – стерические взаимодействия, заряд – электростатические. В качестве меры зависимости позиций белка использовались коэффициенты корреляции.

Слайд 62.

На рисунке представлена схема пакета программ CRASP. На вход подается множественное выравнивание белков, затем эти данные преобразуются в матрицу чисел,

по которой оцениваются корреляции. Матрица коэффициентов корреляции выдается как в текстовом, так и графическом виде. Отдельный блок программ позволяет проводить анализ интегральных характеристик. Этот пакет программ имеет гипертекстовый интерфейс и расположен на сайте лаборатории теоретической генетики ИЦиГ.

Слайд 63.

Множественные выравнивания белков для анализа могут быть получены как на основе сравнения гомологичных последовательностей, так и в результате экспериментов по искусственной селекции белка.

Слайд 64.

Нами были исследованы особенности координированных замен в протеин-киназах. Эти белки принимают участие в каскадах передачи клеточных сигналов и активно исследуются как мишени для дизайна лекарств.

На рисунке приведена сеть взаимодействий, контролирующая рост клетки. Киназы являются важнейшими участниками этой сети.

Слайд 65.

Основной мишенью при разработке ингибиторов протеин-киназ является сайт связывания АТФ. В этот сайт входит АТФ связывающая петля из 20 аминокислот в N-терминальном домене протеин-киназ. Мы оценили консервативность суммарных величин нескольких физико-химических свойств остатков этой петли, с целью выявления наиболее консервативных свойств (по степени вклада координированных замен). Предполагалось, что взаимодействия, определяемые этими свойствами, являются наиболее важными для функционирования петли.

Слайд 66.

Для анализа использовалось множественное выравнивание 388 гомологичных протеин-киназ. Мы выбрали 20 свойств аминокислот из базы данных AAIIndex, которые включали гидрофобность, объем, заряд, вероятность встраивания во вторичную структуру, полярность и др.

Консервативность свойства оценивалась параметром λ (отношение вариабельности в приближении независимых замен остатков к реальной вариабельности свойства). Если $\lambda > 1$, то физико-химическое свойство является консервативным и определяет ключевые взаимодействия петли.

Слайд 67

По результатам анализа к консервативным характеристикам можно отнести шкалы гидрофобности, изоэлектрической точки, характеризующей заряд, и вероятности встраивания в бета-структуру. Их постоянство свидетельствует о функциональной важности соответствующих взаимодействий. Интересно отметить, что согласно существующим моделям остатки петли имеют бета-структурную конформацию, часть из них участвует в формировании гидрофобного кармана, в котором располагается аденозиновое кольцо АТФ, в то же время фосфатная группа АТФ несет на себе отрицательный заряд, который может участвовать в электростатических взаимодействиях. На основании полученных нами результатов можно предположить, что указанные особенности взаимодействия АТФ и протеин-киназ сохраняются в ходе эволюции, в том числе и за счет координированных замен.

Слайд 68.

Пример анализа консервативности суммарной величины изоэлектрической точки АТФ связывающей петли показан на слайде. На рисунке слева показаны остатки АТФ-связывающей петли киназы сАРКа, замены в парах которых происходят компенсаторным образом по отношению к заряду (значимые отрицательные корреляции заряда показаны голубыми стрелками). Показано, что в реальных белках дисперсия суммарного заряда петли значимо меньше, чем в сгенерированных случайных выборках с отсутствием корреляций (постоянство заряда).

Слайд 69.

Нами так же проанализированы координированные замены в ДНК-связывающих доменах семейства гомеодомен. Его структура представлена на рисунке. Выборка составила 372 последовательности, анализировались шкалы изоэлектрической точки.

Слайд 70.

В результате анализа был выделен кластер позиций гомеодомена, у которых все значимые коэффициенты корреляции отрицательны. На рисунке представлено положение этих остатков в структуре гомеодомена, а отрицательные зависимости показаны голубыми стрелками. Замены остатков в этом кластере являются компенсаторными по отношению к суммарному заряду. Статистический критерий показал, что суммарная величина изоэлектрической точки в этих позициях значимо консервативна. Ни в одной из 100000 случайных выборок (с отсутствием парных корреляций) величина дисперсии этой характеристики не была меньше реального значения. Интересно, что все остатки этого кластера расположены в районе контакта двух спиралей, а часть образует солевые мостики. Так как величины изоэлектрической точки характеризуют заряд остатка, мы предположили, что компенсаторные замены могут обуславливать постоянство суммарного заряда остатков в районе контакта спиралей 1 и 2, необходимого для стабилизации их упаковки.

Слайд 71.

Одним из новых методов белковой инженерии является искусственная селекция белков по их способности связываться с молекулами мишени (ДНК, органических соединений, других белков). Схема типичного эксперимента представлена на рисунке. Этот эксперимент позволяет из огромного числа пептидов (свыше миллиона) отбирать наиболее оптимальные по аффинности к субстрату.

[не обязательно]

Суть методики в следующем (см. рис. 7, стр. 34). Для небольшого пептида (размер обычно составляет несколько десятков аминокислот), моделирующего активный сайт белка, генерируется огромное число мутантов (клонов), представляющих максимально возможное сочетание наборов различных аминокислот в позициях пептида. Размер такой пептидной библиотеки обычно составляет 2×10^6 – 2×10^8 клонов. Далее с помощью методов геной инженерии эти пептиды встраиваются в белки оболочки бактериофага так, что оказываются на его поверхности (рис.7). Такие модифицированные фаги помещаются в раствор с магнитными частицами, которые покрыты молекулами субстрата. Происходит связывание пептидов с молекулами субстрата. Пептиды с наибольшим сродством к субстрату отделяются от остальных с помощью магнита и помещаются в колонию бактерий. Происходит заражение бактерий фагом и размножение тех фагов, которые имеют на своей поверхности пептиды с наибольшей аффинностью. После нескольких циклов эксперимента селектированные клоны выделяются, устанавливается их аминокислотная последовательность. Такие клоны содержат в позициях наборы аминокислот, оптимизированные для связывания субстрата.

Слайд 72.

Нами так же были проанализированы последовательности ДНК-связывающих доменов класса "цинковый палец", структура которого представлена на рисунке. Последовательности были получены в результате искусственной селекции в работе Чу и Ключа. Наиболее интенсивные корреляции наблюдались для величин изоэлектрической точки. Их матрица представлена на рисунке. Все значимые корреляции отрицательны и показаны синим цветом.

Слайд 73.

Нам удалось выделить кластер значимо коррелирующих позиций. Все они располагаются в районе контакта большой бороздки ДНК и альфа-спирали цинкового пальца. Так как значимые коэффициенты корреляции в этом кластере все отрицательны, мы предположили, что сумма значений изоэлектрической точки остатков будет консервативной. Это было подтверждено статистическими тестами с достоверностью выше 99.99%. Мы предположили, что возможная функциональная роль данной характеристики состоит в обеспечении неспецифического связывания ДНК и белка за счет электростатических взаимодействий.

Слайд 74.

Нами были так же проанализированы последовательности иммуноглобулин-связывающего домена белка L. Отбор последовательностей в эксперименте по искусственной селекции производился по критерию быстрой и правильной укладки белка. При анализе корреляций отобранных белков использовались два набора выравниваний.

Слайд 75.

Нами были проанализированы корреляции для 4 физико-химических свойств, объема боковых групп, изоэлектрической точки, полярности и гидрофобности. Пары остатков, замены в которых происходят зависимым образом по отношению к этим свойствам, показаны на рисунке. Видно, что значимые парные корреляции наблюдаются в районах стыковки сегментов вторичной структуры белка (терминальные участки вторичных структур).

Слайд 76.

Значимые корреляции наблюдаются и между парами позиций, принадлежащим разным участкам вторичной структуры.

Поскольку отбор последовательностей производился по скорости фолдинга, то можно предположить, что данные пары остатков обуславливают стабилизацию структуры белка, а взаимодействия между ними определяют взаимную укладку элементов вторичной структуры белка.

Слайд 77.

Данные по результатам экспериментов фаговой селекции собраны в базе данных ASPD, созданной в нашей лаборатории. Мы провели массовый анализ координированных замен для множественных выравниваний последовательностей из этой базы данных.

Рассчитывались корреляции по гидрофобности, полярности, объему и изоэлектрической точке. Проанализировано 153 записей из базы данных ASPD. Число проанализированных пар позиций составило 6308 для шкалы изоэлектрической точки и 6334 для трех других характеристик. Результаты анализа в виде электронной коллекции вошли в базу данных ASPD.

Слайд 78.

По результатам массового анализа мы провели исследование режимов координированных замен по 4 свойствам остатков. Известно, что замены остатков по отношению к сумме

свойств в паре позиций могут быть компенсаторными (отрицательные корреляции) и антикомпенсаторными (положительные корреляции). Поэтому в качестве характеристики режима эволюции рассматривалось отношение доли значимых положительных коэффициентов корреляции к доле значимых отрицательных для четырех интервалов значимости. Результаты анализа показали, что для объема и заряда (изоэлектрическая точка) замены преимущественно компенсаторные, а для гидрофобности и полярности – антикомпенсаторные. Эти результаты в целом согласуются с природой стерических, электростатических, гидрофобных и полярных взаимодействий. Они свидетельствуют, что выбранный нами метод адекватно отражает характер взаимодействий остатков.