

(19)



**Евразийское  
патентное  
ведомство**

(11) 016884

(13) B1

## (12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента

2012.08.30

(21) Номер заявки

200970180

(22) Дата подачи заявки

2007.07.24

(51) Int. Cl. G06F 19/16 (2011.01)

### (54) СПОСОБ ПРЕДСКАЗАНИЯ АГРЕГАЦИИ БЕЛКОВ И КОНСТРУИРОВАНИЯ ИНГИБИТОРОВ АГРЕГАЦИИ

(31) 60/821,553

(32) 2006.08.04

(33) US

(43) 2009.08.28

(86) PCT/GB2007/002789

(87) WO 2008/015384 2008.02.07

(71)(73) Заявитель и патентовладелец:

ЛОНЗА БАЙОЛОДЖИКС ПиЭлСи  
(GB)

(72) Изобретатель:

Колхофф Кай Й., Сурдо Хесус,  
Вендрусколо Микеле (GB)

(74) Представитель:

Медведев В.Н. (RU)

(56) DE GROOT NATALIA SÁNCHEZ ET AL.: "Prediction of hot spots of aggregation in disease-linked polypeptides" BMC STRUCTURAL BIOLOGY, BIOMED CENTRAL LTD., LONDON, GB, vol. 5, no. 1, 30 September 2005 (2005-09-30), page 18, XP021006180, ISSN: 1472-6807, whole document, in particular abstract, p.9, left. col. whole first paragraph, p.11., right col. second paragraph, p.12, left col.

PAWAR A.P. ET AL.: "Prediction of "Aggregation-prone" and "Aggregation-susceptible" Regions in Proteins Associated with Neurodegenerative Diseases", JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 350, no. 2, 8 July 2005 (2005-07-08), pages 379-392, XP004927302, ISSN: 0022-2836, whole document, in particular, abstract, p. 381, right col. - p. 382, left col., first paragraph, p. 382, left col. Third paragraph, p. 387, right col., last 2 paragraphs  
WO-A-2005045442

ARAVINDA SUBRAYASHASTRY ET AL.: "Structure and Assembly of Designed beta - Hairpin Peptides in Crystals as Models for beta - Sheet Aggregation", BIOCHEMISTRY; BIOCHEMISTRY FEB 24 2004, vol. 43, no. 7, 24 February 2004 (2004-02-24), pages 1832-1846, XP002456617, the whole document

B1

016884

(57) Изобретение относится к способам предсказания агрегации белков и конструирования ингибиторов агрегации. Один из таких способов предсказания потенциальных пептидных последовательностей, ингибирующих агрегацию белка, включает стадии: а) идентификации пептидной последовательности, образующей по меньшей мере часть области агрегации в белке-мишени; б) тестирования образования указанной пептидной последовательностью  $\beta$ -слоя; с) если на стадии б) достигнут положительный результат, извлечения соседних цепей слоя; д) идентификации остатков в цепях, расположенных рядом с указанной пептидной последовательностью, боковые цепи которых взаимодействуют с указанной пептидной последовательностью, т.е. тех остатков, которые образуют потенциальную пептидную последовательность, ингибирующую агрегацию белка. Также настоящее изобретение относится к способом конструирования соединений с применением остатков, идентифицированных в указанном выше способе; к соединениям, продуцируемым этими способами, и к компьютерным программам для осуществления указанных выше способов.

016884 B1

Изобретение относится к способам предсказания агрегации белков и конструирования ингибиторов агрегации. В частности, в качестве неисключающих примеров, оно относится к способам, которые помогают конструировать соединения для стабилизации белков относительно агрегации, таким образом, потенциально повышая время полужизни белков, снижая иммуногенность белков и повышая выход в системах трансляции *in vitro*.

### **Предпосылки изобретения**

Выявлено, что отложения неправильно свернутых белков в клетках или во внутриклеточном пространстве играют роль в ряде тяжелых медицинских нарушений, к которым относятся заболевания, такие как болезнь Альцгеймера, болезнь Паркинсона и диабет 2 типа. Средства, затрачиваемые системами медицинского обслуживания по всему миру для лечения этих медицинских состояний, являются огромными, так же как и влияние на жизни больных и их семей.

Вероятно, с увеличением продолжительности жизни количество случаев будет постоянно возрастать. Для решения этой нарастающей проблемы разрабатывают новые способы терапии на основе препятствия возможности образования агрегатов белков на ранних стадиях.

Типичный жизненный цикл белка в клетке начинается с синтеза полипептида в рибосоме и продолжается от исходно несвернутого состояния путем сворачивания, который может включать один или несколько промежуточных продуктов, до биологически активного нативного состояния белка. Для большинства белков это нативное состояние соответствует плотно свернутой конформации, хотя существуют некоторые исключения, одним из которых является  $\alpha$ -синуклеин, который в нативном состоянии является несвернутым (Uversky VN (2002) Natively unfolded proteins: A point where biology waits for physics, Protein Sci. 11:739-756). Жизненный цикл завершается денатурацией и деградацией.

Клетка обладает сложными механизмами контроля качества, которые способствуют процессу сворачивания белка. Первым из них является непосредственно рибосома. Во-вторых, белок поддерживают белки теплового шока и шапероны, которые действуют в качестве катализатора или промотора для правильного свертывания белка или для повторного свертывания неправильно свернутых белков (Evans MS, Clarke TF IV, Clark PL (2005) Conformations of Co-Translational Folding Intermediates, Prot, Pept. Let. 12(2): 189-195).

Если повторного сворачивания не происходит, белки с нарушенным сворачиванием перерабатываются системой убиквитин-протеосома. На первой стадии убиквитин присоединяется к нарушенным структурам. Эта метка маркирует полипептидную цепь для направления на деградацию, а эту задачу выполняет протеосома. Более подробное описание процессов сворачивания и нарушенного сворачивания можно найти в Dobson CM (2003) Protein folding and misfolding, Nature 426: 884-890 и Vendruscolo M, Zurdo J, MacPhee CE, Dobson CM (2003) Protein folding and misfolding: a paradigm of self-assembly and regulation in complex biological systems, Phil. Trans. R. Soc. Lond. A 361: 1205-1222). Однако контроль качества в клетке по многим причинам может нарушаться, что ведет к накоплению белков с нарушенным сворачиванием. Затем белки могут агрегировать, образуя плотные структуры, называемые амилоидными фибрillами с центральной областью, состоящей из сплошных скоплений  $\beta$ -слоев (Dobson CM (2005) Prying into prions, Nature 435: 747-749).

В живой ткани отложение белков (часто в форме амилоидных агрегатов) часто ассоциировано с различными заболеваниями, многие из которых связаны со старением. Например, эти заболевания включают нейродегенеративные заболевания, такие как болезнь Паркинсона, болезнь Альцгеймера и губкообразные энцефалопатии, а также системные (такие как амилоидозы легкой цепи иммуноглобулина или транстиретина) и периферические тканевые нарушения (такие как диабет 2 типа). У человека известно более 30 различных нарушений, ассоциированных с отложением белка.

В частности, в развитом мире, где продолжает непрерывно расти продолжительность жизни, серьезную проблему для общества представляет постоянно увеличивающееся количество людей, страдающих этими заболеваниями.

Согласно оценкам, в 2000 году только в США приблизительно 4,5 миллионов человек страдали болезнью Альцгеймера, а к 2050 году количество случаев может возрасти до 16 миллионов (Hebert LE, Scherr PA, Bienias JL, Bennett DA, Evans DA (2003) Alzheimer Disease in the U.S. Population: Prevalence Estimates C using the 2000 Census, Arch. Neurol. 60: 1119-1122). Риск людей заболеть этим нейродегенеративным заболеванием оценивается как 1 из 10 для людей старше 60 лет и практически 1 из 2 для людей старше 85 лет (Evans DA, Funkenstein HH, Albert MS, Scherr PA, Cook NR, Chown MJ, Hebert LE, Hennekens CH, Taylor JO (1989) Prevalence of Alzheimer Disease in a Community Population of Older Persons. Higher, than Previously Reported, Jama 262: 2551-2556). Влияние на систему здравоохранения является очень значительным, и некоторые авторы предсказывают, что нейродегенеративные заболевания могут стать основной причиной смерти (Lozano AM, Kalia SK (2005) New Movements in Parkinson's, Sci. Am., 291(1): 58-65).

Кроме того, склонность биомолекул к образованию агрегатов в растворе всегда была одной из основных проблем при создании лекарственных средств. Терапевтические молекулы должны быть как растворимыми, так и реакционноспособными и не должны образовывать агрегаты при введении в относи-

тельно высоких концентрациях или при хранении в течение длительных периодов времени. Во многих случаях, выявление условий, при которых такие полипептиды являются достаточно стабильными, требуют много времени и является дорогостоящим, а иногда даже невозможным при доступных в настоящее время способах. Таким образом, выявление способов препятствования процессу сворачивания в целях предотвращения образования агрегатов может повысить эффективность разработки лекарственных средств.

### **Сущность изобретения**

Таким образом, для предпочтительного содействия решению указанных выше задач желательно иметь возможность создавать соединения, которые взаимодействуют с патологическим белком или с терапевтической молекулой в растворе, так чтобы соединение конкурентно связывало и блокировало наиболее важные участки, запускающие процесс агрегации. Одно из приближений к достижению этой цели может включать конструирование образованных из пептидов молекул, которые будут препятствовать процессу агрегации.

Таким образом, в широком смысле, один из аспектов настоящего изобретения относится к способу конструирования пептидов, ингибирующих агрегацию белков, включающему идентификацию пептидных последовательностей, боковые цепи которых будут взаимодействовать со склонной к агрегации областью в белке-мишени.

Первый аспект настоящего изобретения относится к способу предсказания потенциальных пептидных последовательностей, ингибирующих агрегацию белков, включающему стадии:

а) идентификации пептидной последовательности, образующей по меньшей мере часть области агрегации в белке-мишени;

б) тестирования, образует ли указанная пептидная последовательность части  $\beta$ -слоя;

с) если на стадии б) достигнут положительный результат, извлечение соседних цепей того слоя;

д) идентификации остатков в цепях, расположенных рядом с указанной пептидной последовательностью, боковые цепи которых взаимодействуют с указанной пептидной последовательностью, т.е. тех остатков, которые образуют потенциальную пептидную последовательность, ингибирующую агрегацию белка.

Предпочтительно стадию тестирования проводят на множестве белков, гетерологичных белку-мишени. Таким образом, даже когда область агрегации в белке-мишени не образует часть  $\beta$ -слоя, можно идентифицировать соответствующие пептидные последовательности из сходных или идентичных последовательностей, встречающихся в других белках, которые образуют часть  $\beta$ -слоя.

Это тестирование можно проводить с использованием базы данных белковых структур и предпочтительно тестирования образования указанной пептидной последовательностью части  $\beta$ -слоя в любой из известных белковых структур, присутствующих в этой базе данных белков. В вариантах осуществления этого изобретения используют данные из Protein Data Bank (PDB) из Research Collaboratory for Structural Bioinformatics (RCSB), который содержит большое количество экспериментальных структур и теоретических моделей (Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank, Nuc. Acids Res. 28: 235-242). На момент 27 июня 2005 года в PDB было 31639 структур. Однако можно использовать другие структурные базы данных или банки данных.

Первой стадией предсказания пептида, ингибирующего агрегацию, является идентификация одной или нескольких областей агрегации в белке-мишени и пептидных последовательностей, образующих по меньшей мере часть этой области.

Предпочтительным способом идентификации этой области является использование профиля агрегации амилоида. Этот теоретический способ предсказания горячих точек агрегации в полипептидных цепях описан в DuBay KF, Pawar AP, Chiti F, Zurdo J, Dobson CM, Vendruscolo M (2004) Prediction of the Absolute Aggregation Rates of Amyloidogenic Polypeptide Chains, J. Mol. Biol. 341: 1317-1326 и в Pawar AP, DuBay KF, Zurdo J, Chiti F, Vendruscolo M, Dobson CM (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases, J. Mol. Biol. 350: 379-392. В способе предоставлен алгоритм, который, на основе присущих аминокислотам свойств, можно использовать для вычисления профиля склонности к амилоидной агрегации для любого белка.

После получения профиля амилоидной агрегации для белка-мишени можно идентифицировать области агрегации с учетом частей белка, для которых склонность к агрегации в профиле превышает заданную величину (например, 1).

Альтернативно или дополнительно, области агрегации можно идентифицировать посредством экспериментальных измерений, например, посредством систематической мутации каждого из остатков в пептиде или белке, или их фрагменте и синтеза фрагментов таких пептидов или белков, или фрагментов и анализа их склонности к агрегации в анализах *in vitro*.

Предпочтительно стадия тестирования включает подстадии: идентификации содержащейся в базе данных белковых структур группы белков, содержащей пептидные последовательности, родственные пептидной последовательности, образующей по меньшей мере часть области агрегации в белке-мишени; и идентификации в этой группе тех белков, в которых указанные родственные пептидные последова-

тельности образуют часть  $\beta$ -слоя.

Для повышения возможного количества совпадений (т.е. идентифицированных белков, в которых родственные пептидные последовательности образуют часть  $\beta$ -слоя) родственные пептидные последовательности предпочтительно включают представляющую интерес пептидную последовательность и фрагменты указанной пептидной последовательности.

Альтернативно или дополнительно, для повышения возможного количества совпадений родственные пептидные последовательности могут включать последовательности, которые включают консервативные замены одной или нескольких аминокислот в представляющей интерес пептидной последовательности.

В рамках этой стадии консервативная замена представляет собой замену, которая сохраняет связанные с агрегацией свойства подлежащей замене аминокислоты. В частности, консервативные заместители могут быть выбраны из аминокислот со склонностью к агрегации при pH 7,0, отличающимися друг от друга в пределах 0,2. Альтернативно или дополнительно консервативные заместители могут быть выбраны из остатков со сходными свойствами (например, основные, кислые, гидрофобные, полярные, ароматические).

Более предпочтительно родственные пептидные последовательности включают одну или несколько из указанных выше пептидных последовательностей как в прямом, так и обратном порядке.

В одном из вариантов осуществления настоящего аспекта подстадия идентификации группы белков, содержащихся в базе данных, включает сравнение родственных пептидных последовательностей с остатками, содержащимися в строках "SHEET" в файле PDB для рассматриваемого белка.

Альтернативно или дополнительно подстадия идентификации группы белков, содержащихся в базе данных, включает идентификацию тех остатков в родственных пептидных последовательностях, которые образуют друг с другом водородные связи.

Одним из способов, которым можно идентифицировать остатки, образующие друг с другом водородные связи, является вычисление евклидова расстояния между каждой парой остатков, которые отделены по меньшей мере тремя остатками, причем полагают, что водородные связи образуются, если это расстояние составляет менее 3,2 Ангстрем, а более предпочтительно, если это расстояние составляет менее 3,075 Ангстрем. В одном из вариантов осуществления эти евклидовы расстояния вычисляют с использованием записей "ATOM" из файла PDB для рассматриваемого белка.

Предпочтительно способ включает дополнительную стадию отображения идентифицированных пар остатков, которые образуют водородные связи друг с другом, и образованных водородных связей.

Для проверки достоверности результатов при любом из указанных выше способов идентификации групп белков результаты обоих способов можно использовать для перекрестной проверки остатков, идентифицированных другим способом. Такую перекрестную проверку, альтернативно или дополнительно, можно проводить другими способами идентификации групп белков.

Способ предпочтительно включает дополнительную стадию отображения остатков, идентифицированных на стадии d), а также он может включать стадию отображения информации о белке, в котором идентифицированы эти остатки. Это отображение также может принимать форму 3-мерного расположения идентифицированных остатков в  $\beta$ -слое.

Предпочтительно способ включает дополнительную стадию модификации остова соседних цепей и боковых цепей потенциальной пептидной последовательности, ингибирующей агрегацию белка, не принимающих участие во взаимодействии с указанной пептидной последовательностью, образующей часть области агрегации в белке-мишени, для максимизации взаимодействия с указанной пептидной последовательностью и повышения потенциальных свойств ингибирования агрегации потенциальной ингибирующей агрегацию последовательности. После идентификации одной или нескольких пептидных последовательностей или "матриц" можно сконструировать и синтезировать пептидную библиотеку, которая вносит вариабельность в область матрицы, не вовлеченные непосредственно во взаимодействие с рассматриваемой областью агрегации. Затем эту библиотеку можно подвергать скринингу с использованием запатентованных биохимических анализов агрегации и цитотоксичности для исследования изменений скоростей агрегации и токсичности белков в присутствии различных соединений.

Предпочтительно библиотеку создают посредством добавления модификаций к аминокислотной последовательности(ям)-кандидату, которые повышают свойства, такие как стабильность и растворимость.

Как описано в настоящем документе, способ, например, может включать: e) синтез пептидной библиотеки, причем члены указанной пептидной библиотеки содержат остатки, идентифицированные на стадии d), и f) определение аффинности связывания членов указанной библиотеки с белком-мишенью.

В библиотеке можно идентифицировать один или несколько пептидов, которые связываются с белком-мишенью с высокой аффинностью относительно контролей. Такие пептиды могут представлять собой пептиды-кандидаты, ингибирующие агрегацию белка.

Пептид, для которого выявлено, что он с высокой аффинностью связывается с мишенью, можно выделять, очищать и/или синтезировать.

Пептидные последовательности, предсказанные или идентифицированные, как описано выше, можно подвергать скринингу в отношении препятствия клеточным процессам (т.е. токсикологию). Например, способ может включать: тестирование наличия взаимодействия остатков, идентифицированных на стадии d), указанной выше, с одним или несколькими не являющимися мишенью белками, находящимися в Protein Data Bank (или любой другой базе данных белков), предпочтительно не являющимися мишенью белками, которые могут опосредовать важные клеточные процессы, такие как метаболические каскады, структурными белками гомеостаза ионов, белками, вовлеченными в ответ на стресс, регуляцию экспрессии генов, репарацию ДНК и т.д.

Стадию тестирования можно проводить на множестве белков, гетерологичных указанному белку-мишени, предпочтительно с использованием базы данных белковых структур. Например, тестирование можно проводить посредством идентификации в базе данных групп белков, которые содержат родственные пептидные последовательности, сходные с тестируемой пептидной последовательностью; и идентификации в группе тех белков, в которых родственные пептидные последовательности взаимодействуют с остатками, идентифицированными на стадии d), представленной выше.

Можно идентифицировать пептидные последовательности-кандидаты, ингибирующие агрегацию белка, которые не взаимодействуют с белками, опосредующими важные клеточные процессы.

Эффективность идентифицированной пептидной последовательности, как описано выше, можно определять в модели заболевания, связанного с нарушением сворачивания белка.

Модели заболеваний, связанных с нарушением сворачивания белка, хорошо известны в данной области. Пригодные модели включают клетки, которые сверхэкспрессируют склонный к агрегации белок, и модели с трансгенными животными, такие как модели на мышах или дрозофилах, которые сверхэкспрессируют склонный к агрегации белок и которые могут подвергаться другим воздействиям, таким как окислительный стресс, и т.д., или могут не подвергаться им.

Например, см.: Junn E, Mouradian MM, Human alpha-synuclein over-expression increases intracellular reactive oxygen species levels and susceptibility to dopamine, Neurosci Lett. 2002; 320: 146-50.

Lev N, Melamed E, Offen D, Proteasomal inhibition hypersensitizes differentiated neuroblastoma cells to oxidative damage, Neurosci Lett. 2006; 399: 27-32.

McGowan E, Eriksen J, Hutton M, A decade of modeling Alzheimer's disease in transgenic mice, Trends Genet. 2006; 22: 281-9.

Whitworth AJ, Wes PD, Pallanck LJ, Drosophila models pioneer a new approach to drug discovery for Parkinson's disease, Drug Discov Today. 2006 Feb; 11 (3-4): 119-26.

Склонные к агрегации белки включают следующие белки, предшественники или их фрагменты: а-синуклеин (либо дикого типа, либо любая из мутантных форм, ассоциированных с болезнью Паркинсона), хантингтин (а также другие белки с протяженными полиглутаминовыми или полиаланиновыми повторами), амилоидный бета-пептид (A $\beta$ 42), прионный белок, островковый амилоидный полипептид (hIAPP) или амилин, супероксиддисмутаза, Тац, альфа-1-антитрипсин и другие серпины, лизоцим, витронектин, кристаллинов, альфа-цепь фибриногена, аполипопротеин AI, цистатин C, гельзолин, лактоферрин, кератоэпителилин, кальцитонин, предсердный натрийуретический фактор, пролактин, кератин, медин (или полноразмерный лактадгерин), легкие цепи иммуноглобулинов, транстиреин (TTR), сывороточный амилоидный апобелок A (SAA), бета-2-микроглобулин, тяжелые цепи иммуноглобулинов или любой другой белок, ассоциированный с каким-либо нарушением сворачивания белка.

Можно определять способность пептидной последовательности, идентифицированной, как описано выше, к одному или более из следующих:

- i. Стабилизация белка в отношении агрегации;
- ii. Снижение скорости потери активности белка при хранении;
- iii. Снижение опосредуемой агрегацией иммуногенности белка;
- iv. Повышение выхода белков в системах трансляции *in vitro*;
- v. Повышение стабильности в растворе состава для терапевтического применения;
- vi. Ингибирование одного или более клеточных процессов;
- vii. Предотвращение олигомеризации или мультимеризации белка.

Дополнительные аспекты настоящего изобретения относятся к способам конструирования ингибиторов агрегации для лечения заболеваний, связанных с нарушением сворачивания белка, с использованием остатка(ов), идентифицированного на стадии d) указанного выше первого аспекта; конструирования соединений для стабилизации белков (например, биофармацевтических соединений, антител, ферментов и т.д.) в отношении агрегации в составах, носителях и других растворах с использованием идентифицированного остатка(ов); конструирования соединений для повышения срока хранения таких белков с использованием идентифицированного остатка(ов); конструирования соединений для снижения иммуногенности белков вследствие агрегации с использованием идентифицированного остатка(ов) и конструирования соединений для повышения выхода белков в системах трансляции *in vitro* с использованием идентифицированного остатка(ов).

Другой аспект настоящего изобретения относится к компьютерной программе, которая, при выпол-

нении на компьютере, осуществляет способ по любому из указанных выше аспектов.

Другой аспект настоящего изобретения относится к носителю компьютерных данных, содержащему компьютерную программу в соответствии с предыдущим аспектом.

Дополнительный аспект настоящего изобретения относится к компьютеру, настроенному на выполнение способа по любому из указанных выше аспектов способа. Предпочтительно этот компьютер представляет собой компьютер общего назначения, который настроен для доступа к базам данных, содержащим информацию об известных белках, для применения в целях получения предсказания. Такие базы данных могут храниться локально, например, на жестком диске или в памяти, но предпочтительно они хранятся дистанционно и доступны через каналы связи, такие как сеть или интернет.

#### **Краткое описание иллюстраций**

Варианты осуществления этого изобретения далее описаны на основании прилагаемых иллюстраций, где

- на фиг. 1 представлены профили амилоидной агрегации А $\beta$ 42 и  $\alpha$ -синуклеина при pH 7;
- на фиг. 2 представлен профиль амилоидной агрегации мутантной формы  $\gamma$ -кристаллина D при pH 7;
- на фиг. 3 представлены выходные данные программы PeptideSearch для остатков с 1 по 8 mutCRYD34-58;
- на фиг. 4 представлена блок-схема процесса поиска пептидов;
- на фиг. 5 представлены извлечения из файлов pdb\_seqres.txt и ss.txt, полученных из PDB;
- на фиг. 6 представлено предсказание водородных связей в трех соседних цепях в  $\beta$ -слое;
- на фиг. 7 представлен пример выходных данных программы для предсказания водородных связей между остатками;
- на фиг. 8 представлены обобщенные результаты для совпадения при поиске пептидов; и
- на фиг. 9 представлены профили склонности к агрегации для mutCRYD34-58, А $\beta$ 42 и  $\alpha$ -синуклеина с указанием положений, по которым были выявлены совпадения коротких пептидов в PDB.

#### **Подробное описание предпочтительных вариантов осуществления**

Далее описаны варианты осуществления настоящего изобретения. Способы настоящего варианта осуществления будут продемонстрированы на мутантной форме  $\gamma$ -кристаллина D (mutCRYD). Для целей иллюстрации также рассмотрены следующие два белка:  $\alpha$ -синуклеин, который вовлечен в болезнь Паркинсона, и А $\beta$ 42, который ассоциирован с болезнью Альцгеймера.

mutCRYD распространен в клетках хрусталика глаза человека. При нарушении сворачивания он может формировать агрегаты, которые проявляются в качестве катаракт, ведущих к неясному зрению или слепоте (Heon E, Priston M, Schorderet DF, Billingsley GD, Girard PO, Lubsen N, Munier FL (1999) The  $\gamma$ -Crystallins and Human Cataracts: A Puzzle Made Clearer, Am. J. Hum. Genet. 65: 1261-1267, и Dahm R (2004) Dying to see, Sci. Am., 291(4): 52-59).  $\gamma$ -Кристаллин D и мутантная форма отличаются тремя остатками, R58H, R36S и R14C, которые, как было выявлено, усиливают агрегацию (Pande A, Pande J, Asherie N, Lomakin A, Ogun O, King J, Benedek GB (2001) Crystal cataracts: Human genetic cataract caused by protein crystallization, PNAS, 98(11): 6116-6120).

$\alpha$ -Синуклеин представляет собой белок из 140 остатков, который встречается в качестве основного компонента в тельцах Леви (плотные отложения, встречающиеся в головном мозге пациентов с болезнью Паркинсона, которые могут вызывать нейродегенерацию - Spillantini MG, Schmidt ML, Lee VMY, Trojanowski JQ, Jakes R, Goedert M (1997)  $\alpha$ -Synuclein in Lewy bodies, Nature. 388: 839-840).

А $\beta$ 42 представляет собой небольшой гидрофобный пептид из 42 остатков, который может играть непосредственную роль в прерывании синаптической функции при болезни Альцгеймера, наиболее распространенной форме патологии нервов (Selkoe DJ (2002) Alzheimer's disease is a synaptic failure. Science 298:789-790).

Склонность к образованию плотных подобных  $\beta$ -слою структур, т.е. к агрегации, варьирует между белками, в зависимости от аминокислотного состава и последовательности.

Целью способа настоящего варианта осуществления была идентификация пептидных последовательностей-кандидатов, которые могут позволить снизить скорость образования и количество агрегатов посредством блокирования склонных к агрегации областей белков-мишеней.

#### **Идентификация областей агрегации**

Так называемые области, чувствительные к агрегации, в частях двух белков:  $\alpha$ -синуклеина и А $\beta$ 42 были предсказаны Pawar AP, DuBay KF, Zurdo J, Chiti F, Vendruscolo M, Dobson CM (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases, J. Mol. Biol. 350: 379-392.

В более общем смысле, в указанной выше статье предложен алгоритм, который позволяет вычислять профиль склонности к агрегации для белка и который, как было показано, приводит к результатам, которые тесно согласуются с широким диапазоном экспериментальных измерений. Для любой данной пептидной последовательности он позволяет вычислить склонность каждого отдельного остатка или усредненную склонность для окна из нескольких аминокислот. В табл. 1 приведена склонность для каждой аминокислоты отдельно при pH 7,0.

Таблица 1. Склонность к агрегации для каждой из 20 встречающихся в природе аминокислот отдельно при pH 7,0

Остаток	Склонность	Остаток	Склонность
Trp	2,92	Gly	-3,96
Phe	2,80	His	-4,31
Cys	1,61	Ser	-5,08
Tyr	1,03	Gln	-6,00
Ile	0,93	Asn	-6,02
Val	0,49	Asp	-9,42
Leu	-0,25	Lys	-9,55
Met	-1,06	Glu	-10,38
Thr	-2,12	Arg	-11,93
Ala	-3,31	Pro	-11,96

На фиг. 1 представлены профили склонности к агрегации для  $\alpha$ -синуклеина и А $\beta$ 42, вычисленные в соответствии с этим алгоритмом (и представленные в указанной выше статье). На фигуре слева представлен профиль склонности к агрегации А $\beta$ 42 при pH 7. Профиль агрегации для неизменной последовательности А $\beta$ 42 дикого типа нанесен на график в виде средней линии, вместе с максимальными и минимальными значениями склонности, полученными после тестирования в отношении какой-либо возможной мутации в каждом аминокислотном положении на протяжении последовательности. Линия  $Z_{\text{agg}}^{\text{prof}} = 1$  изображена для облегчения идентификации обеспечивающих агрегацию областей. Те же области, полученные экспериментальными способами, представлены в виде затемненных областей.

Аналогично, на фигуре справа представлен профиль амилоидной агрегации  $\alpha$ -синуклеина при pH 7, также включающий линию  $Z_{\text{agg}}^{\text{prof}} = 1$  для облегчения идентификации областей, способствующих агрегации. На этой фигуре светло-серым оттенком представлена большая область белка, которая, как полагают, является структурированной в фибриллы; более темным серым оттенком показана более высоко амилоидогенная область NAC; и штриховыми линиями показана область 69-79, которая, как было выявлено, является особенно амилоидогенным сегментом в этой области NAC.

Считают, что части белка, для которых показатель пересекает порог (например,  $Z_{\text{agg}}^{\text{prof}} = 1$  на фиг. 1), имеют наиболее высокую склонность к агрегации и полагают, что они являются центральными участками агрегата и образования фибрилл *in vivo*. Предсказанные центральные участки агрегации из фиг. 1 используют в этом варианте осуществления в качестве последовательностей-мишеней для детекции любых партнеров по взаимодействию, которые могут действовать в качестве ингибиторов.

Способы по настоящему варианту осуществления будут продемонстрированы на mutCRYD.

На фиг. 2 показан профиль склонности к агрегации mutCRYD, усредненный для семи остатков. Как в случае профилей склонности к агрегации, представленных на фиг. 1, для облегчения отличия между областями, которые следует рассматривать как чувствительные к агрегации, и другими частями белка, на уровне значения склонности к агрегации 1,0 добавлена прямая.

Для дальнейшего поиска были отобраны остатки с 34 по 58.

Эта область содержит два пика, для которых предсказаны горячие точки агрегации, и содержит две мутации. Для полного анализа рассматривают все области со склонностью к агрегации, превышающей 1, но выбранный короткий фрагмент является достаточным для демонстрации используемого способа. Исследованная последовательность из 24 остатков представляет собой SARVDSGCWMLYEQPNYSGLQYFL, и она обозначена как mutCRYD34-58.

#### Поиск - обзор

Хотя поиск последовательности в одном файле является стандартным способом и его можно проводить с помощью большинства редактирующих текст программ или средств командной строки, в файле для коротких последовательностей вполне могут существовать сотни или тысячи совпадений. Рассмотрение каждого из них вручную, поиск информации о вторичной структуре, а затем проверка web-сайта PDB для получения большей информации и применение 3D-средства для визуализации структуры и поиска взаимодействий между остатками в целях извлечения пептидов, которые тесно связаны водородными связями, может отнять недели или месяцы для каждого белка.

Таким образом, для управления процессом было разработано проприетарное программное обеспечение. Программное обеспечение было разработано в виде большого числа небольших инструментов, которые затем поэтапно были включены в одну комплексную программу, названную PeptideSearch.

Для написания PeptideSearch был выбран язык программирования PERL, поскольку все используемые данные были доступны в качестве обычных текстовых файлов, а PERL предоставляет мощные процедуры для анализа этих файлов, и извлечения и переработки содержащейся в них информации. Также

PERL является свободно доступным программным обеспечением, по этим причинам известным биоинформатикам.

На фиг. 4 представлена блок-схема программного обеспечения и способов по настоящему варианту осуществления и приведен обзор подробного объяснения способа, который представлен ниже.

PeptideSearch основан на вызовах ряда внешних программ для расширения его функциональности. Эти вызовы также представлены на фиг. 4.

При использовании программы, вводимыми пользователем данными являются последовательность-мишень и ряд параметров, а программа в итоге выдает краткую информацию о пептидах-кандидатах, включая трехмерную визуализацию пептидов и их партнеров по взаимодействию. Таким образом, время, требуемое для идентификации возможных ингибиторов любого белка, было снижено с потенциальных нескольких недель до нескольких часов.

PeptideSearch создает HTML-файл, называемый "result.html" для сохранения записей для каждого из совпадений. Этот и все другие файлы, создаваемые обеими программами и внешними инструментами, сохраняются в новой поддиректории. Название этой директории образуется из текущей даты и времени, так что результаты всех запусков программы остаются доступными на диске для последующего применения. При проведении поиска, PeptideSearch также выводит информацию о совпадениях в командную строку или на консоль, вместе с дополнительными сообщениями о состоянии. После завершения выполнения программы "result.html" содержит обзор всех выявленных совпадений, что упрощает отбор пептидов-кандидатов. Выходные данные программы после поиска остатков mutCRYD34-58 с 1 по 8 представлены на фиг. 3.

#### **Поиск - подробное описание**

В настоящем варианте осуществления поиск был основан на записях в Protein Data Bank (PDB) Research Collaboratory for Structural Bioinformatics (RCSB), поскольку этот свободно доступный источник содержит большое число экспериментальных структур и теоретических моделей. Таким образом, выявление всей искомой последовательности или ее части в одной из пептидных последовательностей в PDB дает доступ к родственной структуре, позволяя идентификацию вторичной структуры и партнеров по взаимодействию.

Первой стадией поиска является поиск последовательности-мишени во всех записях в PDB. Этого можно достичь копированием файлов PDB на локальный компьютер и анализом каждого из них по отдельности на последовательность остатков. Однако этот подход может быть неэффективным с точки зрения времени и пространства: на момент 27 июня 2005 года в PDB находилось 31639 структур, многие из которых представляют собой мультимеры с несколькими цепями.

Альтернатива, которая была принята для этого варианта осуществления, представляла собой однократное копирование отдельного файла, содержащего всю информацию о последовательностях и последующий поиск в нем для детекции совпадений. Такой файл можно получить с ftp-сервера PDB RCSB по адресу [ftp://ftp.rcsb.org/pub/pdb/derived\\_data](ftp://ftp.rcsb.org/pub/pdb/derived_data).

Файл, pdb\_seqres.txt, представляет собой список всех последовательностей PDB в формате FASTA. Аналогично, по тому же адресу можно найти файл, содержащий все вторичные структуры в PDB в формате FASTA, ss.txt. Версии файла, используемые для этого варианта осуществления, датируются 27 июня 2005 года.

Файл для вторичной структуры представляет собой первый и быстрый источник для проведения сопоставления последовательности и вторичной структуры и для детекции вовлечения совпадения в конформацию  $\beta$ -слоя.

На фиг. 5 представлены короткие извлечения из двух файлов: первые десять строк из seq\_res.txt и ss.txt соответственно. В табл. 2, ниже, представлено значение сокращений, используемых в ss.txt.

Таблица 2

Сокращение	Структурный элемент	Сокращение	Структурный элемент
H	Спираль	T	Поворот с водородными связями
E	Протяженная бета-цепь	S	Изгиб
G	310-спираль	B	Остаток в отдельном бета-мостике
I	Pi-спираль		

### **Очистка файлов с последовательностями и структурами**

Два указанных выше файла были получены автоматически с помощью PDB с использованием данных из всех записей PDB. Информация о вторичной структуре в файлах PDB определяется рядом различных программ.

Однако два файла имеют неравную длину. На извлечениях, представленных на фиг. 5, показана одна из причин этого: строки заголовка FASTA в файле с последовательностями являются более длинными, включая более подробную информацию.

Учитывая источник данных, можно ожидать, что для каждой последовательности должна существовать одна запись о структуре и наоборот. Также, сначала, кажется, что количество строк в файлах может быть идентичным, что очень удобно для сопоставления содержания двух файлов.

Однако ни одно из этих предположений не является истинным: количества строк не всегда совпадают. Скрипт PERL, просматривающий файл с извлечением всех строк заголовков FASTA, вернул 74560 записей для pdb\_seqres.txt и 69903 записей для ss.txt. Существует ряд других проблем, которые необходимо было решить перед тем, как использовать файлы в качестве входных данных. Например, в этих двух файлах не было согласовано использование ID. В файле с последовательностями используются ID, такие как 1tsv\_, а различные цепи в нем представлены как 1thl\_, lthl\_1, 1thl\_2 или как 1pr2\_A, 1pr2\_B, в то время как в ss.txt вместо этого используются 1TSV:\_1, 1THL:\_1, 1THL:\_1, 1PR2:A, 1PR2:B. Также в файле с последовательностями существует много ID, которые не имеют соответствующей записи в файле со вторичными структурами, и наоборот: такие записи, как 1J3W, 10G8 и IVGS, содержатся в ss.txt, но не в pdb-seqres.txt. Они были заменены на ID 2CVZ, 2BUH и 2CV4 из Protein Data Bank, который учитывается только в pdb\_seqres.txt.

Для удаления всех уникальных записей из обоих файлов понадобился ряд скриптов PERL. Полученные после этого файлы содержали 68383 совпадающих записи.

Тем не менее, остается потенциальная проблема, состоящая в том, что некоторые из последовательностей не совпадают по длине. Программу PeptideSearch настроили для указания таких несоответствий во время выполнения.

### **Сопоставление искомой последовательности с последовательностями**

Представляющие интерес экспериментальные структуры представляют собой структуры, в которых встречается аминокислотная последовательность, состоящая из искомой строки. Это означает, что необходим поиск точных совпадений строки в прямом направлении, для mutCRYD34-58:

SARVDSCGWMLEYQPNTSGLQYFL

а также для инвертированной последовательности, здесь:

LFYQLGSYNPQEYLMWCSDVRAS.

Очевидно, что нахождение в PDB точного совпадения последовательностей для длинных пептидов значительно менее вероятно, чем для коротких. Действительно, при поиске совпадений для полного белка, очень вероятно, что будет найдено только одно: сам белок.

Даже при сосредоточении на чувствительной области белка может быть желательным поиск совпадений с помощью участков из аминокислот, которые просто являются слишком длинными для того, чтобы дать точные совпадения в PDB. Поскольку структура мутантной формы γ-кристаллина D не включена в PDB, то не существует результата для полной последовательности из 24 остатков.

Для повышения выхода совпадений применяли два дополнительных способа.

Во-первых, белковую последовательность-мишень разделяли на две подстроки и партнеры по взаимодействию подвергали поиску по этим подстрокам. Это позволяет идентифицировать пептиды, которые связываются с частью последовательности. Затем их можно тестировать по отдельности, например, в сочетании с другими пептидами, или конструировать в более длинный пептид, соединяя несколько таких пептидов вместе.

В программе PeptideSearch, пользователь может установить минимальную длину  $n$  и максимальную длину  $m$  каждой искомой последовательности. После испытаний было выявлено, что длина искомой последовательности в диапазоне от 5 до 8 остатков приводит к достаточному количеству совпадений, выявляя партнеров, которые все еще были достаточно длинными для того, чтобы все еще быть пригодными. Однако также можно использовать подстроки другой длины.

Программа настроена для автоматического проведения полного поиска по всем непрерывным подстрокам исходной последовательности с длиной в данном диапазоне. Таким образом, общее количество искомых последовательностей  $q$ , которые ищут для полной искомой последовательности длины  $l$ , составляет:

$$q = \frac{(l-n+1) \cdot (l-n+2) - (l-m) \cdot (l-m+1)}{2}$$

Во-вторых, программа позволяет использовать регулярные выражения. Это позволяет вносить вариабельность в поисковую последовательность. Таким образом, PeptideSearch позволяет определять консервативные замены, которые учитываются при поиске по файлу с аминокислотными последовательностями.

Например, в соответствии с табл. 1, Gln (Q) и Asn (N) обладают сходной склонностью к агрегации, и, таким образом, одним вариантом увеличения количества пригодных совпадений может быть учет их обеих при поиске. Тогда, регулярное выражение в строке поиска, например, QANT может представлять собой [QN]A[QN]T. Аналогично можно использовать другие замены, исходя из склонности остатков к агрегации.

Два способа можно комбинировать, так что каждая возможная подстрока вплоть до минимальной длины превращается в регулярное выражение, которое учитывает консервативные замены, описанные выше.

### **Выявление элементов вторичной структуры**

Данные как от pdb\_seqres.txt, так и от ss.txtчитываются один раз при запуске программы и сохраняются в больших массивах. Затем записи из обоих массивов обрабатываются совместно. Поскольку, как отмечалось выше, длина аминокислотной последовательности и структура не всегда совпадают, номера строк нельзя использовать в качестве указателя. Вместо этого, для того, чтобы убедиться, что поиск учитывает записи, которые соответствуют друг другу, и что информация в обоих файлах остается синхронизированной, в качестве ориентира служат строки заголовков FASTA.

Если записи разделяются на несколько строк, их объединяют, удаляя разрывы строк между ними. Это создает две строковых переменных для каждой записи: одну для полной аминокислотной последовательности и одну для соответствующей вторичной структуры. Если на этом этапе две строки обладают неравной длиной, то данную информацию более не рассматривают как достоверную. В этом случае программа отображает предупреждение с извещением пользователя о несовместимости и предлагает исследование совпадения вручную.

PERL предоставляет простую для использования процедуру поиска строки последовательности на встречаемость искомой последовательности, либо поиском абсолютных совпадений, либо сопоставлением регулярных выражений. Если существуют какие-либо совпадения, то PeptideSearch извлекает их из строки последовательности и выравнивает их с соответствующим участком строки вторичной структуры. Это дает возможность первой визуальной проверки того, сколько остатков в подобранных последовательностях находятся в структуре протяженного  $\beta$ -слоя.

Поскольку окружение структуры также может представлять интерес, программа настроена также для вывода ряда остатков и информации о структуре как слева, так и справа от совпадения. Размер этого окна может быть установлен в параметрах программы.

В идеальном случае, эту информацию можно просто использовать для отбора всех тех записей, для которых файл вторичной структуры указывает на наличие  $\beta$ -слоев, и, затем, далее исследовать только их. Однако, для используемых файлов, это может привести к множеству ложноположительных, а также ложноотрицательных результатов. Основная причина этого состоит в том, что предсказание в ss.txt часто является неточным. Это было выявлено при сравнении выравнивания с существующей структурой в файле PDB.

Таким образом, получается, что лучше рассматривать каждое совпадение в последовательности независимо от информации в ss.txt. Это повышает количество совпадений, которые необходимо обрабатывать, далее требуя доступа к PDB, но также это приводит к тому, что ни один важный пептид-кандидат не упускается.

### **Получение файлов PDB**

После детекции искомой последовательности в одной из записей, требуется получить больше информации и необходимо преобразовать существующий файл PDB. PDB ID извлекается из строки заголовка FASTA. В PeptideSearch определяется локальная директория PDB.

Если файл PDB, который относится к этому ID, доступен в локальной директории, в сжатом или несжатом состоянии, тогда используется эта локальная копия. В ином случае PeptideSearch использует одну из его операций для автоматического доступа к зеркалу ftp-сервера RCSB Protein Data Bank в Cambridge Crystallographic Data Centre (CCBC), <ftp://pdb.ccsc.cam.ac.uk/rcsb/data/structures/divided/pdb/>, и загружает сжатую копию файла PDB. Затем этот .Z-архив извлекается посредством системного вызова программы WinRAR.

### **Обработка файлов PDB**

Перед открытием файла PDB программа проверяет, использовался ли тот же PDB ID для предсказания совпадения. Это может происходить в случае, когда существует несколько совпадений в одной аминокислотной последовательности или когда структура в файле представляет собой мультимер, так что существуют аналогичные совпадения на каждой цепи. Если ID является таким же, тогда данные все еще содержатся в массивах данных программы и их можно повторно использовать, экономя значительное количество времени. В ином случае, открывается локальная копия файла PDB, и данныечитываются в память.

Существуют две части каждого файла, которые представляют интерес. Первая часть представляет собой строки, которые можно использовать для идентификации перекрывания искомой последовательности с  $\beta$ -слоями. Они начинаются с метки "SHEET", которая показывает положения в  $\beta$ -слоях.

Вторая относится к записям "ATOM", которые дают координаты каждой части молекулы. Эта информация позволяет предсказывать водородные связи, поскольку знание того, какие остатки в действительности взаимодействуют, обеспечивает важную информацию и дает исследователю больше свободы при конструировании ингибиторных пептидов. Например, остатки, боковые цепи которых не вовлечены во взаимодействие, можно заменить другими, которые дают пептид с лучшими биохимическими свойствами.

Для каждой цепи в  $\beta$ -слое существует одна строка "SHEET". Она содержит индекс первого и последнего остатка в цепи. Если эта цепь не является первой в слое, тогда строка также содержит два остатка, где эта и предыдущая цепи совпадают, т.е. индексы одной пары остатков, для которых выявлена связь водородной связью. Полную информацию о записях "SHEET" и "ATOM" можно найти на [http://www.rcsb.org/pdb/docs/format/pdbguide2.2/guide2.2\\_frame.html](http://www.rcsb.org/pdb/docs/format/pdbguide2.2/guide2.2_frame.html).

Извлечение информации из этих строк и начальная и конечная точки совпадения в аминокислотной последовательности обеспечивают тест на перекрывание: если стартовый индекс слоя не превышает индекс конца совпадения и индекс конца слоя не является меньшим, чем начало совпадения, тогда существует перекрывание, в случае которого вычисляется количество перекрывающихся остатков.

В идеальном случае, количество перекрывающихся остатков не должно быть слишком маленьким, поскольку оно определяет длину потенциального ингибиторного пептида. Ингибитор, который состоит только из небольшого количества остатков, может оказаться не очень эффективным, и будет трудно сконструировать стабильный пептид с благоприятными свойствами. Программа позволяет пользователю определять порог и исключает все совпадения с перекрыванием, более коротким, чем этот порог из вывода в "result.html".

Затем  $\beta$ -цепь, которая перекрывается с искомой последовательностью, и соседние цепи (либо одна, либо две, в зависимости от положения первой цепи в слое) отображаются в выравнивании, в котором учитывается ориентация (параллельная или антипараллельная) и совмещение цепи.

Например, если пептидная последовательность представляет собой "ADDYYTATGHWYAT" и цепи проходят для поиска остатков с 1 по 4, с 6 по 10 и с 12 по 14, соответственно, в антипараллельной структуре  $\beta$ -слоя с совмещением в 3 и 7, и 9 и 13, тогда выравнивание дает в результате:

```
ADDY
HGTAT
YAT
```

Затем в "result.html" остатки искомой последовательности в этом выравнивании будут показаны красным цветом. Если поиск осуществлялся, например, на наличие "ATGHW", который был выявлен в индексах с 7 по 11 в пептидной последовательности, представленной выше, тогда остатки "HGTA" при выравнивании будут выделены цветом.

Выравнивание обеспечивает быстрый скрининг совпадений на совпадения с большим перекрыванием между искомой последовательностью и  $\beta$ -слоем. Оно показывает, сколько и какие из остатков формируют  $\beta$ -слой, и, следовательно, сколько и какие из остатков могли бы блокироваться, если бы их партнеры по взаимодействию были включены в короткие ингибиторные пептиды.

Несмотря на то что указанный выше способ поиска обеспечивает один путь идентификации пептидов-кандидатов, он не является полностью надежным при доступных текущих данных. В частности, индексы остатков, находящиеся в строках "SHEET", относятся к индексам в строках "ATOM", но не в записях "SEQRES" в том же файле. Protein Data Bank гарантирует, что записи "ATOM" протестированы на отсутствие ошибок. Однако данные в "pdb\_seqres.txt" были получены из записей "SEQRES", и их индексы часто не совпадают. Например, запись "SEQRES" может начинаться с MET, в то время как первая запись "ATOM" представляет собой SER, или последовательности могут быть одинаковыми, но запись "ATOM" начинается с индекса остатка 21, а не 1.

Это означает, что, хотя указанный выше способ, демонстрирующий выравнивание, может воспроизводить относительные положения цепи  $\beta$ -слоя корректно, идентификатор остатков может быть взят из неправильной части аминокислотной последовательности. Также это означает, что идентификация перекрывания с  $\beta$ -слоем может быть некорректной в первом положении: искомая последовательность в действительности может быть расположена в другом в структуре.

#### peptideAlign.class

Для того чтобы убедиться, что такое событие можно выявлять, при исследовании обобщенных результатов для этого совпадения позднее, был разработан следующий способ выравнивания с использованием записей "ATOM". Это обеспечило альтернативный способ выявления пептидов-кандидатов или перекрестной проверки результатов, полученных при поиске "SHEET", описанном выше.

Этот способ выравнивания также включает детекцию и отображение водородных связей между остатками, делая выравнивание значительно более информативным.

Для поиска в данных "ATOM" был изобретен следующий алгоритм: для всех пар атомов, которые отделены по меньшей мере тремя остатками и которые находятся в пределах представляющей интерес части  $\beta$ -слоя, вычисляется евклидово расстояние между ними (с использованием стандартной формулы

$$d = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$$

где  $d$  представляет собой эвклидово расстояние, и  $\Delta x$ ,  $\Delta y$  и  $\Delta z$  представляют собой разности соответствующих координат). Если вычисленное расстояние составляет менее 3,075 Ангстрем, то индексы в этой паре остатков перемещают в массив для хранения водородных связей. Предел длины был выбран из литературы и он является достаточно большим для детекции большинства водородных связей между атомами остова и боковыми цепями, и в то же время он является достаточно маленьким для поддержания малого количества ложноположительных результатов.

После завершения предсказания, информация о цепях и связях записывается в файл, называемый "bonds.txt". Этот файл используется в качестве файла для хранения промежуточных данных для объединения информации с другим приложением: "peptideAlign.class", основанным Java инструментом, который был создан для изображения графического выравнивания  $\beta$ -слоя. Хотя PERL обладает мощными процедурами для обработки больших количеств текста, Java оказывается легче для использования графических функций. Таким образом, свойства обоих языков программирования были объединены посредством использования системного вызова "PeptideSearch.pl" для внешней связи с PeptideAlign. Последний создает снимок экрана выравнивания, который затем программа PERL включает в обзор текущего совпадения в "result.html".

Применение графических изображений облегчает указание на водородные связи между остатками, которое можно осуществлять простым проведением линий. PeptideAlign не учитывает непосредственно антипараллельную форму  $\beta$ -слоя, но графические выходные данные указывают на то, следует ли инвертировать ориентацию цепи, поскольку в этом случае водородные связи пересекаются. Например, синими линиями, указывающими на водородные связи на фиг. 6, показано, что ориентацию средней цепи необходимо инвертировать. Для правильной ориентации цепей и сдвига их относительно друг друга для минимизации общей длины связей используют дополнительный алгоритм.

PeptideAlign считывает информацию о пептидах и связях из "bonds.txt" вместе с названием для ожидаемого выходного файла с изображением. Названия файлов с изображениями имеют форму hitX.png, где X указывает на текущий номер совпадения. Программа переводит информацию в файле в графическое представление и возвращает управление PeptideSearch. Пример выходных данных представлен на фиг. 7: файл содержит путь и название файла, под которым предполагается сохранить снимок экрана, количества атомов и связей в файле с последующими ID остатков для всех атомов и информацией о том, какие из них соединены связями; графическое представление справа показывает, какие остатки связаны друг с другом.

### Минимальный файл PDB

Выравнивание, проводимое поиском "ATOM", является корректным, благодаря отбору по PDB. Если оно соответствует выравниванию, проводимому PeptideSearch, тогда запись "SEQRES" также согласуется. В ином случае, пользователю необходимо будет проверить графический обзор визуально в отношении перекрывания. Если оно отсутствует, требуется проверка файла PDB вручную, чтобы видеть существует ли отличающийся  $\beta$ -слой, который содержит часть искомой последовательности или всю ее.

Следует отметить, что, в настоящее время, невозможно создать аминокислотную последовательность заново из записи "ATOM" и провести ее поиск в отношении наличия искомой строки. Это является следствием того, что в записи часто отсутствуют остатки. Например, в pdblp9w.ent после остатка 108 следует остаток 121. Брешь можно объяснить рядом причин, например недостаточными экспериментальными данными, низким разрешением ЯМР или рентгенографии, или высокоподвижными боковыми цепями, которые приводят к размытым сигналам, что делает невозможным точное определение положений атомов. Другое объяснение этого просто состоит в том, что в научной работе, в которой получены данные, иногда интерес представляет только конкретная часть белка, например активный участок, и, таким образом, не предоставляется какого-либо предсказания об оставшейся структуре. Однако, когда доступны полные данные, такой способ поиска может быть подходящей отдельной альтернативой поиску по "SHEET", рассмотренному выше.

После нахождения совпадения в  $\beta$ -слое можно видеть, какие из остатков взаимодействуют, и построить остав для конструирования пептидов-кандидатов. Например, остатки, которые не взаимодействуют с остатками в искомой последовательности, можно заменять другими аминокислотами, обеспечивая больший диапазон пептидов, подлежащих конструированию и тестированию.

Например, в коротком выравнивании, представленном на фиг. 6, SARV взаимодействует с VTY. Водородные связи предсказаны между A и V, и V и Y. Однако T по-видимому не взаимодействует с искомой последовательностью вообще, так что пептид VXY можно использовать в качестве остова для создания пептидов-кандидатов, причем X является заменяемым различными аминокислотами. В реальных условиях такой остав, наиболее вероятно, является слишком коротким, чтобы быть эффективным, таким образом, предпочтительным является поиск пептидов с длинами из 5 остатков и более.

Для каждого успешного совпадения PeptideSearch создает новый файл .pdb, содержащий координаты всех атомов, входящих в  $\beta$ -слой. Для этого он извлекает эту часть из полной структуры, которая со-

держит  $\beta$ -цепи с фрагментом искомой последовательности и ее партнерами по взаимодействию. Файл можно открыть с помощью любой из обычных программ для просмотра 3D-изображений, таких как Rasmol или vmd, и он позволяет эффективно проверить предсказанные водородные связи и пространственную ориентацию боковых цепей.

#### **Peptide3D.class**

Для предварительного просмотра структуры в новом файле .pdb в записи "results.html" включаются снимки экрана. Для осуществления этого необходима программа для просмотра 3D-изображений, которая может обеспечить прорисовку молекул, помещение их в информативной ориентации в пространстве, и получение снимка экрана, все это без необходимости взаимодействия с пользователем.

#### **Results.html**

После завершения PeptideSearch анализа всех возможных подстрок и мутаций во всех записях в pdb\_seqres.txt он открывает окно браузера посредством системного вызова и предоставляет все совпадения, которые были выявлены, в виде обзора. Этот обзор предоставляет ссылки на страницу отчета PDB для ID PDB, связанным с каждым совпадением, и ссылку на Peptide3D, которая применяется в качестве апплета для просмотра структуры в формате 800×600 пикселей.

#### **Результаты**

Выполнение программы с mutCRYD34-58 в качестве входных данных, с установкой минимальной и максимальной длины подстрок на пять, с обеспечением отсутствия мутаций и с исключением всех совпадений, в которых перекрывание с  $\beta$ -слоем является меньшим четырех, дало 153 точных совпадения в pdb seqres.txt, 47 из которых были в  $\beta$ -слоях.

Только восемь из этих результатов были длиннее трех остатков и последние были включены в "result.txt". Фактически, этот набор в действительности дает только два отличающихся совпадения, поскольку первые два и последние шесть совпадений являются практически идентичными, так как их источником являются различные цепи мультимера.

Первое из этих совпадений относится к последовательности SARVD (остатки mutCRYD34-58 с 1 по 5), которая была выявлена в предполагаемом транскрипционном репрессоре системы отщепления глицина (PDB ID 1U8S). Для этого совпадения результат представлен на фиг. 3. Перекрывание из четырех остатков представлено на фиг. 6.

Второе совпадение было выявлено для последовательности PNYSG (остатки mutCRYD34-58 с 15 по 20) в Endo-1, 4-Beta-Xylanase II. Обобщенные результаты для этого совпадения из файла HTML представлены на фиг. 8.

На фиг. 9 цветом выделены два совпадения в профиле склонности к агрегации mutCRYD34-58. Расположения точных совпадений рядом с горячими точками агрегации двух белков  $\alpha$ -синуклеина и А $\beta$ 42, или вблизи них, также представлены в соответствующих профилях на фиг. 9. Для  $\alpha$ -синуклеина длина подстрок была установлена в диапазоне от 6 до 7, а для А $\beta$ 42 она была установлена на 5.

Для mutCRYD первое совпадение может быть не очень эффективным в качестве кандидата для ингибитора агрегации, поскольку он не находится в области, чувствительной к агрегации. Однако второе совпадение имеет потенциал в отношении блокирования в пептиде иницииации образования агрегатов в этой области. Для других двух белков, каждый из партнеров по взаимодействию, соответствующий показанным совпадениям, обладает потенциалом к действию в качестве ингибитора агрегации. Количество совпадений для  $\alpha$ -синуклеина значительно возрастает при установке длины подстроки на 5.

Результаты анализа  $\alpha$ -синуклеина являются объектом дополнительной патентной заявки настоящего заявителя.

С использованием описанного выше программного обеспечения сконструирована серия пептидов для взаимодействия с участками 61-66 (EQVTN) и 71-76 (VTGVT)  $\alpha$ -синуклеина.

В этом анализе идентифицирован пептид, состоящий из последовательности D-аминокислот, которая взаимодействует с областью  $\alpha$ -синуклеина между остатками 61-66 (EQVTN). Один из подходящих пептидов содержит или последовательность D-аминокислот QYSVLI (ZP-0195, описанную в таблице 3, ниже), или состоит из нее. Другие подходящие пептиды могут содержать последовательность D-аминокислот QYSVLI с одной, двумя или тремя аминокислотными заменами, или состоять из нее. Например, пептид может состоять из последовательности D-аминокислот, выбранной из группы, состоящей из: qykvli, qysvpi, qyspli, qypyli, rysvli, qysvli, qytvli, pysvli или qysvlv.

Пептид может содержать один, два или три дополнительных N-концевых остатка.

Например, пептид может содержать последовательность, выбранную из группы, состоящей из: ekysvli и drysvli, или состоять из нее.

Также в анализе идентифицирован пептид, состоящий из последовательности D-аминокислот, которая взаимодействует с областью  $\alpha$ -синуклеина между остатками 71-76 (VTGVT). Например, пептид может состоять из последовательности D-аминокислот hhviva (ZP-0158) или он может содержать последовательность D-аминокислот hhviva с одной, двумя или тремя аминокислотными заменами, или состоять из нее. Предпочтительно N-концевые остатки гистидина не замещены.

Например, пептид может содержать последовательность, выбранную из группы, состоящей из

hhvvva, hhviva, hhvkva, hhveva, hpviva, hhvip, hhvivv, hhvivt, hhvivy, hhvivw, hhtivv, hhtivk, hhtvva, hhtlva, hhtlvv, hhtevy и hhttvy, или может состоять из нее.

Для области 61-65 сконструирован взаимодействующий пептид Ac-qysvli-NH2 (ZP-0195) для предотвращения агрегации. Также тестировали варианты этой последовательности, где одна или несколько аминокислот в любом данном положении замещена другой, и были проведены изменения на N-конце, такие как добавление дополнительной аминокислоты и ацетилирование (с ZP-0195 по ZP-0230).

Для области 71-75 сконструирован взаимодействующий пептид Ac-hhviva-NH2 (ZP-0158) для предотвращения агрегации. Также тестировали варианты этой последовательности, где одна или несколько аминокислот в любом данном положении замещены другой. Все протестированные пептиды были ацетилированы на N-конце и во всех конструкциях в начале последовательности постоянно оставляли 2 гистидина (с ZP-0158 по ZP-0194).

Все пептиды тестировали в отношении ингибирования агрегации ASYN в TBS поведением анализов агрегации с 50 мкМ ASYN и 100 мкМ ингибитором с 50 мМ tris и 150 мМ NaCl, и 20 мкМ тиофлавином Т. Объем реакции составлял 200 мкл. Каждую реакцию проводили в 96-луночном планшете из полипропилена с контролем в виде только ASYN и только буфера. Реакционные смеси инкубировали при 37°C при встряхивании в течение 48 часов и мониторинг агрегации проводили считыванием флуоресценции тиофлавина Т.

Кинетические кривые строили с использованием программного обеспечения Zyentiafit, которое подбирает для данных сигмовидную функцию  $f(x)=k+A/(1+\exp(-(b(t-t_0)))$ , из которой можно вычислить период задержки, скорость агрегации и общее изменение флуоресценции ThT.

Пептиды ранжировали в соответствии с их эффективностью и в табл. 3, ниже, представлены последовательности, выбранные для последующего исследования. Выбор был основан на пептиде, имеющем более чем 20% повышение времени задержки и/или более чем 20% снижение флуоресценции ThT или скорости агрегации.

Таблица 3. Пептидные последовательности, сконструированные для взаимодействия с областями 61-66 и 71-76 ASYN, которые показывают эффективность в отношении предотвращения агрегации ASYN in-vitro

Код Zyentia	Последовательность	% снижение флуоресценции ThT	% снижение скорости агрегации	% повышение фазы задержки
ZP-0158	Ac-hhviva-NH2	15.1	18.4	32.6
ZP-0159	Ac-hhvvva-NH2	19.8	18.6	36.6
ZP-0160	Ac-hhv1va-NH2	6.6	10.5	36.5
ZP-0161	Ac-hhvkvva-NH2	-5.9	2.3	19.7
ZP-0162	Ac-hhveva-NH2	6.3	6.6	33.4
ZP-0164	Ac-hpviva-NH2	3.7	-2.3	43.0
ZP-0168	Ac-hhvivp-NH2	-8.6	-6.4	29.8
ZP-0169	Ac-hhvivv-NH2	-0.5	-13.3	26.1
ZP-0170	Ac-hhvivt-NH2	24.5	29.3	29.4
ZP-0171	Ac-hhvivy-NH2	13.9	13.2	43.6
ZP-0172	Ac-hhvivw-NH2	-1.7	6.9	28.1
ZP-0175	Ac-hhtivv-NH2	21.9	16.8	89.7
ZP-0179	Ac-hhtivk-NH2	0.6	-0.6	30.4
ZP-0180	Ac-hhtvva-NH2	19.8	18.4	30.3
ZP-0181	Ac-hhtlva-NH2	22.8	-1.2	29.4
ZP-0186	Ac-hhtlvv-NH2	30.7	29.7	16.1
ZP-0193	Ac-hhtevy-NH2	8.7	-15.8	31.4
ZP-0194	Ac-hhttvy-NH2	11.6	7.7	38.8

ZP-0202	Ac-qykvli-NH2	53.5	54.5	-28.8
ZP-0204	Ac-qysvpi-NH2	1.4	-5.6	23.6
ZP-0205	Ac-qyspli-NH2	16.6	-11.9	30.6
ZP-0206	Ac-qypvli-NH2	3.3	8.4	25.9
ZP-0207	Ac-qpsvli-NH2	5.5	3.6	29.7
ZP-0212	Ac-rysvli-NH2	51.1	44.9	-14.4
ZP-0213	Ac-ekysvli-NH2	25.6	12.9	29.4
ZP-0214	Ac-drysvli-NH2	20.4	4.8	-12.6
ZP-0215	NH3-qysvli-NH2	27.6	17.5	3.1
ZP-0221	NH3-qytvli-NH2	24.7	22.8	-22.4
ZP-0222	NH3-qykvli-NH2	42.3	34.3	7.7
ZP-0228	NH3-pysvli-NH2	40.8	46.2	38.2
ZP-0229	NH3-qysvlv-NH2	17.8	24.3	33.5

### Конструирование соединений на основе результатов

Совпадения, идентифицированные способами или программными средствами, описанными выше, можно использовать в качестве матриц для конструирования ингибиторов или стабилизаторов агрегации. Для тестирования библиотек соединений на их аффинность к последовательности-мишени на основе этих матриц можно использовать молекулярную динамику или другие пригодные компьютерные способы.

В компьютерных способах в целях максимизации взаимодействия между ингибитором и агрегирующим полипептидом-мишенью для множества ингибиторов на основе матрицы можно использовать конкретные процедуры минимизации энергии и силовых полей (см. Das B, Meirovitch H, Navon IM, Performance of hybrid methods for large-scale unconstrained optimization as applied to models of proteins, J Comput. Chem. 2003;24:1222-31 и de Bakker PI, DePristo MA, Burke DF, Blundell TL, Ab initio construction of polypeptide fragments: Accuracy of loop decoy discrimination by an all-atom statistical potential and the AMBER force field with the Generalized Born solvation model, Proteins, 2003; 51 (1): 21-40).

Отобранные библиотеки соединений можно тестировать в анализах агрегации *in vitro* в целях идентификации основных соединений, которые могут ингибировать агрегацию полипептидов-мишней.

Когда необходимо, чтобы стабилизатор/ингибитор агрегации был устойчивым к другим протеазам, можно использовать ретро-энантио-производные (с реверсией последовательности C-N и D-аминокислот).

Идентифицированный стабилизатор/ингибитор агрегации можно подвергать слиянию с другими белками/пептидами. Белок/пептид, слитый со стабилизатором/ингибитором агрегации, может действовать в качестве носителей для направления доставки в конкретные области организма, конкретные органы, конкретные типы клеток и т.д.; облегчать внутриклеточную доставку; облегчать проникновение через гематоэнцефалический барьер; повышать время полужизни в плазме и/или взаимодействовать с другим белком или рецептором.

### ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ предсказания потенциальных пептидных последовательностей, ингибирующих агрегацию белков, включающий следующие стадии:

а) идентификация пептидной последовательности, образующей по меньшей мере часть области агрегации в белке-мишени;

б) тестирование, образует ли указанная пептидная последовательность часть  $\beta$ -слоя, причем стадия тестирования включает подстадии идентификации в базе данных белковых структур группы гетерологичных белков, содержащей пептидные последовательности, родственные указанной пептидной последовательности; и идентификации в указанной группе тех белков, в которых указанные родственные пептидные последовательности образуют часть  $\beta$ -слоя;

с) если на стадии б) достигнут положительный результат, извлечение из указанной базы данных белковых структур остатков, которые образуют соседние цепи указанного  $\beta$ -слоя;

д) идентификация остатков в цепях, расположенных рядом с указанной пептидной последовательностью, боковые цепи которых взаимодействуют с указанной пептидной последовательностью, т.е. тех

остатков, которые образуют потенциальную пептидную последовательность, ингибирующую агрегацию белка,

где подстадия идентификации включает сравнение указанных родственных пептидных последовательностей с остатками, содержащимися в строках "SHEET" в файле PDB для представляющего интерес белка.

2. Способ по п.1, где указанные родственные пептидные последовательности включают указанную пептидную последовательность и фрагменты указанной пептидной последовательности.

3. Способ по п.1 или 2, где указанные родственные пептидные последовательности включают последовательности, содержащие консервативные замены одной или более аминокислот в указанной пептидной последовательности.

4. Способ по любому из пп.1-3, где подстадия идентификации включает идентификацию тех остатков в указанных родственных пептидных последовательностях, которые образуют друг с другом водородные связи; предпочтительно в целях идентификации тех остатков, которые образуют водородные связи друг с другом, вычисляют евклидово расстояние между каждой парой остатков, которые отделены по меньшей мере тремя остатками, и считают, что водородная связь образуется, если расстояние составляет менее 3,075 ангстрем.

5. Способ по любому из предшествующих пунктов, где остатки, идентифицированные на стадии d), представляют собой те остатки, боковые цепи которых взаимодействуют с указанной пептидной последовательностью посредством водородных связей.

6. Способ по любому из предшествующих пунктов, где на стадии идентификации используют профиль склонности к агрегации.

7. Способ по любому из пп.1-5, где стадию идентификации проводят экспериментально.

8. Способ по любому из предшествующих пунктов, дополнительно включающий стадии тестирования, взаимодействуют ли остатки, идентифицированные на стадии d), с одним или более другими белками; где предпочтительно стадию тестирования:

(i) проводят на множестве белков, гетерологичных указанному белку-мишени, и/или

(ii) проводят с использованием базы данных белковых структур, которые необязательно опосредуют важные клеточные процессы,

и она необязательно включает подстадии идентификации группы белков, содержащихся в указанной базе данных, которые содержат родственные пептидные последовательности, сходные с указанной пептидной последовательностью; и идентификации в указанной группе тех белков, в которых указанные родственные пептидные последовательности взаимодействуют с указанными идентифицированными остатками.

9. Способ по любому из предшествующих пунктов, где часть области агрегации, идентифицированная на стадии a), представляет собой часть спирали, петлю,  $\beta$ -поворот или  $\beta$ -изгиб.

10. Способ по любому из предшествующих пунктов, включающий получение ингибирующего агрегацию белка пептида, содержащего остатки, идентифицированные на стадии d).

11. Способ по любому из предшествующих пунктов, дополнительно включающий стадии:

е) синтеза пептидной библиотеки, причем члены указанной библиотеки содержат остатки, идентифицированные на стадии d); и

f) определения аффинности членов указанной библиотеки в отношении указанного белка-мишени.

12. Способ получения пептида, ингибирующего агрегацию белка, включающий следующие стадии:

а) идентификация пептидной последовательности, образующей по меньшей мере часть области агрегации в белке-мишени;

б) тестирование, образует ли указанная пептидная последовательность часть  $\beta$ -слоя;

с) если на стадии б) достигнут положительный результат, извлечение из указанной базы данных белковых структур остатков, которые образуют соседние цепи указанного  $\beta$ -слоя;

д) идентификация остатков в цепях, расположенных рядом с указанной пептидной последовательностью, боковые цепи которых взаимодействуют с указанной пептидной последовательностью, т.е. тех остатков, которые образуют потенциальную пептидную последовательность, ингибирующую агрегацию белка;

е) синтез пептидной библиотеки, причем члены указанной библиотеки содержат остатки, идентифицированные на стадии d); и

f) определение аффинности членов указанной библиотеки к белку-мишени.

13. Способ по п.11 или 12, включающий идентификацию в указанной библиотеке пептида, который проявляет высокую аффинность к белку-мишени относительно контролей, в качестве пептида, ингибирующего агрегацию белка, где предпочтительно способ дополнительно включает:

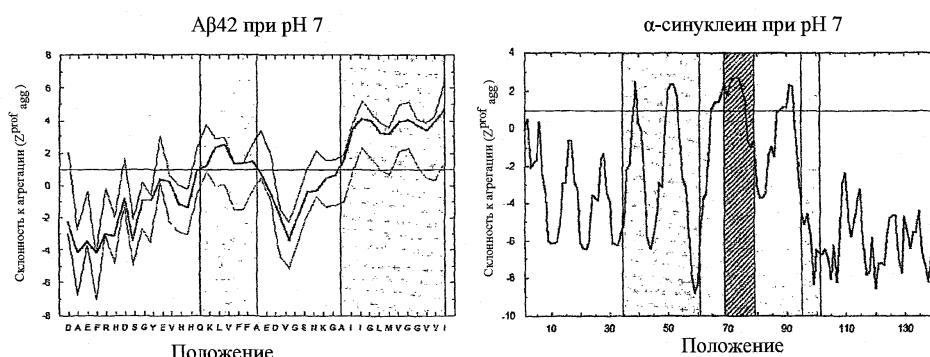
(i) выделение пептида, идентифицированного в указанной библиотеке, или

(ii) синтез пептида, идентифицированного в указанной библиотеке.

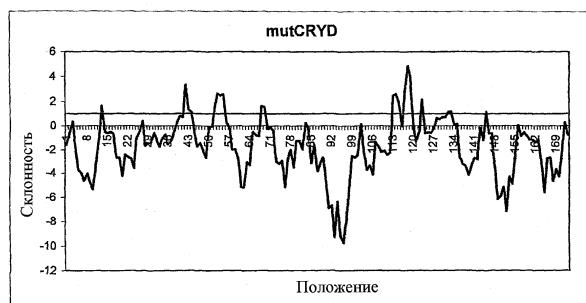
14. Способ по п.10, включающий определение способности указанного пептида к одному или более из следующих:

- i) стабилизация белка в отношении агрегации;
- ii) снижение скорости потери активности белка при хранении;
- iii) снижение опосредуемой агрегацией иммуногенности белка;
- iv) повышение выхода белков в системах трансляции *in vitro*;
- v) повышение стабильности в растворе состава для терапевтического применения;
- vi) ингибирование одного или более клеточных процессов;
- vii) предотвращение олигомеризации или мультимеризации белка.

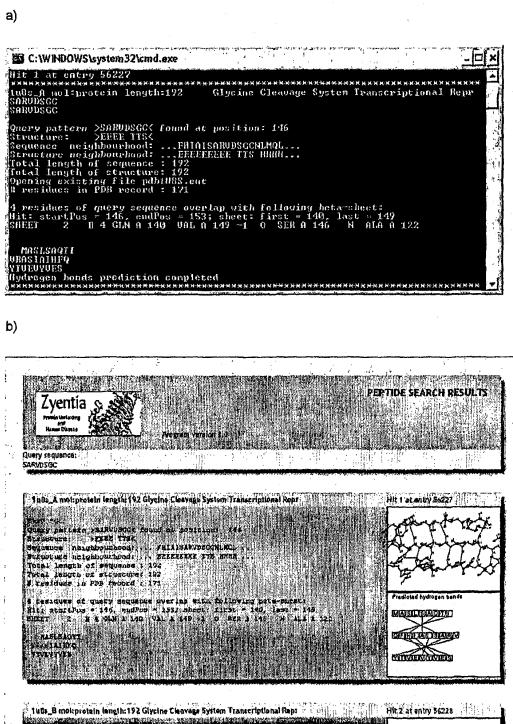
15. Способ конструирования соединения для (A) стабилизации белка в отношении агрегации или (B) повышения стабильности состава для терапевтического применения в растворе, включающий стадии предсказания пептидов, ингибирующих агрегацию белков по любому из пп.1-9 со стадией тестирования, проводимой на множестве белков, гетерологичных указанному белку-мишени, и применения остатка(ов), идентифицированного на стадии d) этого предсказания, для конструирования соответственно (A) соединения для стабилизации белка в отношении агрегации или (B) соединений с повышенной стабильностью в растворе.



Фиг. 1

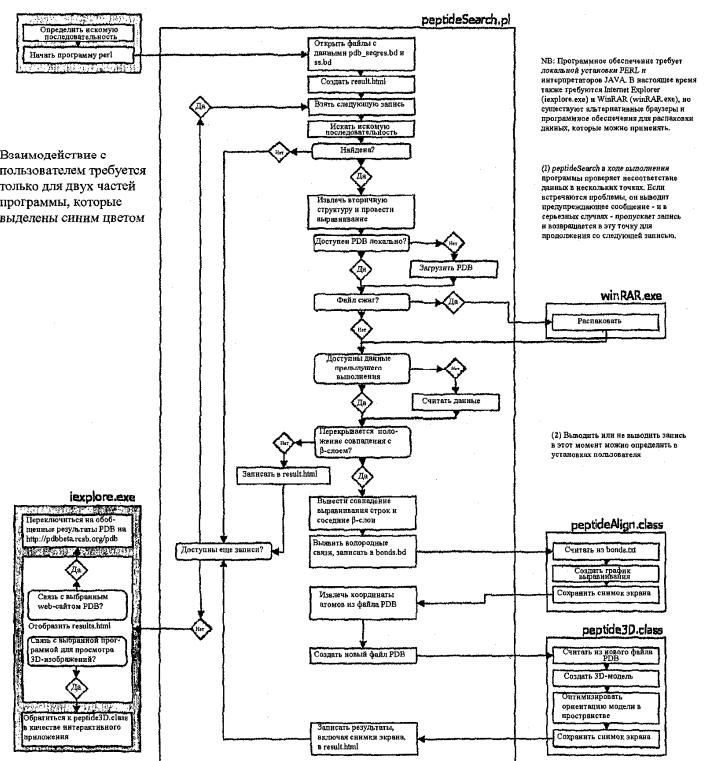


Фиг. 2



ФИГ. 3

## БЛОК-СХЕМА ПРОЦЕССА ПОИСКА ПЕПТИДОВ



ФИГ. 4

## Первые десять строк из pdb\_seqres.txt

```
>1pr2_A mol:protein length:239 Purine Nucleoside Phosphorylase  
  
MATPHINAEKGDFADVVLMPGDPLRAKYIAETPLEDAREVVNNVRGMLGFTGTGYKGRKISVNGHGMGIPSCSIYTK  
  
ELITDGFVKKIIRVGSCGAVLPHVKLDRDVVIGMGACTDSKVNRIFPKDHDPAAIAFDPMVRNAVDAAKALGIDAR  
  
VGNLFSADLFYSPDGFEMDVMEMEKGYGILGVEMEAAGIYGVAEEFGKAALKTICTVSDDHIRTBEQTAAERQTTFNDM  
  
IKIALESVLLGDKE  
  
>1pr2_B mol:protein length:239 Purine Nucleoside Phosphorylase  
  
MATPHINAEKGDFADVVLMPGDPLRAKYIAETPLEDAREVVNNVRGMLGFTGTGYKGRKISVNGHGMGIPSCSIYTK  
  
ELITDGFVKKIIRVGSCGAVLPHVKLDRDVVIGMGACTDSKVNRIFPKDHDPAAIAFDPMVRNAVDAAKALGIDAR  
  
VGNLFSADLFYSPDGFEMDVMEMEKGYGILGVEMEAAGIYGVAEEFGKAALKTICTVSDDHIRTBEQTAAERQTTFNDM  
  
IKIALESVLLGDKE
```

## Первые десять строк из ss. txt

>1PR2:A

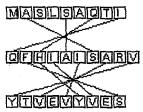
BBTB TTSS SEEEEESSTHHHHHHHHHT BS EE B GGG EEEEEEETTEEEEEE SSHHHHHHHHHH  
HHHHHS BEEEEEEER STTTTT BEEESEBEES SHHHHHHTTS B HHHHHHHHHHHHHHTT EE  
EEEEE S SS S TTHHHHHHHHHTT EEESSHHHHHHHHHHTT EEEEEEETTS HHHHHHHHHHH  
HHHHHHHHHHHHHHH

>1PR2:B

SS TTSS SEESE SSSHHHHHHHHHH BS EE B GGG EEEEEEETTEEEEEE SSHHHHHHHHH  
HHHHHS BEEEEEEER STTTTT BEEESEBEES SHHHHHHTTS B HHHHHHHHHHHHHHTT E  
EEEEE S SS S STTHHHHHHHHHHTT EEESSHHHHHHHHHHTT EEEEEEETTT HHHHHHHHHHH  
HHHHHHHHHHHHHHH

Фиг. 5

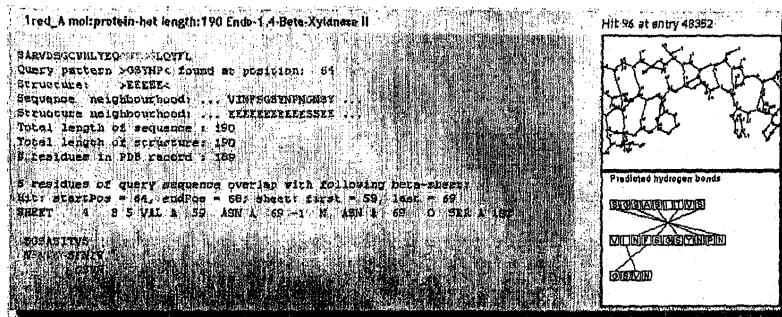
### Предсказанные водородные связи



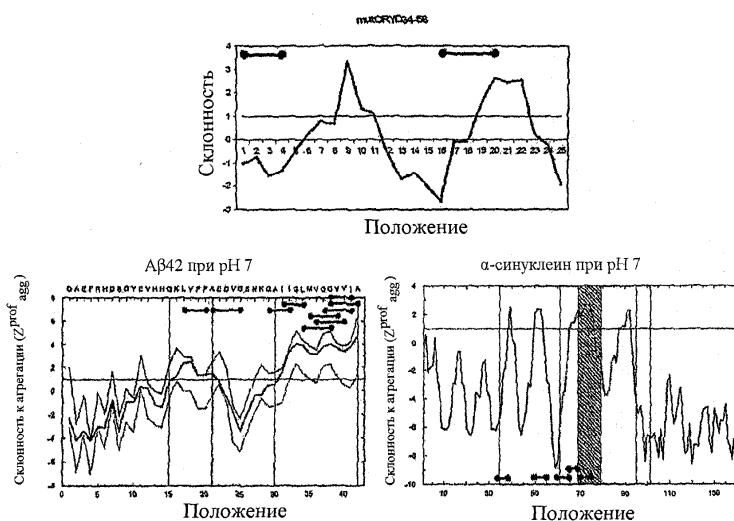
Фиг. 6

OUTPUTTO 2005_08_19_14h1ImOls/images/hit38.png	Predicted hydrogen bonds
NUMATOMS 9	
NUMBONDS 4	
RESID THR 188	
RESID VAL 189	
RESID SER 190	
RESID ILE 60	
RESID ASN 61	
BOND ASN 61SER 188	
RESID PHE 62	
RESID SER 63	
BOND SER 63THR 180	
RESID GLY 148	
BOND GLY 148PHE 61	
RESID SER 149	
BOND SER 149ASN 62	

Фиг. 7



Фиг. 8



Фиг. 9