

Е.А. Бородин, П.Е. Бородин

БИОИНФОРМАТИКА КАК МЕТОД СИСТЕМНОГО АНАЛИЗА В БИОЛОГИИ И МЕДИЦИНЕ. БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ СЕРПИНОВ РАСТЕНИЙ – ОСНОВЫ ЛЕКАРСТВЕННЫХ СРЕДСТВ ДЛЯ КОРРЕКЦИИ НАРУШЕНИЙ ГЕМОСТАЗА И ФИБРИНОЛИЗА

Раскрывается сущность биоинформатики как метода системного анализа в биологии и медицине. Обосновывается возможность использования методов биоинформатики для создания на основе растительных ингибиторов протеаз-серпинов лекарственного препарата, предназначенного для регуляции процессов гемостаза и фибринолиза.

Ключевые слова: биоинформатика, серпины, гемостаз, фибринолиз.

BIOINFORMATICS AS A METHOD OF SYSTEM ANALYSIS IN BIOLOGY AND MEDICINE. BIOINFORMATIC ANALYSIS OF PLANT SERPINS AS A GROUND FOR THE DEVELOPMENT OF THE DRUGS DESIGNED FOR THE REGULATION OF HEMOSTASIS AND FIBRINOLYSIS

Bioinformatics (the essence) is a method of system analysis in biology and medicine. The possibility of exploiting of bioinformatic methods for the development of plant serpins is based on a protease inhibitor drug designed for the regulation of hemostasis and fibrinolysis (which are presented).

Key words: bioinformatics, serpins, hemostasis, fibrinolysis.

В 1970 г. Paulien Hogeweg впервые использовала термин «биоинформатика» применительно к изучению информационных процессов в биотических системах [1]. Современная биоинформатика является составной частью триады новых биологических наук, включающей геномику, протеомику и биоинформатику [2]. Своим возникновением эти науки обязаны международному научному проекту «Геном человека» (1989-2002 гг.), позволившему установить последовательность $3,2 \cdot 10^9$ пар нуклеотидов в молекуле ДНК Homo sapiens. Биоинформатика основана на использовании персональных компьютеров для хранения и обработки информации, полученной в ходе секвенирования геномов живых организмов. Основным предметом биоинформатики являются нуклеиновые кислоты и белки как соединения, хранящие и реализующие генетическую информацию. Типичные задачи биоинформатики включают хранение информации о первичной структуре генов (т.е. нуклеиновых кислот) и белков; глобальное и локальное выравнивание последовательностей генов и белков, а также поиск их гомологов, прогнозирование 3D-структуры и функций белков на основе их первичной структуры; моделирование белок-белковых и белок-лигандных взаимодействий (docking algorithms); высокоточный анализ изображений; вычислительную эволюционную биологию и компьютерный дизайн лекарств.

Нами предпринята попытка использовать биоинформатический анализ растительных ингибиторов протеаз для оценки возможности создания на основе этих белков новых лекарственных средств с целью коррекции нарушений гемостаза и фибринолиза.

Гемостаз (свертывание крови) и фибринолиз (растворение кровяного сгустка) представляют каскады протеолитических ферментативных реакций, поддерживающих надлежащий ток крови в кровеносном русле. Нарушение тонкого баланса между этими процессами может привести к внутрисосудистому свертыванию крови (тромбозу) или к аномальным кровотечениям. В регуляции гемостаза и фибринолиза важнейшая роль принадлежит серпинам – ингибиторам сериновых протеаз, найденных во всех формах жизни (археи, бактерии и эукариоты). Большинство серпинов являются необратимыми ингибиторами сериновых протеиназ семейства химотрипсина, включающих белковые факторы свертывания крови (тромбин, факторы Ха и VIIa и др.) и растворения тромба (альфа-антиплазмин, антитромбин и др.) [3].

С учетом важной роли серпинов в регуляции биологических процессов целесообразно использовать их в качестве лекарственных средств. Тем не менее к настоящему времени только два серпина животного происхождения нашли применение в качестве зарегистрированных лекарственных форм – антитромбин III и апротинин, выпускаемый с различными торговыми названиями (тразилол, гордокс, контрикал и др.). Детальное исследование серпинов представляет императиву для создания совершенных лекарств, способных поддерживать баланс между свертывающей, противосвертывающей и фибринолитической системами крови. Серпины растений могут выступить кандидатами на эту роль [4]. Традиционный подход при оценке способности серпина ингибировать ту или иную протеазу состоит в кинетическом анализе реакции, катализируемой протеазой в присутствии серпина, либо в выявлении серпин-протеиназных комплексов с помощью электрофореза в SDS-геле [5]. Альтернативный подход заключается в использовании биоинформатики, позволяющей сделать аналогичные выводы на основе выравнивания первичных и 3D-структур белков. В настоящем исследовании предпринята попытка сравнительной биоинформатической характеристики растительных серпинов с серпинами, вовлеченными в гемостаз и фибринолиз, с целью выявить в их строении общие и специфические черты.

Мы использовали базы данных UniProt <http://www.uniprot.org> и NCBI Protein <http://www.ncbi.nlm.nih.gov/protein> для нахождения первичных структур, характерных особенностей строения, активных центров, функциональной активности, а также для выполнения множественного и глобального попарного выравнивания первичных структур ингибиторов протеаз. Локальное выравнивание последовательностей серпинов осуществляли на сервере EMBL-EBI по алгоритму EMBOSS Water (PROTEIN) http://www.ebi.ac.uk/Tools/psa/emboss_water. 3-D структуры серпинов находили в RCSB PDB <http://www.rcsb.org/> и осуществляли их выравнивание с использованием опции Analyze Option Sequence and Structure Alignment (Java-утилита RCSB PDB Protein Comparison Tool в оффлайн-режиме).

Суперсемейство серпинов на филогенетической основе разбито на 16 кладов, имеющих определенные структурно-функциональные различия [6]. Из 7 выбранных нами для исследования ингибиторов протеаз человека, участвующих в процессах гемостаза и фибринолиза, 5 относятся к серпинам: антитромбин III (serpinC1), протеин Z-зависимый ингибитор протеаз (serpinA10), ингибитор α_1 -антитрипсина (serpinA1), плазменный ингибитор протеазы C1 (serpinG1), α_2 -антиплазмин (serpinF2). Буквы А – F означают клады, а цифры – индивидуальный номер серпина в кладе. Ингибиторы пути тканевого фактора 1 и 2 относят к кунинам – белкам, гомологичным панкреатическому ингибитору трипсина – апротинину. Ингибиторы протеаз растений классифицируются в соответствии с базой данных MEROPS(8). Для сравнения с названными выше серпинами человека мы выбрали 7 растительных ингибиторов протеаз: серпин-1 резуховидки Таля (*Arabidopsis Thailana*) (serpin1), трипсин – ингибитор А соевых бобов (*Glycine max*) (L3a), трипсин/фактор XIIA – ингибитор кукурузы сахарной (*Zea Mays*) (L6), трипсин – ингибитор 1 горькой тыквы (*Momordica charantia*) (L7) ингибитор 1 цистеиновых протеиназ японского риса (*Oryza sativa subsp. Japonica*) (L25b), субтилизин-химотрипсин – ингибитор WSCI пшеницы мягкой (*Triticum aestivum*) (L13) и ингибитор трипсинатипа Боумана

Бирка соевых бобов (*Glycine max*) (L12). Сведения об идентификационных номерах (ID) в базе UNIPROT, принадлежности к определенным кладам и длине аминокислотных (АМК) цепей выбранных ингибиторов протеаз – в табл. 1.

Таблица 1

Ингибиторы гемостаза и фибринолиза человека и растительные ингибиторы протеаз, выбранные для сравнительного исследования

Ингибитор	UNIPROT ID	Клад*/семейство**	Длина цепи (кол-во АМК)
Ингибиторы протеаз гемостаза и фибринолиза человека			
Антитромбин III	<u>P01008</u>	serpinC1	464
Протеин Z-зависимый ингибитор протеаз	<u>Q9UK55</u>	serpinA10	444
α_1 -антитрипсин	<u>P01009</u>	serpinA1	418
Плазменный ингибитор протеазы C1	<u>P05155</u>	serpinG1	500
α_2 -антиплазмин	<u>P08697</u>	serpinF2	491
Ингибитор пути тканевого фактора 1	<u>P10646</u>		304
Ингибитор пути тканевого фактора 2	<u>P48307</u>		235
Растительные ингибиторы протеаз			
Серпин-1 резуховидки Таля (<i>Arabidopsis Thailana</i>)	<u>Q9S7T8</u>	<u>I04.087</u>	391
Трипсини нгибитор А соевых бобов (<i>Glycine max</i>)	<u>P01070</u>	<u>I03.001</u>	216
Трипсин/фактор ХПА – ингибитор кукурузы сахарной (<i>Zea Mays</i>)	<u>P01088</u>	<u>I06.001</u>	155
Трипсин ингибитор 1 горькой тыквы (<i>Momordica charantia</i>)	<u>P10294</u>	<u>I07.001</u>	30
Ингибитор 1 цистеиновых протеиназ японского риса (<i>Oryzasativasubsp. Japonica</i>)**	<u>P09229</u>	<u>I25.028</u>	140
Субтилизин-химотрипсин ингибитор WSCI – пшеницы мягкой (<i>Triticum aestivum</i>)	<u>P82977</u>	<u>I13.012</u>	84
Ингибитор трипсина типа Боумана-Бирка соевых бобов (<i>Glycinemax</i>)	<u>P01055</u>	<u>I12.001</u>	110

*– классификация по кладам; ** – MEROPSID.

Обращает на себя внимание различная длина АМК-цепей серпинов человека и растительных ингибиторов протеаз. Первые обычно содержат от 400 до 500 остатков АМК, а – вторые 100-200 [7]. Среди выбранных нами серпинов человека длина цепи варьировала в пределах 418-500, а у растительных ингибиторов протеаз – от 30 до 391 (табл. 1). У ингибиторов тканевого пути свертывания 1 и 2, относящихся к семейству кунинов, длина цепи была заметно меньше – 304 и 235 остатков АМК.

Наибольшее сходство последовательностей с серпинами гемостаза выявляет серпин-1 резуховидки Таля Q9S7T8. Степень его идентичности по отношению к отдельным представителям последних варьирует в пределах 9-24 (табл. 2). Несколько меньшую идентичность демонстрируют трипсин ингибитор А соевых бобов P01070 и трипсин/фактор ХПА – ингибитор кукурузы сахарной P01088. Гомология последовательностей остальных растительных ингибиторов протеаз с серпинами гемостаза в целом существенно меньше, но с отдельными представителями достаточно высока. Например, идентичность первичных структур ингибитора трипсина типа Боумана-Бирка соевых бобов P01055 ингибитором тканевого пути 2 P48307 составляет 9,79. Если сопоставить результаты глобального попарного выравнивания последовательностей животных и растительных ингибиторов протеаз (табл. 2) с длиной их АМК-цепей (табл. 1), то нетрудно заключить, что степень идентичности тем выше, чем ближе длина сопоставляемых последовательностей. Именно этим обстоятельством можно объяснить наибольшую идентичность серпинам гемостаза серпина-1 резуховидки Таля Q9S7T8 (391 АМК) и наименьшую – трипсин ингибитора 1 горькой тыквы P10294 (30АМК).

Таблица 2

Результаты попарного выравнивания АМК последовательностей животных и растительных ингибиторов протеаз (идентичность в %)

Ингибитор протеаз	Антитромбин III <u>P01008</u>	α_1 -антитрипсин <u>P01009</u>	Плазменный ингибитор протеазы С1 <u>P05155</u>	α_2 -антиплазмин <u>P08697</u>	Протеин Z-зависимый ингибитор протеаз <u>Q9UK55</u>	Ингибитор тканевого пути 1 <u>P10646</u>	Ингибитор тканевого пути 2 <u>P48307</u>
Серпин-1 резуховидки Таля <u>Q9S7T8</u>	23,78	22,37	16,04	17,59	20,69	13,01	9,11
Трипсин-ингибитор А соевых бобов <u>P01070</u>	7,83	9,37	9,52	6,96	3,14	10,72	11,82
Трипсин/фактор ХПА-ингибитор кукурузы сахарной <u>P01088</u>	7,07	9,15	7,18	7,07	7,83	10,09	13,00
Трипсин-ингибитор 1 горькой тыквы <u>P10294</u>	1,72	0,69	1,2	1,63	1,80	3,94	3,74
Ингибитор 1 цистеиновых протеиназ японского риса <u>P09229</u>	5,77	6,68	4,64	7,07	8,24	9,90	8,06
Субтилизин-химотрипсин – ингибитор WSCI – пшеницы мягкой <u>P82977</u>	3,88	3,83	3,6	3,67	3,15	5,73	5,74
Ингибитор трипсина типа Боумана-Бирка соевых бобов <u>P01055</u>	3,98	4,84	5,2	5,38	3,15	5,14	9,79

Результаты глобального выравнивания не в полной мере отражают сходства в функциональной активности сравниваемых белков, поскольку последняя может определяться не всей АМК цепью, а отдельными доменами. Так, идентичность последовательностей ингибитора трипсина типа Боумана-Бирка соевых бобов и антитромбина III составляет всего 3,97 (рис. 1), что не удивительно, если учесть различную длину полипептидных цепей этих белков – 110 и 464 АМК соответственно. Если же осуществить локальное выравнивание участков цепей ингибитора трипсина типа Боумана-Бирка с 5 по 45 АМК и антитромбина III с 15 по 69 АМК, то идентичность составит 20 (рис. 2).

Аналогичный вывод можно сделать на основании результатов структурного выравнивания этих ингибиторов (рис. 3), свидетельствующих, что при практическом отсутствии сходства 3D-структур белков в целом (идентичность 0, схожесть 7,84) отдельные домены этих белков весьма схожи.

Сравнительный биоинформатический анализ растительных серпинов с серпинами, вовлеченными в гемостаз и фибринолиз, свидетельствует, что, несмотря на различия в их первичных и третичных структурах, между ними имеются определенные черты сходства. Для 3D-структуры серпинов характерно наличие консервативного домена из 9 α -спиралей (А-I) и трех β -структур (А-С). Главный участок β -структуры содержит петлю в активном центре белка (reactive center loop), образованную 20

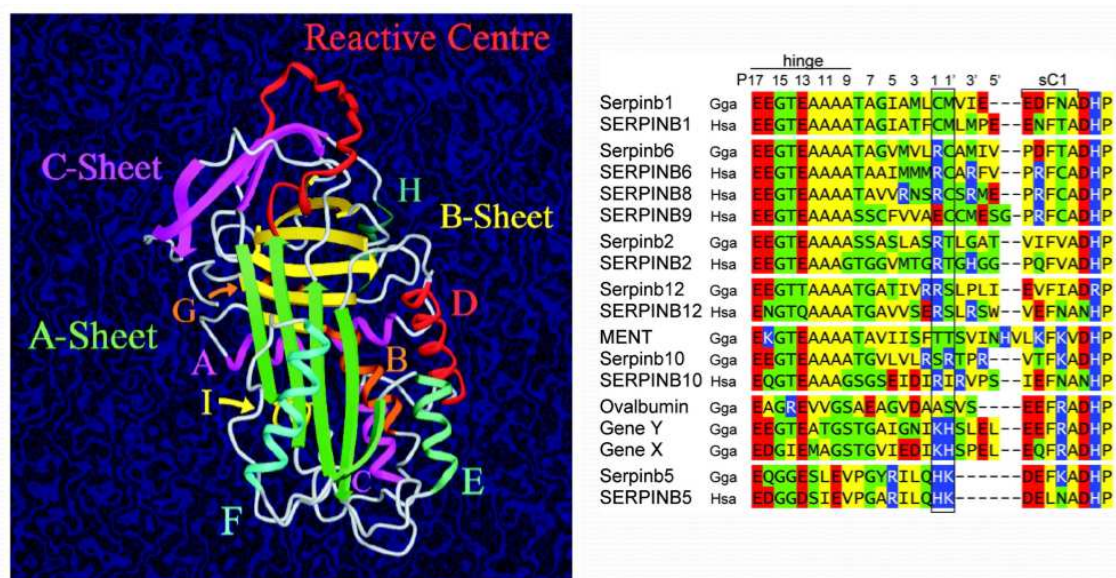


Рис. 4. Элементы 3D-структуры серпинов (слева) и АМК последовательность консервативного домена у С-конца цепи (справа).

1. Hesper, B., Hogeweg, P. Bioinformatica: eenwerkconcept // Kameleon. – 1970. – № 1 (6). – P. 28-29.
2. Арчаков, А.И. Биоинформатика, геномика и протеомика -науки о жизни XXI столетия // Вопросы медицинской биохимии. – 2000. – Т. 47, № 1. – С. 2-9.
3. Rau, J.C., Beaulieu, L.M., Huntington, J.A., Church, F.C. Serpins in thrombosis, hemostasis and fibrinolysis // J. Thromb Haemost. – 2007.– № 5 (Suppl. 1). – P. 102-15.
4. Borodin, Eugene A., Pamirsky, Igor E., Shtarberg, Mikhail A., Dorovskikh, Vladimir A., Korotkikh, Alexander V. Chie Tarumizu, Kiyoharu Takamatsu and Shigeru Yamamoto Effects of Soy Bean Trypsin Inhibitor on Hemostasis. // Soybean - A Review (ed. by Hany A. El-Shemy). – In-Tech, Croatia. – 2013. – P. 495-512.
5. Patston, P.A., Church, F.C. & Olson, S.T. Serpin-ligand interactions// Methods. – 2004. – № 32.– P. 93-109.
6. Irving, J.A., Pike, R.N., Lesk, A.M., Whisstock, J.C. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function // Genome Res. – 2000. – № 10. – P. 1845-1864.
7. Law, R.H.P., Zhang Q.W., McGowan, S., Buckle, A.M., Silverman, G.A., Wong, W., Rosado, C.J., Langendorf, C.G., Pike, R.N., Bird, P.I., Whisstock, J.C. An overview of the serpin superfamily // Genome Biol. – 2006. – № 7. – P. 216 (1-11).