



УДК 541.69:577.122; 615.017:616.079

КОМПЬЮТЕРНОЕ ПРОГНОЗИРОВАНИЕ СПЕКТРА БИОЛОГИЧЕСКОЙ АКТИВНОСТИ НИЗКОМОЛЕКУЛЯРНЫХ ПЕПТИДОВ И ПЕПТИДОМИМЕТИКОВ

© 2000 г. Н. Б. Мартынова[#], Д. А. Филимонов, В. В. Поройков

НИИ биомедицинской химии РАМН, 119832, г. Москва, Погодинская ул., 10

Поступила в редакцию 10.06.99 г. Принята к печати 17.11.99 г.

Разнообразие известных видов биологического действия пептидов и их высокая активность обуславливают актуальность поиска новых базовых структур лекарств в данном классе соединений. Отбор наиболее перспективных веществ может быть осуществлен на основе компьютерного прогноза спектра биологической активности с использованием компьютерной системы PASS (Prediction of Activity Spectra for Substances). Поскольку система PASS изначально конструировалась для прогноза активности низкомолекулярных органических “drug-like” соединений, нами проведено тестирование ее прогностических возможностей на примере выборки из 134 пептидов и пептидомиметиков с 9 известными биологическими активностями. Показано, что средняя точность предсказания составляет около 97%. Это позволяет рекомендовать компьютерный прогноз как для оценки эффектов и механизмов действия эндогенных и синтетических пептидов, так и для отбора среди них наиболее перспективных базовых структур новых лекарственных соединений.

Ключевые слова: пептиды, пептидомиметики; биологическая активность; компьютерное прогнозирование, поиск базовых структур новых лекарственных соединений.

ВВЕДЕНИЕ

Пептиды принимают участие в регуляции почти всех известных биологических процессов, протекающих в организме человека, и не имеют конкурентов по эффективности действия и разнообразию точек приложения [1, 2]. В связи с этим перспективно создание новых лекарственных препаратов на их основе.

Одним из путей исследований в этой области является химическая модификация известных физиологически активных пептидов. Цель модификаций – получение более активных аналогов с повышенной избирательностью и, соответственно, с меньшими, чем у исходного пептида, побочными эффектами. Это обстоятельство дает возможность использовать такие модифицированные аналоги в практической медицине. Примерами такого рода препаратов являются стимулятор памяти – препарат “Семакс” (фрагмент АКТГ-(4-7) [3], не влияющий, в отличие от АКТГ, на обмен глюкокортикоидов, но обладающий его эффектами на ЦНС); противоязвенный препарат “Даларгин” (*D-Ala-Leu*-энкефалин) [4], иммуномодулятор природного происхождения “Ликопид” [5] и другие.

С другой стороны, пептидная химия предоставляет широкие возможности комбинаторного синтеза [6], что позволяет осуществлять поиск базовых структур новых лекарств методами высокопроизводительного (high-throughput) скрининга [6]. Анализ данных по тестированию комбинаторных библиотек дает возможность строить модели фармакофоров, ответственных за проявление активности, и проводить целенаправленную химическую модификацию найденных базовых структур. При этом критериями оптимизации структуры являются повышение у получаемого соединения основной активности, снижение побочного действия и токсичности, увеличение биодоступности и устойчивости к протеолизу. Так, например, Овенс с сотр. [7] для выявления ингибиторов протеиназы ВИЧ-1 синтезировали и исследовали библиотеку тетрапептидов на основе “строительных блоков”, включающих 22 различные аминокислоты. Наиболее активные ингибиторы были использованы для построения модели фармакофора. В результате последующих экспериментов был получен высокоактивный тетрапептид статин, величина IC_{50} которого составляет 200 нМ.

Последовательное применение методов молекулярного моделирования и комбинаторной химии позволяет перейти от исходных базовых структур, построенных на основе пептидов, к пеп-

[#] Автор для переписки (тел.: (095) 245-27-53; e-mail: vvp@ibmh.msk.su).

тидомиметикам, лишенным известных недостатков (недостаточная стабильность, неэффективность при пероральном применении и др.).

Несмотря на огромные возможности комбинаторного синтеза и высокопроизводительного скрининга, позволяющие изучать сотни тысяч веществ еженедельно, естественным ограничителем данного подхода является высокая суммарная стоимость исследований. Отбор наиболее перспективных структур на ранних стадиях исследования возможен путем применения виртуального компьютерного скрининга, позволяющего предсказать биологическую активность веществ еще до проведения синтеза на основе анализа сконструированных структур. При этом традиционно применяются методы QSAR [8] и молекулярного моделирования [8], ориентированные на изучение взаимодействия лигандов с одной или небольшим количеством макромолекулярных мишеней.

Нами недавно предложен подход к прогнозированию спектра биологической активности вещества по структурной формуле, реализованный в компьютерной системе PASS (Prediction of Activity Spectra for Substances) [9–11], способной предсказать наряду с основной активностью также побочные эффекты вещества. Таким образом, как благодаря дополнительной информации о полезных фармакотерапевтических свойствах веществ, так и за счет отбраковки соединений с нежелательными видами активности, возможности для отбора перспективных соединений расширяются. Краткое описание компьютерной системы PASS приведено в разделе “Эксперимент. часть”. Имеется возможность получения прогноза биологической активности на основе структурной формулы вещества, вводимой в компьютер с помощью химического редактора ISIS/Draw и представленной в виде MOL-файла (<http://www.mdli.com>) через Интернет (<http://www.ibmh.msk.su/PASS>).

Поскольку тестирование компьютерной системы PASS осуществлялось ранее на примерах низкомолекулярных органических, так называемых “drug-like”, соединений, возможности ее применения для прогноза биологической активности пептидов и пептидомиметиков нуждаются в специальном изучении. Цель настоящей работы – оценка точности предсказания системой PASS биологической активности для ди-, три- и тетрапептидов.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Мы выполнили прогноз спектра биологической активности для 134 описанных в литературе веществ (тестовой выборки), включающей пептиды и пептидомиметики с известной биологиче-

ской активностью. Выборка достаточно разнообразна по проявляемым веществами биологической активности: ингибиторы нитросинтазы (6 соединений), ингибиторы фарнизилтрансферазы (10), ингибиторы ангиотензинпревращающего фермента и нейтральной эндопептидазы (17), агонисты холецистокинина (20), антагонисты эндотелиновых рецепторов (16), противогрибковые соединения – ингибиторы миристоилтрансферазы (31), ингибиторы эластазы (34).

Для всех соединений тестовой выборки с помощью системы PASS был выполнен прогноз спектра биологической активности в режиме исключения из обучающего набора возможных структурных аналогов соединений тестовой выборки. Результаты прогноза в сравнении с известными экспериментальными данными приведены в табл. 1. Здесь П/Э – сопоставление данных прогноза и экспериментальных данных для противогрибковой активности; ++ означает наличие активности как в прогнозе, так и в эксперименте; -/+ – отсутствие активности в прогнозе и наличие в эксперименте (ошибка 1-го рода). Поскольку в публикациях, как правило, представлены только данные об активных соединениях, оценить величины точности предсказания отсутствия активности (-/-) и ошибок 2-го рода (+/-) не представляется возможным. P_a – вероятность проявления прогнозируемой активности; P_i – вероятность отсутствия данной активности. Некоторую оценку ошибки 2-го рода можно получить при скользящем контроле, так как параметры системы установлены таким образом, чтобы при обучении величина ошибки 1-го рода была равна ошибке 2-го рода (<http://www.ibmh.msk.su/PASS>).

Суммарные оценки качества прогноза для веществ тестовой выборки приведены в табл. 2. Как видно из приведенных данных, большая часть активностей предсказывается с высокой точностью (более 90%). Средняя точность прогноза составляет 97.25%.

Лишь для одной группы веществ, обладающих противогрибковой активностью, точность прогноза несколько ниже (83.87%). По-видимому, это объясняется тем, что их структура обладает большей новизной по отношению к соединениям обучающей выборки, имеющим противогрибковую активность. Для повышения точности прогноза противогрибковой активности эти вещества должны быть добавлены в обучающую выборку системы PASS (<http://www.ibmh.msk.su/PASS>).

На рис. 1 приведено распределение числа соединений по величине вероятности $P_a/(P_a + P_i)$, с которой прогнозируется та или иная известная из литературы активность. Из рисунка видно, что для большинства (80%) соединений тестовой

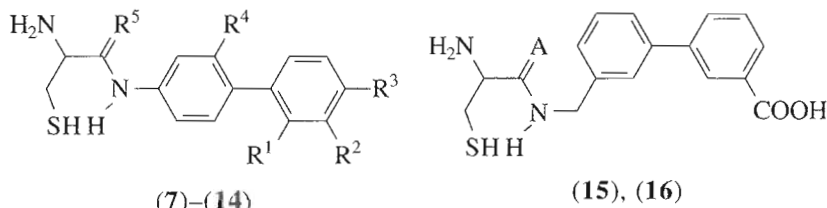
Таблица 1. Структурные формулы веществ тестовой выборки и сопоставление данных прогноза с экспериментальными данными

а) Ингибиторы нитросинтазы [12]

Номер	Структура*	П/Э	P_a	P_i
1	Arg(NO ₂)-Phe	+/+	0.818	0.002
2	Phe-Arg(NO ₂)	+/+	0.819	0.002
3	Phe-Arg(NO ₂)-OCH ₃	+/+	0.788	0.002
4	Arg(NO ₂)-Phe-OCH ₂ C ₆ H ₅	+/+	0.773	0.002
5	Arg(NO ₂)-Phe-OCH ₃	+/+	0.788	0.002
6	Phe-Arg(NO ₂)-OCH ₂ C ₆ H ₅	+/+	0.773	0.002

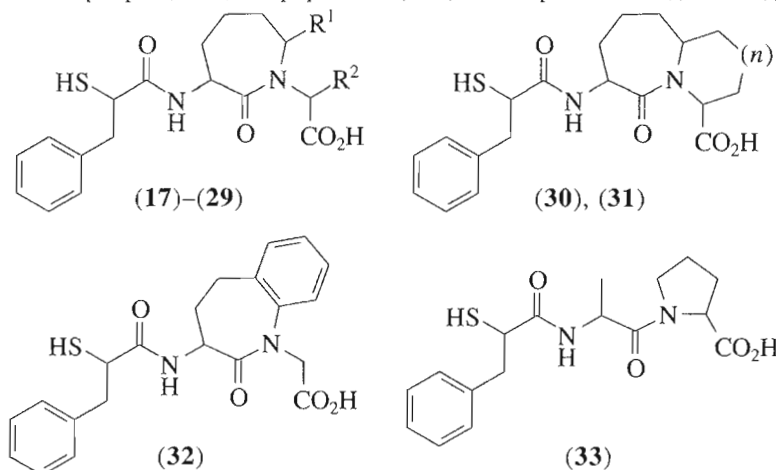
* Arg(NO₂) – нитроаргинин.

б) Ингибиторы фарнизилтрансферазы [13]



Номер	R ¹	R ²	R ³	R ⁴	R ⁵	A	П/Э	P_a	P_i
7	COOH	H	H	H	H,H	-	+/+	0.847	0.002
8	H	COOH	H	H	H,H	-	+/+	0.813	0.002
9	H	H	COOH	H	H,H	-	+/+	0.861	0.002
10	H	COOH	H	H	O	-	+/+	0.646	0.003
11	H	H	H	H	H,H	-	+/+	0.969	0.001
12	H	CH ₃	H	H	H,H	-	+/+	0.890	0.002
13	H	COOCH ₃	H	H	H,H	-	+/+	0.811	0.002
14	H	COOH	H	OCH ₃	H,H	-	+/+	0.695	0.003
15	-	-	-	-	-	O	+/+	0.648	0.003
16	-	-	-	-	-	H,H	+/+	0.768	0.002

в) Ингибиторы ангиотензинпревращающего фермента (АФ) и нейтральной эндопептидазы (НЭ) [14]

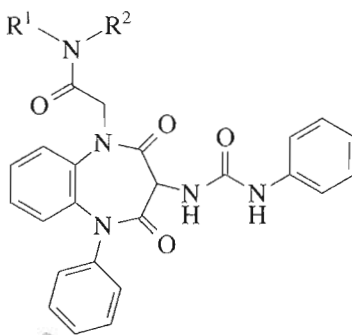


Номер	R ¹	R ²	П/Э	АФ		НЭ	
				P_a	P_i	P_a	P_i
17	H	H	+/+	0.868	0.002	0.854	0.002
18	H	CH ₃	+/+	0.869	0.002	0.889	0.002
19	H	C ₆ H ₅ CH ₂	+/+	0.860	0.003	0.856	0.002
20	H	(CH ₃) ₂ CH	+/+	0.678	0.004	0.622	0.003
21	CH ₃	H	+/+	0.853	0.003	0.818	0.002
22	C ₃ H ₇	H	+/+	0.812	0.003	0.784	0.002
23	CH ₂ =CHCH ₂	H	+/+	0.783	0.003	0.765	0.002
24	HOCH ₂ CH ₂	H	+/+	0.819	0.003	0.789	0.002
25	CH ₃ c(C ₃ H ₄)	H	+/+	0.761	0.003	0.722	0.002

Таблица 1. (Продолжение)

Номер	R ¹	R ²	П/Э	АФ		НЭ	
				P _a	P _i	P _a	P _i
26	(CH ₃) ₂ CHCH ₂	H	+/+	0.815	0.003	0.789	0.002
27	c(C ₅ H ₉)	H	+/+	0.863	0.002	0.799	0.002
28	C ₃ H ₇	CH ₃	+/+	0.712	0.003	0.659	0.003
29	C ₃ H ₇	C ₂ H ₅	+/+	0.669	0.004	0.621	0.003
30 (n = 0)	–	–	+/+	0.961	0.002	0.940	0.001
31 (n = 1)	–	–	+/+	0.961	0.002	0.940	0.001
32	–	–	+/+	0.849	0.003	0.872	0.002
33	–	–	+/+	0.905	0.002	0.662	0.003

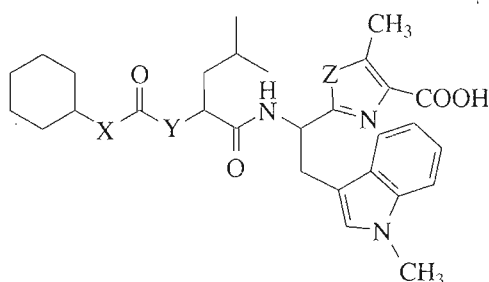
г) Агонисты холецистокинина [15]



Номер	R ¹	R ²	П/Э	P _a	P _i
34	H	C ₆ H ₅	+/+	0.676	0.003
35	CH ₃	То же	+/+	0.894	0.002
36	CH ₃ CH ₂	»	+/+	0.893	0.002
37	CH ₃ CH ₂ CH ₂	»	+/+	0.869	0.003
38	CH ₃ CH ₂ CH ₂ CH ₂	»	+/+	0.873	0.003
39	NCCH ₂ CH ₂	»	+/+	0.807	0.003
40	HOOCCH ₂	»	+/+	0.837	0.003
41	C ₂ H ₅ OOCCH ₂	»	+/+	0.694	0.003
42	H ₂ NCH ₂ CH ₂	»	+/+	0.877	0.002
43	(CH ₃) ₂ CH	»	+/+	0.954	0.003
44	c(C ₆ H ₁₁)	»	+/+	0.473	0.004
45	C ₆ H ₅	»	+/+	0.592	0.004
46	(CH ₃) ₂ CH	4-НО-C ₆ H ₄	+/+	0.936	0.002
47	То же	4-CH ₃ O-C ₆ H ₄	+/+	0.935	0.002
48	»	4-морфолино-C ₆ H ₄	+/+	0.926	0.002
49	»	4-НОOC-C ₆ H ₄	+/+	0.927	0.002
50	»	4-F-C ₆ H ₄	+/+	0.938	0.002
51	»	C ₆ H ₅ CH ₂	+/+	0.454	0.004
52	»	(CH ₃) ₂ CH	+/+	0.307	0.006
53	CH ₃ CH ₂	CH ₃ CH ₂	+/+	0.299	0.006

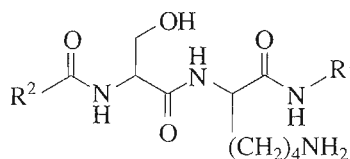
Таблица 1. (Продолжение)

д) Антагонисты эндотелиновых рецепторов [16]



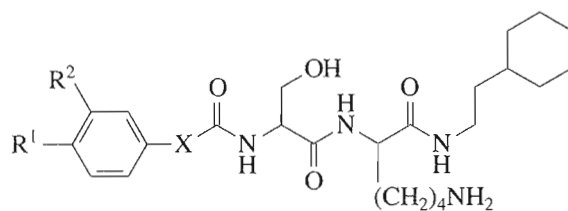
Номер	X	Y	Z	П/Э	P_a	P_i
54	NH	NH	NH	+/+	0.191	0.017
55	NCH ₃	NH	NH	+/+	0.177	0.024
56	O	NH	NH	+/+	0.175	0.024
57	CH ₂	NH	NH	+/+	0.181	0.021
58	NH	NCH ₃	NH	+/+	0.178	0.023
59	NH	O	NH	+/+	0.195	0.016
60	NH	CH ₂	NH	+/+	0.171	0.027
61	NH	NH	O	+/+	0.380	0.007
62	NCH ₃	NH	O	+/+	0.302	0.007
63	O	NH	O	+/+	0.329	0.007
64	CH ₂	NH	O	+/+	0.354	0.007
65	NH	NCH ₃	O	+/+	0.327	0.007
66	NH	O	O	+/+	0.393	0.007
67	NH	CH ₂	O	+/+	0.322	0.007
68	NCH ₃	O	O	+/+	0.442	0.006
69	O	O	O	+/+	0.315	0.007

е) Противогрибковые соединения (ПГ) [17] ингибиторы миристоилтрансферазы (ИМ) [17]

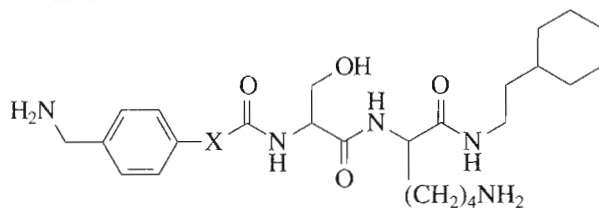


Номер	R ¹	R ²	ПГ			ИМ	
			П/Э	P_a	P_i	P_a	P_i
70	(CH ₃) ₂ CHCH ₂ CH ₂	NH ₂ (CH ₂) ₁₀	+/+	0.270	0.132	0.369	0.002
71	(CH ₃) ₂ CHCH ₂	То же	+/+	0.244	0.159	0.389	0.002
72	(CH ₃) ₃ CCH ₂ CH ₂	»	+/+	0.221	0.182	0.345	0.002
73	C ₆ H ₅ CH ₂ CH ₂	»	+/+	0.213	0.193	0.461	0.002
74	c(C ₆ H ₁₁)CH ₂ CH ₂	»	+/+	0.275	0.127	0.655	0.001
75	c(C ₆ H ₁₁)CH ₂ CH ₂ CH ₂	»	+/+	0.257	0.145	0.455	0.002
76	c(C ₈ H ₁₅)CH ₂ CH ₂ CH ₂	»	+/+	0.257	0.145	0.455	0.002
77	NH ₂ OC(CH ₂ CH(CH ₃) ₂)CH	»	+/+	0.234	0.169	0.335	0.002
78	NH ₂ OCCH ₂	»	-/+	-	-	0.467	0.002
79	NH ₂ OC(CH ₃)CH	»	+/+	0.215	0.189	0.383	0.002
80	NH ₂ OC((CH ₃) ₂ CH)CH	»	+/+	0.233	0.170	0.335	0.002
81	NH ₂ OC(CH(CH ₃)C ₂ H ₅)CH	»	+/+	0.285	0.118	0.302	0.003
82	NH ₂ OC(CH ₂ C ₆ H ₅)CH	»	-/+	-	-	0.329	0.002
83	NH ₂ OC(CH ₂ c(C ₆ H ₁₁))CH	»	+/+	0.232	0.171	0.414	0.002
84	NH ₂ OC(c(C ₆ H ₁₁))CH	»	-/+	-	-	0.305	0.003
85	CH ₃ O ₂ C((CH ₃) ₂ CHCH ₂)CH	»	+/+	0.216	0.189	0.240	0.003
86	То же	NH ₂ (CH ₂) ₉	+/+	0.216	0.189	0.240	0.003
87	»	NH ₂ (CH ₂) ₈	+/+	0.216	0.189	0.197	0.004
88	»	NH ₂ (CH ₂) ₆	+/+	0.216	0.189	0.190	0.004
89	»	NH ₂ (CH ₂) ₁₁	+/+	0.216	0.189	0.240	0.003
90	»	CH ₃ NH(CH ₂) ₁₀	-/+	-	-	0.197	0.004
91	»	(CH ₃) ₂ N(CH ₂) ₁₀	-/+	-	-	0.190	0.004

Таблица 1. (Продолжение)

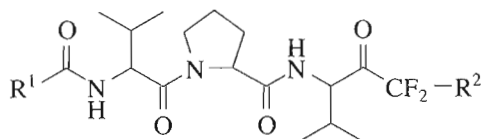


Номер	R ¹	R ²	X	П/Э	ПГ		ИМ	
					P _a	P _i	P _a	P _i
92	NH ₂ (CH ₂) ₅	H	CH ₂	+/+	0.262	0.140	0.622	0.001
93	NH ₂ (CH ₂) ₆	H	CH ₂	+/+	0.230	0.173	0.541	0.002
94	NH ₂ (CH ₂) ₈	H	CH ₂	+/+	0.262	0.140	0.622	0.001
95	H	NH ₂ (CH ₂) ₅	CH ₂	+/+	0.249	0.154	0.589	0.001
96	NH ₂ CH ₂	H	(CH ₂) ₆	+/+	0.280	0.122	0.643	0.001
97	NH ₂ (CH ₂) ₆	H	CH(CH ₃)	+/+	0.225	0.178	0.421	0.002



Номер	X	П/Э	ПГ		ИМ	
			P _a	P _i	P _a	P _i
98	C≡C(CH ₂) ₄	+/+	0.317	0.096	0.410	0.002
99	C≡CC ₆ H ₄ CH ₂	+/+	0.337	0.085	0.654	0.001
100	(CH ₂) ₂ C ₆ H ₄ CH ₂	+/+	0.262	0.139	0.642	0.001

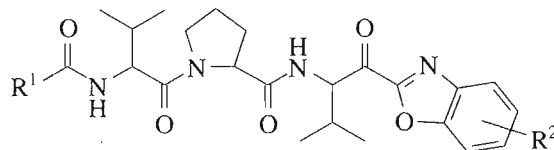
ж) Ингибиторы эластазы [18]



Номер	R ¹	R ²	П/Э	P _a	P _i
101	4-CH ₃ O-C ₆ H ₄	F	+/+	0.681	0.002
102	C ₆ H ₅ CH ₂ O	»	+/+	0.704	0.002
103	C ₆ H ₅ O	»	+/+	0.707	0.002
104	c(C ₅ H ₉)O	»	+/+	0.685	0.002
105	(CH ₃) ₂ CHCH ₂ O	»	+/+	0.710	0.002
106	(CH ₃) ₂ CHO	»	+/+	0.737	0.002
107	CH ₃ CH ₂ O	»	+/+	0.719	0.002
108	CH ₃ O	»	+/+	0.722	0.002
109	C ₆ H ₅ OCH ₂	»	+/+	0.679	0.002
110	4-CH ₃ O-C ₆ H ₄	CH ₃ (CH ₂) ₂ NHCO	+/+	0.599	0.003
111	То же	(2-пиридил)CH ₂ NHCO	+/+	0.501	0.003
112	»	CH ₃ NHCO	+/+	0.603	0.003
113	C ₆ H ₅ CH ₂ O	CH ₃ NHCO	+/+	0.629	0.003

Таблица 1. (Окончание)

Номер	R ¹	R ²	П/Э	P _a	P _i
114	C ₆ H ₅ CH ₂ O	NH ₂ CO	+/+	0.597	0.003
115	C ₆ H ₅ O	CH ₃ NHCO	+/+	0.626	0.003
116	C ₆ H ₅ O	(2-пиридил)CH ₂ NHCO	+/+	0.521	0.003
117	4-CH ₃ O-C ₆ H ₄	C ₄ H ₉ OC(O)NHCH ₂	+/+	0.440	0.004
118	То же	NH ₂ CH ₂	+/+	0.504	0.003
119	»	C ₆ H ₅ CH ₂ NHCH ₂	+/+	0.473	0.003
120	»	C ₆ H ₅ C(O)NHCH ₂	+/+	0.516	0.003
121	C ₆ H ₅ CH ₂ O	C ₄ H ₉ OC(O)NHCH ₂	+/+	0.480	0.003
122	То же	NH ₂ CH ₂	+/+	0.527	0.003
123	»	C ₆ H ₅ CH ₂ NHCH ₂	+/+	0.506	0.003
124	C ₆ H ₅ O	C ₄ H ₉ OC(O)NHCH ₂	+/+	0.460	0.003
125	То же	NH ₂ CH ₂	+/+	0.518	0.003
126	»	C ₆ H ₅ CH ₂ NHCH ₂	+/+	0.480	0.003
127	»	CH ₃ C(O)NHCH ₂	+/+	0.511	0.003
128	»	(3-пиридил)C(O)NHCH ₂	+/+	0.436	0.004



Номер	R ¹	R ²	П/Э	P _a	P _i
129	C ₆ H ₅ CH ₂ O	H	+/+	0.613	0.003
130	C ₆ H ₅ O	3-CO ₂ CH ₃	+/+	0.519	0.003
131	C ₆ H ₅ CH ₂ O	2-CO ₂ CH ₃	+/+	0.537	0.003
132	4-CH ₃ O-C ₆ H ₄	2-CO ₂ CH ₃	+/+	0.524	0.003
133	4-CH ₃ O-C ₆ H ₄	2-NHCOCH ₃	+/+	0.519	0.003
134	C ₆ H ₅ O	2-CO ₂ CH ₃	+/+	0.544	0.003

Таблица 2. Оценка качества прогноза для веществ тестовой выборки

Активность	Количество веществ	Ошибки 1-го рода, %
Ингибиторы фарнизилтрансферазы	10	0
Ингибиторы ангиотензинпревращающего фермента	17	0
Ингибиторы нейтральной эндопептидазы	17	0
Агонисты холецистокинина	20	0
Антагонисты эндотелиновых рецепторов	16	0
Противогрибковые	31	16, 13
Ингибиторы миристоилтрансферазы	31	0
Ингибиторы нитросинтазы	6	0
Ингибиторы эластазы	34	0

выборки известная активность предсказывается с вероятностью, превышающей 90%.

Таким образом, в настоящей работе нами показана высокая точность прогноза биологической активности ди-, три- и тетрапептидов и пеп-

тидомиметиков с использованием системы PASS. Этот результат свидетельствует о перспективности использования данной системы для оценки фармакологических эффектов и биохимических механизмов действия небольших эндогенных и

синтетических пептидов, конструирования и скрининга базовых структур новых лекарств на их основе.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Тестовая выборка. Основные требования, предъявляемые к тестовой выборке таковы: (1) вещества должны принадлежать к классу пептидов или пептидомиметиков; (2) сравнительно небольшой размер молекулы: ди-, три-, тетрапептиды (“drug-like substances” [19]); (3) наличие не менее 5 соединений с одинаковой биологической активностью; (4) разнообразие представленных видов биологической активности.

Как видно из приведенных на рис. 2 данных, молекулярная масса соединений используемой нами тестовой выборки в среднем выше, чем для известных лекарств [19]. Это обусловлено тем, что относительно тяжелые пептиды и пептидомиметики составляют пока лишь небольшую часть используемых в медицинской практике лекарственных препаратов.

База данных по пептидам и пептидомиметикам. Отобранные в тестовую выборку вещества были введены в специальную базу данных, созданную нами на основе широко используемой системы управления химическими базами данных (СУБД) ISIS/Base (MDL Information Systems, Inc., <http://www.mdli.com>). ISIS/Base обеспечивает ввод химических структурных формул и сопутствующей текстовой информации в компьютер; хранение, поиск и обработку данных. Созданная нами база данных содержала следующие поля: (1) индивидуальный номер (ID) – уникальный для каждого вещества, отражающий в закодированной форме данные о литературном источнике, из которого взята информация; (2) структурная формула соединения (Structure); (3) молекулярная масса (MW); (4) известная биологическая активность (Biological activity); (5) ссылка на литературный источник (Reference), содержащий данные об этом соединении; (6) результаты компьютерного прогноза спектра биологической активности соединения (Prediction). ISIS/Base позволяет экспортировать структурные данные в виде SDF- и MOL-файлов, которые являются стандартом “де-факто” для представления химической структурной информации и используются в системе PASS при прогнозировании биологической активности. Нами также была разработана компьютерная программа, генерирующая виртуальные комбинаторные библиотеки пептидов и пептидомиметиков, которые могут храниться в ISIS/Base [20]. Пример представления информации, содержащейся в нашей базе данных, на дисплее компьютера приведен на рис. 3.

Прогнозирование спектра биологической активности. Для прогнозирования спектра биологи-

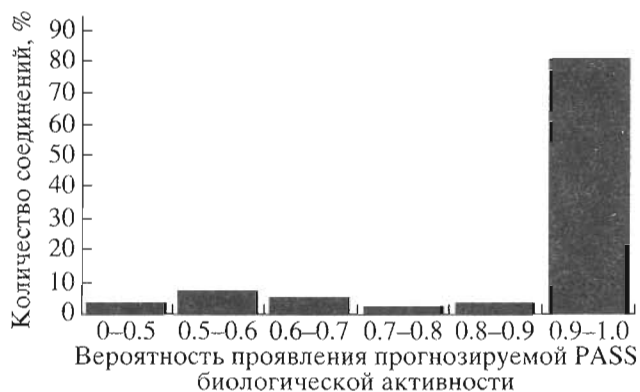


Рис. 1. Диаграмма распределения соединений тестовой выборки по вероятности, с которой прогнозируется известная из литературы активность. По оси абсцисс – прогнозируемая вероятность биологической активности ($P_a/(P_a + P_i)$), известная из литературы (0–0.5 – данная активность не прогнозируется: ошибка 1-го рода); по оси ординат – количество соединений.

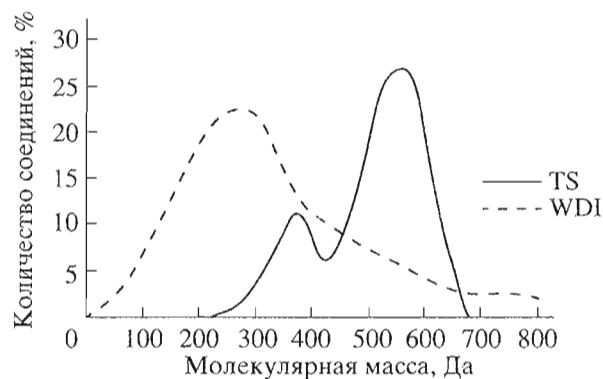


Рис. 2. Диаграмма распределения соединений использованной нами тестовой выборки (TS) и соединений из World Drug Index (WDI) [19] по молекулярной массе (в дальтонах).

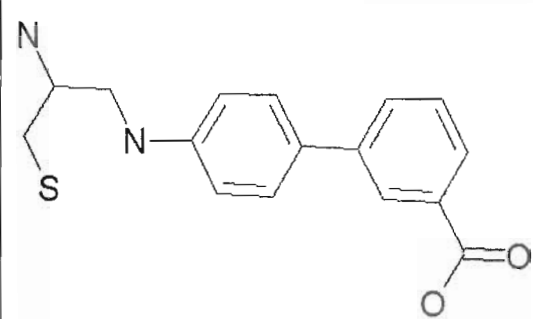
ческой активности соединений тестовой выборки использована компьютерная система предсказания спектра биологической активности веществ по структурной формуле PASS. Функционирование системы PASS основано на процедуре обучения с использованием обучающей выборки химических соединений с известной биологической активностью. Основные элементы системы PASS включают в себя: описание химической структуры, представление биологической активности, обучающую выборку, математический алгоритм установления зависимости структуры–активности. Биологическая активность в системе PASS прогнозируется качественным образом (наличие/отсутствие). Общее число прогнозируемых видов активности (спектр биологической активности) включает свыше 400 фармакологических эффектов, механизмов действия, а также канцерогенность, тератогенность/эмбриотоксичность, мута-

ISIS/Base - [PEPT4.DB/4]

File Edit Options Object Database Search List Window Help

Forms Query Browse Update <Root> 5 of 134
Search Domain: All

STRUCTURE



ID
JMC-1996-39-1-218-4

REFERENCE
Journal of Medicinal Chemistry, 1996,
Vol.39, No.1, P.218

BIOLOGICAL ACTIVITY
Farnesyltransferase inhibitor

Mol.weight 302.40

Prediction of Activity Spectra
Probability of to be active as: 4 Possible activities at $P_a > P_i$.

Activity	P_a	P_i
0.999 Farnesyltransferase inhibitor	0.813	0.002
0.990 Antibacterial	0.151	0.008
0.978 Antibiotic Quinolone-like	0.176	0.051
0.733 Alpha adrenoreceptor antagonist	0.097	0.020
0.714 Enkephalinase inhibitor		
0.708 Beta adrenoreceptor antagonist		
0.664 Adrenergic blocker		
0.607 Antibiotic Naphthyridine-like		

Рис. 3. Пример данных для одного из веществ тестовой выборки, представленный в базе данных ISIS/Base.

генность. Для каждого из прогнозируемых видов активности с помощью специально разработанного математического алгоритма [9] рассчитываются вероятности наличия и отсутствия данной активности (P_a и P_i соответственно). Результаты прогноза представляются в виде списка активностей, упорядоченных по убыванию разности вероятностей их наличия и отсутствия, причем в данном исследовании мы рассматривали только наиболее вероятные активности, для которых $P_a/(P_a + P_i) > 0.5$. Более подробно система PASS описана в литературе [9] и на соответствующих страницах в Интернете (<http://www.ibmh.msk.su/PASS>). Примеры применения системы PASS в практических целях представлены в работах [21–23]. Ограничения метода детально рассмотрены в статье [24].

СПИСОК ЛИТЕРАТУРЫ

1. Ашмарин И.П. // Эволюц. биохимия и физиология. 1982. Т. 18. С. 3–10.
2. Гомазков О.А. Физиологически активные пептиды. М.: ИПГМ, 1995. 144 с.
3. Пономарева-Степная М.А., Незавибатько В.Н., Антонова Л.В., Алфеева Л.Ю., Потаман В.Н., Каменский А.А., Ашмарин И.П. // Хим.-фарм. журн. 1984. Т. 7. С. 790.
4. Виноградов В.В., Сневак С.Е., Ярыгин К.Н., Коробов Н.В., Соловьева А.И., Шехтер А.Б., Титов М.И. // Бюлл. эксперимент. биологии и медицины. 1987. Т. 104. С. 89–91.
5. Иванов В.Т. // Вопр. мед. химии. 1984. Т. 30. С. 23–31.
6. Hobbs de Witt. Combinatorial Libraries and High-Throughput Synthesis in The Practice of Medicinal Chemistry. London: Academic Press, 1996. P. 117–134.
7. Owens R.A., Gesellchen P.D., Houchins B.J., Dimarichi R.D. // Biochem. Biophys. Res. Commun. 1991. V. 181. P. 402–408.
8. Norinder U. // Computer-Aided Drug Design Industrial Research / Ed. E.C. Herrmann, R. Franke. Berlin: Springer Verlag, 1995. P. 99–109.
9. Глориозова Т.А., Филимонов Д.А., Лагунин А.А., Поройков В.В. // Хим.-фарм. журн. 1998. Т. 32. С. 32–39.
10. Поройков В.В., Филимонов Д.А., Степанчикова А.В., Будунова А.П., Шилова Е.В., Рудницких А.В., Селезнева Т.М., Гончаренко Л.В. // Хим.-фарм. журн. 1996. Т. 30. С. 20–23.
11. Филимонов Д.А., Поройков В.В., Карачева Е.И., Казарян Р.К., Будунова А.П., Михайловский Е.М., Рудницких А.В., Гончаренко Л.В., Буров Ю.В. // Эксперимент. и клин. фармакология. 1995. Т. 58. С. 56–62.

12. Silverman R.B., Hui Huang, Marletta M.A., Martasek P. // *J. Med. Chem.* 1997. V. 40. P. 2813–2817.
13. Sebt S.M., Qian Y., Vogt A., Hamilton A.D. // *J. Med. Chem.* 1996. V. 39. P. 217–223.
14. Robl J.A., Cimarusti M.P., Simpkins L.M., Brown B., Ryono D.E., Bird J.E., Asaad M.M., Schaeffer T.R., Trippodo N.C. // *J. Med. Chem.* 1996. V. 39. P. 494–502.
15. Sugg E.E., Aquino C.J., Armour D.R., Berman J.M., Birkemo L.S., Carr R.A.E., Croom D.K., Dezube M., Dougherty R.W., Jr., Ervin G.N., Grizzle M.K., Heard J.E., Hirst G.C., James M.K., Johnson M.F., Miller L.J., Queen K.L., Rimele T.J., Smith T.H. // *J. Med. Chem.* 1996. V. 39. P. 562–569.
16. Geldern T.W., Hoffman D.J., Kester J.A., Nellans H.N., Dayton B.D., Calzadilla S.V., Marsh K.C., Hernandez L., Chiou W., Dixon D.B., Wu-Wong J.R., Oppenorth T.J. // *J. Med. Chem.* 1996. V. 39. P. 982–991.
17. Sikorski J.A., Devadas B., Nagarajan S.R., Zupec M.E., Freeman S.K., Brown D.L., Hwang-Fun Lu, Mehta P.P., Kishore N.S., McWherter C.A., Getman D.P., Godon J.I. // *J. Med. Chem.* 1997. V. 40. P. 1422–1438.
18. Veale C.A., Bernstein P.R., Bohnert C.M., Brown F.J., Bryant C., Damewood J.R., Jr., Earley R., Freeney S.W., Edwards P.D., Gomes B., Hulsizer J.M., Kosmider B.J., Krell R.D., Moore G., Salcedo T.W., Vacek E.P., Williams J.C., Wolanin D.J., Woolson S. // *J. Med. Chem.* 1997. V. 40. P. 3173–3181.
19. Gillet V.J., Willet P., Bradshaw J. // *J. Chem. Inf. Sci.* 1998. V. 38. P. 165–179.
20. Мартынова Н.Б., Поройков В.В. Тез. докл. на межд. науч. конф. "Органический синтез и комбинаторная химия". 4–7 марта 1999 г. Москва, Звенигород.
21. Geronikaki A., Poroikov V., Hadjipavlou-Litina D., Filimonov D., Lagunin A., Mgonzo R. // *Quant. Struct.-Activ. Relationships.* 1999. V. 18. P. 16–25.
22. Трапков В.А., Будунова А.П., Бурова О.А., Филимонов Д.А., Поройков В.В. // *Вопр. мед. химии.* 1997. Т. 43. С. 41–57.
23. Майборода Д.А., Бабаев Е.В., Гончаренко Л.В. // *Хим.-фарм. журн.* 1998. Т. 32. С. 24–28.
24. Поройков В.В. // *Химия в России.* 1999. Т. 2. С. 8–12.

Computer Prediction of Biological Activity Spectra for Low-Molecular Peptides and Peptidomimetics

N. B. Martynova[#], D. A. Filimonov, and V. V. Poroikov

Institute of Biomedical Chemistry, Russian Academy of Biomedical Sciences, Pogodinskaya ul. 10, Moscow, 119832 Russia

The wide variety of the biological effects of peptides and their high activity are the main reasons for the search for new basic drug structures among them. The most promising compounds can be selected using the PASS computer system (Prediction of Activity Spectra for Substances). This system was originally developed to predict the activities of low-molecular "drug-like" organic compounds. Its predictive capacity is described here by the example of 134 peptides and peptidomimetics with nine known biological activities. Its average predictive power is shown to be approximately 97%. Such an accuracy demonstrates that computer prediction can be applied both to the evaluation of effects and mechanisms of action of endogenous and synthetic peptides and to the screening of new therapeutic agents among the most promising basic structures.

Key words: peptides, peptidomimetics; biological activity; computer prediction, search for basic structures of new therapeutics

[#] To whom correspondence should be addressed; phone: +7 (095) 245-2753; e-mail: vvp@ibmh.msk.su.

For the full text of English translation of this article, see *Russian Journal of Bioorganic Chemistry*, 2000, vol. 26, no. 5. It is also available (free) in PDF format through the Web Site <http://www.maik.rssi.ru/>.