

БИОИНФОРМАТИКА

УДК 541.69 + 519.25 + 518.5

© Коллектив авторов

КОМПЬЮТЕРНЫЕ КЛАССИФИКАЦИОННЫЕ МОДЕЛИ ВЗАИМОСВЯЗИ СТРУКТУРЫ ХИМИЧЕСКИХ СОЕДИНЕНИЙ И ЛЕКАРСТВ С ИХ СПОСОБНОСТЬЮ ПРОНИКАТЬ ЧЕРЕЗ ГЕМАТОЭНЦЕФАЛИЧЕСКИЙ БАРЬЕР

О.А. Раевский^{1}, С.Л. Солодова¹, О.Е. Раевская¹, Я.В. Липлавский¹, Р. Манхольд²*

¹Институт физиологически активных веществ РАН, 142432 г. Черноголовка
Московской обл., Северный проезд, 1; тел./факс: (495)785-70-24;
эл. почта: raevsky@ipac.ac.ru

²Университет им. Генриха Гейне, 40225 Дюссельдорф, Германия

Способность лекарственных веществ проходить через гематоэнцефалический барьер (ГЭБ), обозначенная как BBB⁺ и BBB для не проникающих ГЭБ, является одним из важнейших свойств соединений, действующих на центральную нервную систему (ЦНС). В работе представлены результаты по моделированию взаимосвязи структуры органических лекарственных препаратов и их способности преодолевать ГЭБ. Использованы литературные данные по BBB⁺/– для 1513 соединений. Для указанного ряда веществ 1276 были отнесены к BBB⁺ и 237 к BBB. Компьютерное моделирование взаимосвязи структура-активность осуществлялось по двум направлениям: метод структурной аналогии ("read-across") и линейный дискриминантный анализ (ЛДА) на основе физико-химических дескрипторов. Главным параметром, определяющим проницаемость через ГЭБ, оказалась сумма водород-связывающих донорно-акцепторных факторов.

Ключевые слова: гематоэнцефалический барьер, структурное сходство, водородная связь, физико-химические дескрипторы, HUBOT

ВВЕДЕНИЕ. Часто нарушения деятельности ЦНС являются результатом влияния на организм различных эндогенных и экзогенных факторов. Для того чтобы ЦНС нормально и правильно функционировала в процессе эволюции организмом был выработан ряд собственных защитных механизмов, которые обеспечивают независимость работы ЦНС даже в агрессивных условиях. Одним из таких важнейших механизмов является гематоэнцефалический барьер (ГЭБ). Под этим термином условно объединяют совокупность анатомических и функциональных систем, которые поддерживают постоянство внутренней среды мозга и регулируют её состав [1, 2].

ГЭБ регулирует прохождение различных веществ между током крови и мозгом и обладает высокой избирательной проницаемостью. Барьер предотвращает проникновение вредных веществ из крови в мозг и активно выводит их в том случае, если они каким-то образом попали в мозг. ГЭБ играет важную роль в патологии таких заболеваний, как инсульты, микробные инфекции, черепно-мозговые травмы, нейродегенеративные заболевания, которые, в свою очередь, могут приводить к изменениям проницаемости ГЭБ. С другой стороны, ГЭБ непроницаем для многих лекарственных препаратов, что может служить препятствием при лечении заболеваний [2].

* - адресат для переписки

Молекулы веществ могут проходить через такую сложно устроенную систему двумя путями: за счёт пассивного и активного транспорта. Пассивная диффузия осуществляется по градиенту концентрации из тока крови в мозг без энергетических затрат. Активный транспорт осуществляется против градиента концентрации, такой перенос требует затраты энергии и осуществляется при помощи белков-транспортёров.

В связи с указанной важностью процесса проникновения лекарств через ГЭБ за последние 20 лет был проведен целый ряд исследований по установлению взаимосвязи между структурой органических соединений и их способностью проникать через ГЭБ. Здесь можно упомянуть об оценке проницаемости Сасо-2 [3], классификации действия соединений на систему ЦНС [4, 5], а также проницаемость через ГЭБ [6, 7]. Во всех указанных публикациях использовался целый набор различных дескрипторов и значимость и влияние этих дескрипторов на процесс транспорта веществ в организме продолжает оставаться предметом оживленных дискуссий. Следует отметить, что с годами росло и число соединений, включённых в расчёт. Так, в публикации [3] были использованы только 17 лекарств, в публикации [5] - 125, а в последующих работах уже сотни химических соединений и лекарств.

В целом сложилось впечатление, что основными факторами, влияющими на проницаемость веществ через ГЭБ, являются стерические взаимодействия, липофильность и способность соединений к образованию водородных связей.

Представленная в настоящем сообщении работа посвящена построению классификационной модели проницаемости веществ через ГЭБ путем пассивного транспорта (проникающие соединения обозначаются как BBB+ и непроникающие как BBB-) на основе рассмотрения данных для 1513 химических соединений и лекарств. Авторы рассматривают эту работу в качестве первого шага в системном исследовании этого чрезвычайно сложного биологического процесса, и главным результатом этой работы должно было стать выявление физико-химических дескрипторов, оказывающих влияние на изучаемую классификацию.

Использовались два независимых метода классификации соединений: метод структурной аналогии ("read-across") и линейный дискриминантный анализ (ЛДА) на основе физико-химических дескрипторов.

МЕТОДИКА.

Данные. Во многих работах проникновение химических соединений через ГЭБ отождествляется с непосредственным действием веществ на ЦНС. Нет сомнений в том, что вещества, оказывающие действие на ЦНС, конечно, проходят через барьер, однако существует и ряд веществ, которые проходят через ГЭБ, но не являются ЦНС-активными. По этой причине для настоящей работы требовался тщательный отбор литературных данных, конкретно связанных с проницаемостью через ГЭБ. Здесь за основу были взяты данные по проницаемости ГЭБ, отобранные и проанализированные в публикации [6]. При этом из указанного массива соединений были исключены данные по проницаемости 91 соединения, которые, согласно [6], проявляют свойства гликопротеиновых субстратов. В качестве веществ, проникающих через ГЭБ (BBB+), рассматривались соединения, имеющие величину $\log BB \geq -1$, а вещества непроникающие (BBB-) с $\log BB < -1$. Среди указанных 1513 соединений 1276 являются BBB+ и 237 BBB-.

Дескрипторы. С помощью имеющихся компьютерных программ (SYBYL, HYBOT и DRAGON) в данной работе был рассчитан целый ряд дескрипторов, связанных с различными видами межмолекулярных взаимодействий:

- Стерическими (MW, α , AREA).
- Электростатическими (ΣQ).

Ввиду хорошо известной исключительной роли водородной связи во взаимодействиях органических лигандов с биологическими мишенями в живых организмах в данной работе были привлечены дескрипторы водородной связи, рассчитываемые различными подходами и авторами:

- Число водородносвязанных доноров (HD) и акцепторов (HA) в молекулах [4].
- Общая водородносвязанная основность (B) и кислотность (A) [7].
- Водородносвязанные донорно-акцепторные факторы ($\Sigma Ca(o)$, ΣCd , $\Sigma Ca(o)d$) [8, 9].
- Поверхность полярных взаимодействий PSA [10].
- Поверхность полярных взаимодействий, рассчитанная с учётом силы водородной связи WEASA, WEDSA, WEASA + WEDSA [11].

Кроме того, в расчётах был использован в качестве дескриптора коэффициент распределения соединений в система вода-останол ($\log P$), который является широко используемым в литературе параметром, связанным с транспортом химических веществ [12].

В таблице 1 представлены все использованные в данной работе физико-химические дескрипторы, их обозначения и ссылки.

Таблица 1. Физико-химические дескрипторы.

N	Название дескриптора	Обозначение	Программа	Ссылка
1	Молекулярный вес	MW	HYBOT	[13]
2	Молекулярная полярность	A	HYBOT	[13]
3	Молекулярная поверхность	AREA	SYBYL	[14]
4	Сумма отрицательных зарядов	ΣQ^-	HYBOT	[13]
5	Число доноров водородной связи	HD	DRAGON	[16]
6	Число акцепторов водородной связи	HA	DRAGON	[16]
7	Общая водородносвязанная кислотность	A	Данные, рассчитанные Algorithm Builder [28] из публикации [7]	[7]
8	Общая водородносвязанная основность	B	Данные, рассчитанные Algorithm Builder [28] из публикации [7]	[7]
9	Сумма водородносвязанных акцепторных факторов молекулы	$\Sigma Ca(o)$	HYBOT	[13]
10	Сумма водородносвязанных донорных факторов молекулы	ΣCd	HYBOT	[13]
11	Абсолютное значение суммы водородносвязанных донорных и акцепторных факторов	$\Sigma Ca(o)d$	HYBOT	[13]
12	Водородносвязанная энтальпийная акцепторная поверхность на Ван дер Ваальсовых сферах	WEASA	3D HYBOT	[15]
13	Водородносвязанная энтальпийная донорная поверхность на Ван дер Ваальсовых сферах	WEDSA	3D HYBOT	[15]
14	Поверхность полярных взаимодействий	PSA	SYBYL	[14]
15	Липофильность	$\log P$	DRAGON	[16]

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.

Использование метода структурной аналогии ("read-across") для классификации химических соединений на BBB+ и BBB-. Метод аналогии производит оценку свойств какого-либо соединения на основе этих же свойств одного или нескольких сходных соединений [17, 18]. В англоязычной литературе этот метод именуется "read-across".

Подлинными основоположниками этого метода являются Д.И. Менделеев [19] и А.М. Бутлеров [20]. Стоит вспомнить, что свойства ещё не открытых элементов Менделеев предсказывал на основании их места: в периодической

КОМПЬЮТЕРНЫЕ КЛАССИФИКАЦИОННЫЕ МОДЕЛИ

системе, по свойствам их соседей справа и слева, сверху и снизу ("read-across"!!!), т.е. по соседним членам периода и группы.

Официальное современное определение метода аналогии ("read-across") формулируется следующим образом: "read-across" - это техника для предсказания целевой информации для какого-либо химического соединения, используя информацию о другом химическом соединении, которое рассматривается как сходное каким-либо образом (структурное сходство, сходство свойств или активности) [18].

Краткие основы метода аналогии и примеры его использования для классификации представлены в публикациях [21-24].

В простейшем случае в методе аналогии данные по свойству одного соединения приписываются другому сходному соединению. В настоящей работе в качестве сходства были выбраны структурные формулы, а степень сходства устанавливалась посредством расчета индексов Танимото (T_c):

$$T_c = C / (A + B - C),$$

где C - общие фрагменты структуры в паре сравниваемых соединений, A и B - число структурных фрагментов в каждом соединении. С целью отбора наиболее родственных структур для каждого из 1513 соединения все остальные 1512 соединений ранжировались в порядке их значений индекса Танимото. Производилась эта процедура с помощью программы MOLDIVS [25, 26]. Для классификации отбиралось нечётное число ближайших структурных аналогов. И рассматриваемому соединению приписывалось свойство, которое было у большинства отобранных аналогов. Результаты этой работы суммированы в таблице 2.

Таблица 2. Структурное сходство.

$T_c \geq$	Число ближайших структурных соседей	Число соединений			Правильное распознавание	
		BBB-	BBB+	BBB+/-	BBB- (%)	BBB+(%)
0,0	1	237	1276	1513	222(93,7)	1262(98,9)
0,0	3	237	1276	1513	222(93,7)	1268(99,4)
0,0	5	237	1276	1513	222(93,7)	1270(99,5)
0,1	1	237	1276	1513	222(93,7)	1262(98,9)
0,1	3	237	1276	1513	222(93,7)	1268(99,4)
0,1	5	237	1276	1513	222(93,7)	1270(99,5)
0,2	1	237	1273	1510	222(93,7)	1259(98,9)
0,2	3	237	1266	1503	222(93,7)	1258(99,4)
0,2	5	237	1258	1495	222(93,7)	1252(99,5)
0,3	1	234	1241	1475	219(93,6)	1227(98,9)
0,3	3	232	1159	1391	219(94,4)	1152(99,4)
0,3	5	225	1105	1330	210(93,3)	1100(99,5)
0,4	1	224	1114	1338	213(95,1)	1103(99,0)
0,4	3	208	900	1108	201(96,6)	899(99,9)
0,4	5	202	770	972	196(97,0)	769(99,9)
0,5	1	218	941	1159	210(96,3)	932(99,0)
0,5	3	188	618	806	185(98,4)	617(99,8)
0,5	5	161	478	639	159(98,8)	478(100)

В качестве первого шага классификация была выполнена без какого-либо порогового значения индекса Танимото для структурных аналогов. Это позволило включить в классификацию все 1513 соединений, включая те, которые не имели "хороших" структурных аналогов. Как видно из таблицы 2, уже использование

одного ближайшего структурного аналога привело к хорошей классификации: правильное отнесение было сделано для 222 BBB- соединений (93,7%) и для 1262 BBB+ соединений (98,9%). Дальнейшее привлечение трех и пяти ближайших структурных аналогов еще улучшили распознавание BBB+ соединений (до 99,4% и 99,5% соответственно).

Повышение порога структурного сходства приводит к улучшению правильного распознавания как BBB+, так и BBB- соединений. Так, при пороге $T_s \geq 0,50$ и использовании пяти ближайших структурных аналогов правильно распознались 100% BBB+ соединений и 98,8% BBB- соединений. Однако, при этом общее число, вовлечённых в классификацию соединений резко сократилось (до 161 BBB+ и 478 BBB- соединений).

Для рассматриваемого ряда соединений оптимальным вариантом классификации, по-видимому, является использование трёх ближайших структурных аналогов без ограничений порога структурного сходства. Это позволяет провести классификацию всех соединений и правильно распознать свойства 98,5% соединений (1490 из 1513).

Полученные хорошие результаты классификации BBB+ и BBB- соединений методом структурной аналогии позволяют предположить наличие в молекулах структурных фрагментов, которые преимущественно находятся в соединениях, проходящих и не проходящих через ГЭБ. Для определения таких структурных фрагментов была оценена их пропорция в соединениях BBB+ и BBB- с помощью программы MOLDIVS [25]. При фрагментации с использованием второй координационной сферы данный массив содержал 3870 фрагментов. Поскольку данный ряд (1513 соединений) состоял из 237 BBB- соединений (что составляет 15,7% от общего числа) и 1276 BBB+ соединений (84,3%), то естественно предположить, что фрагменты не влияющие на классификацию BBB+ и BBB- должны быть распределены с таким же соотношением. В таблице 3 представлены структурные фрагменты, пропорция которых в BBB- больше 30%, а в BBB+ меньше 70% соединениях существенно отличается от указанного соотношения (15,7% и 84,3%).



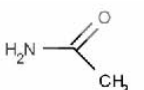
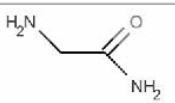
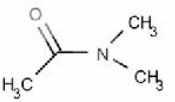
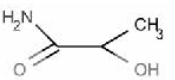
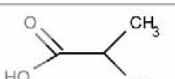
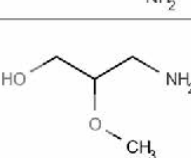
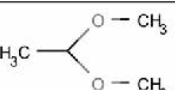
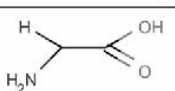
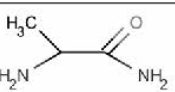
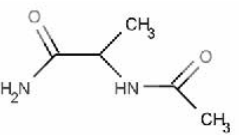
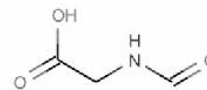
Как видно из таблицы 3, структурные фрагменты, имеющиеся преимущественно в BBB- соединениях, насыщены водородносвязанными донорными и акцепторными группами. Это согласуется с высказанным ранее нами и другими авторами предположением о сильном влиянии водородных связей на процессы транспорта органических соединений в живых организмах. Для дальнейшего обсуждения этого вопроса в настоящей работе были привлечены различные физико-химические дескрипторы.

Взаимосвязи физико-химических дескрипторов с проницаемостью через ГЭБ. На первом этапе этой работы каждый из приведенных в таблице 1 физико-химических дескрипторов ранжировался в порядке их рассчитанных значений. Затем колонка этих значений сопоставлялась с колонкой, содержащей информацию о свойстве (BBB+ или BBB-). Это дало возможность сравнить на качественном уровне способность различных физико-химических дескрипторов дискриминировать соединения BBB+ и BBB-.

В качестве примера на рисунках 1 и 2 представлены зависимости доли BBB+ соединений от липофильности ($\log P$) и суммы водородносвязанных факторов. В случае использования дескриптора $\log P$ зависимость доли BBB+ соединений носит S-образный вид и не дает возможности дискриминировать BBB+ и BBB- соединения. Совершенно иная картина наблюдается при использовании суммы водородносвязанных факторов. В интервале этих факторов от 0,00 до 5,00 все соединения (их 187!) являются BBB+. А при $\sum Ca(o)d \geq 25,00$ все соединения являются BBB-. В интервале же 5,00–25,00 происходит постепенное уменьшение доли BBB+ соединений по мере увеличения значений этого дескриптора. Тем самым очевидна хорошая способность этого водородносвязанного дескриптора дискриминировать соединения BBB+ и BBB-.

КОМПЬЮТЕРНЫЕ КЛАССИФИКАЦИОННЫЕ МОДЕЛИ

Таблица 3. Структурные фрагменты.

Функциональная группа	BBB+/- (1513)	BBB- (237) (15,7%)	BBB+ (1276) (84,3%)
	417	171(41,0%)	246(59,0%)
	17	7(41,2%)	10(58,8%)
	565	175(30,9%)	390(69,1%)
	235	132(56,2%)	130(43,8%)
	309	130(42,1%)	179(57,9%)
	35	22(62,9%)	13(37,1%)
	79	62(78,5%)	17(21,5%)
	58	36(62,1%)	22(37,9%)
	112	62(55,4%)	50(44,6%)
	153	131(85,6%)	22(14,4%)
	164	131(79,9%)	33(20,1%)
	128	120(93,8%)	8(6,2%)
	140	129(92,1%)	11(7,9%)

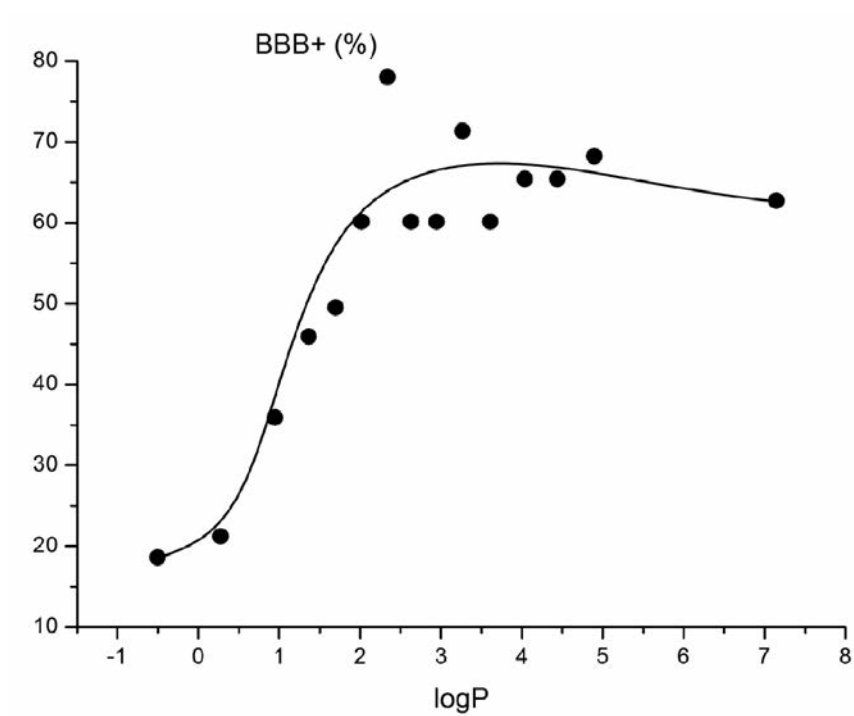


Рисунок 1.

Графическая зависимость доли BBB+ соединений от коэффициента распределения в системе октанол-вода.

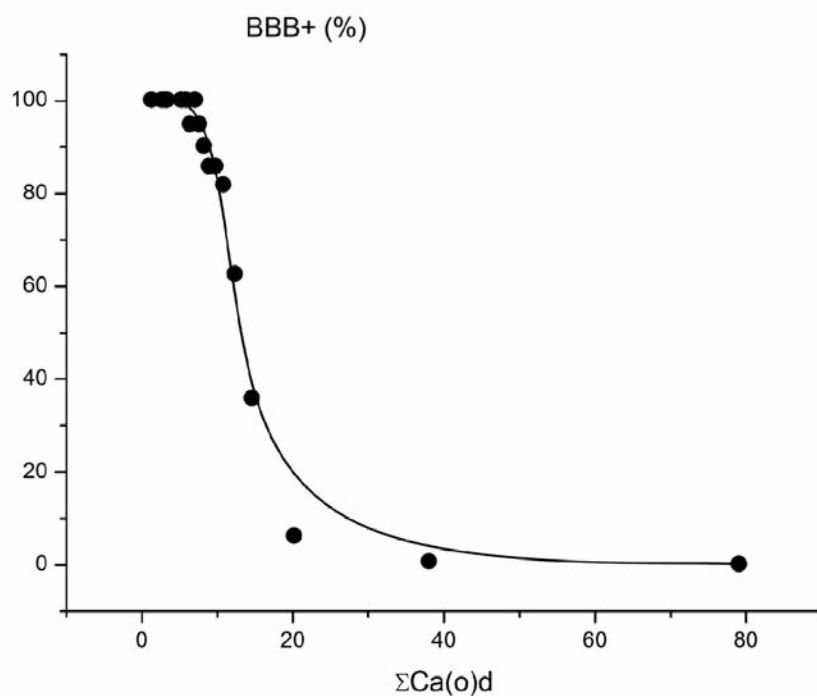


Рисунок 2.

Графическая зависимость доли BBB+ соединений от суммы водородно связанных факторов.

На заключительном этапе работы для классификации BBB+/BBB- соединений были использованы метод линейного дискриминантного анализа (ЛДА) и программа SPSS Statistics 17.0 [27]. Результаты представлены в таблице 4.

Таблица 4. Параметры и статистические критерии дискриминантной функции $y = a_0 + a_1x_1 + \dots + a_nx_n$.

x_1	x_2	x_3	x_4	a_0	A_1	a_2	a_3	a_4	λ	F	p	D ²	TP	Чувствительность, %	FN	TN	Специфичность, %	FP	Точность, %
MW	-	-	-	-2,89	0,008	-	-	-	0,766	461,2	<0,001	2,31	1042	81,7	234	169	71,3	68	80,0
α	-	-	-	-2,96	0,81	-	-	-	0,845	276,2	<0,001	1,38	957	75,0	319	148	62,4	89	73,0
area	-	-	-	-3,79	0,006	-	-	-	0,924	123,5	<0,001	0,62	903	70,8	373	140	59,1	97	68,9
ΣQ	-	-	-	2,18	1,25	-	-	-	0,683	701,1	<0,001	3,51	1132	88,7	144	149	62,9	88	84,7
HD	-	-	-	0,817	0,473	-	-	-	0,842	282,8	<0,001	1,42	1097	86,0	179	137	57,8	100	81,6
HA	-	-	-	-1,66	0,319	-	-	-	0,775	437,6	<0,001	2,19	1085	85,0	191	155	65,4	82	82,0
HD+HA	-	-	-	-1,44	0,209	-	-	-	0,772	445,4	<0,001	2,23	1158	90,8	118	148	62,4	89	86,3
HD+HA	MW	-	-	-2,68	0,126	0,005	-	-	0,689	340,7	<0,001	3,41	1132	88,7	144	163	68,8	74	85,6
HD+HA	MW	ΣQ	-	2,01	-0,089	0,002	1,169	-	0,643	279,9	<0,001	4,21	1152	90,3	124	167	70,5	70	87,2
A	-	-	-	-0,95	2,27	-	-	-	0,831	307,4	<0,001	1,54	1096	85,9	180	152	64,1	85	82,5
B	-	-	-	-1,92	1,08	-	-	-	0,801	376,3	<0,001	1,88	1098	81,6	178	150	63,3	87	82,5
A+B	-	-	-	-0,228	1,228	-	-	-	0,781	423,5	<0,001	2,12	1145	89,7	131	150	63,3	87	85,6
A+B	A	-	-	-2,67	0,587	0,038	-	-	0,743	260,7	<0,001	2,61	1111	87,1	165	153	64,4	84	83,5
A+B	A	ΣQ	-	1,64	-0,30	0,042	1,44	-	0,633	292,2	<0,001	4,39	1153	90,4	123	166	70,0	71	87,2
$\Sigma Ca(0)$	-	-	-	-2,204	0,319	-	-	-	0,573	1124,9	<0,001	5,63	1207	94,6	69	168	70,9	69	90,9
ΣCd	-	-	-	1,09	0,348	-	-	-	0,535	1313,9	<0,001	6,57	1227	96,2	49	165	69,6	72	92,0
$\Sigma Ca(0)d$	-	-	-	-1,91	0,190	-	-	-	0,448	1583,8	<0,001	7,92	1250	98,0	26	182	76,8	55	94,6
$\Sigma Ca(0)d$	A	-	-	-1,67	0,200	-0,010	-	-	0,486	799,7	<0,001	8,01	1247	97,7	29	182	76,8	55	94,4
$\Sigma Ca(0)d$	A	ΣQ	-	-1,71	0,208	-0,005	0,118	-	0,485	800,1	<0,001	8,01	1247	97,7	29	182	76,8	55	94,4
ΣCd	ΣCd	-	-	1,80	-0,159	0,223	-	-	0,486	789,5	<0,001	7,99	1248	97,8	28	180	75,9	57	94,4
$\Sigma Ca(0)$	ΣCd	α	-	1,69	-0,183	0,213	0,007	-	0,486	798,4	<0,001	7,99	1246	97,6	30	182	76,8	55	94,4
$\Sigma Ca(0)$	ΣCd	α	ΣQ	-1,71	0,169	-0,215	0,004	0,10	0,485	798,5	<0,001	7,99	1247	97,7	29	181	76,4	56	94,4
WEASA	-	-	-	-1,89	0,136	-	-	-	0,759	480,9	<0,001	2,41	1136	89,0	140	155	65,4	82	85,3
WEDSA	-	-	-	-0,89	0,310	-	-	-	0,799	380,4	<0,001	1,90	1101	86,3	175	160	67,5	77	83,3
WEASA+	-	-	-	-1,69	0,101	-	-	-	0,746	514,9	<0,001	2,58	1158	90,8	118	157	66,2	80	86,9
WEDSA	-	-	-	-2,62	0,079	0,036	-	-	0,708	311,7	<0,001	3,12	1170	91,7	106	169	71,3	68	88,5
WEASA+	A	-	-	1,71	-0,049	0,039	1,34	-	0,609	323,3	<0,001	4,86	1167	91,5	109	179	75,5	58	88,9
WEDSA	A	ΣQ	-	-1,41	0,018	-	-	-	0,732	554,5	<0,001	2,77	1145	89,7	131	155	65,4	82	85,9
PSA	-	-	-	-2,40	0,015	0,035	-	-	0,692	335,3	<0,001	3,36	1128	88,4	148	163	68,8	74	85,3
PSA	A	ΣQ	-	1,62	-0,009	0,036	1,28	-	0,603	331,8	<0,001	4,99	1148	89,9	128	176	74,3	61	87,5
logP	-	-	-	-1,18	0,491	-	-	-	0,921	130,6	<0,001	0,65	885	69,4	391	146	61,6	91	68,1

Примечание. a_0, a_1, \dots, a_n - коэффициенты функции, x_1, x_2, \dots, x_n - дескрипторы, λ - лямбда Уилкса (изменяется от нуля при идеальной дискриминации до единицы при отсутствии дискриминации), F - критерий Фишера, p - значимость, D² - расстояние Михаланобиса (мера разделения двух групп), TP - число правильно распознанных соединений среди BBV+, TN - число правильно распознанных соединений среди BBV-, FP - число неправильно распознанных соединений среди BBV-, FN - число неправильно распознанных соединений среди BBV+, Чувствительность = TP / (TP + FN), Специфичность = TN / (TN + FP), Точность = (TP + TN) / (TP + TN + FN + FP).

Как очевидно из таблицы 4, наилучший результат классификации с помощью одиночной независимой переменной был получен при использовании НУВОТ-дескриптора $\Sigma Ca(o)d$ (минимальное значение λ Уилкса, максимальное значение критерия Фишера, максимальная чувствительность и точность). Другие водородносвязанные дескрипторы также обеспечивают достаточно хорошую классификацию, показывая несколько худшие статистические критерии дискриминантной функции. Дескриптор, отвечающий за электростатические контакты (ΣQ -), провёл дискриминацию с точностью 84,7%. Одиночное использование дескрипторов, связанных со стерическими взаимодействиями (MW, α , AREA, соответственно) показали средний результат (точность до 80%). При использовании параметра гидрофобности logP был получен наихудший результат классификации.

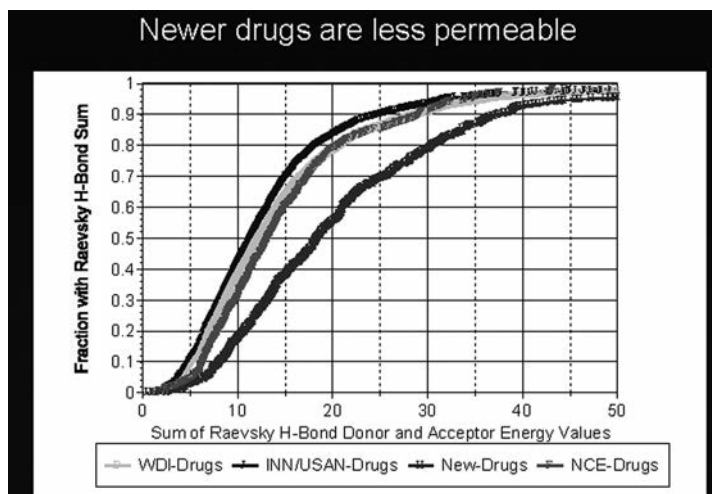


Рисунок 3.

Графическое сравнение фракции химических соединений в пределах рассмотренных в [29] четырёх обширных баз данных, содержащих многие тысячи соединений: Derwent World Drug Index (WDI), International Non-Proprietary Names (INN)/US Adopted Names (USAN), New Chemical Entities (NCE), New Drugs. от $\Sigma Ca(o)d$. Из рисунка очевидно, что только 40% соединений из базы "New Drugs" имеет значение $\Sigma Ca(o)d \leq 15,00$ (примерный порог проницаемости), в то время как для других трех баз данных при указанном значении $\Sigma Ca(o)d$ фракция таких веществ достигает 70%.

Высокая закоррелированность водородносвязанных дескрипторов между собой не позволяет включать в дискриминантный анализ какое-либо их сочетание. Присоединение же к какому-либо водородносвязанному дескриптору параметров, связанных со стерическими и/или электростатическими взаимодействиями немного улучшает статистику. Однако, модели, основанные на этих комбинациях, уступают по статистическим критериям классификации модели с использованием в качестве одиночного дескриптора $\Sigma Ca(o)d$ (точность 94,6%). Модели же с участием $\Sigma Ca(o)d$ и α и/или ΣQ - не увеличили точности классификации, что естественно в виду близости указанных значений к предельной точности самих использованных данных.

ВЫВОДЫ.

1. Методы структурной аналогии и дискриминантного анализа на базе водородносвязанных дескрипторов позволили успешно провести классификацию 1513 разнообразных химических соединений по их способности проникать через ГЭБ по пассивному механизму транспорта.

2. Подтверждается предположение Липинского [29], что сумму водородносвязанных донорных и акцепторных факторов ($\Sigma Ca(o)d$) можно рассматривать в качестве глобальной меры проницаемости веществ в живых организмах и что этот дескриптор должен использоваться при отборе новых эффективных веществ с хорошими параметрами проницаемости (см. рис. 3 из [29]).

Работа выполнена при поддержке МНТЦ (проект 3777) и Минобрнауки (Госконтракт №07.514.11.4118.).

ЛИТЕРАТУРА

1. Бредбери М. (1983) Концепция гематоэнцефалического барьера (пер. с англ.), Медицина, М.
2. Покровский В.М., Коротько Г.Ф., Кобрин В.И. и др. (1997) Физиология человека: Учебник, Медицина, М.
3. Waterbeemd H., Gamenisch G., Folkers G., Raevsky O.A. (1996) Quant. Struct.-Act. Relat., **15**, 480-490.
4. Lipinski C.A., Lombardo F., Dominy B.W., Feeney P.J. (1997) Adv. Drug Delivery Rev., **23**, 3-25.
5. van de Waterbeemd H., Camenisch G., Folkers G., Chretien J.R., Raevsky O.A. (1998) J. Drug Targeting, **6**, 151-165.
6. Adenot M., Lahana R.J. (2004) Chem. Inf. Comput. Sci., **44**, 239-248.
7. Zhao Y.H., Abraham M.H., Ibrahim A., Fish P.V., Cole S., Lewis M.L., de Groot M.J., Reynolds D.P. (2007) J. Chem. Inf. Model., **47**, 170-175.
8. Raevsky O.A. (2004) Minireview in Medicinal Chemistry, **4**, 1041-1052.
9. Raevsky O.A. (2008) In: Molecular Drug Properties. Measurement and Prediction (Mannhold R., ed.). Wiley-VCH Verlag GmbH & Co.KGaA, Weinheim, pp. 127-154.
10. Ertl P. (2008) Polar surface area In: Molecular Drug Properties. Measurement and Prediction (Mannhold R., ed.) Wiley-VCH Verlag GmbH & Co.KGaA, Weinheim, pp. 127-154.
11. Raevsky O.A., Skvortsov V.S. (2005) SAR & QSAR in Environmental Research, **15**, 287-300.
12. Hansch C., Leo A., Hoekman D. (1995) Exploring QSAR: Hydrophobic, Electronic, and Steric Constants. American Chemical Society, Washington, DC.
13. Раевский О.А., Григорьев В.Ю., Трепалин С.В. (26.02.1999) Программа HYBOT, Роспатент № 990090.
14. SYBYL: SYBYL 8.0, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.
15. Раевский О.А., Скворцов В.С. (27.09.2004) Программа 3D HYBOT, Роспатент № 2004612207.
16. DRAGON (<http://www.vcclab.org/lab/edragon/>)
17. Organization for Economic Cooperation and Development. Guidance on Grouping of Chemicals, OECD, Paris (2007).
18. European Chemical Agency, Guidance on Information Requirements and Chemical Safety Assessment, Chapter R6, QSAR and grouping of chemicals, ECHA, Helsinki (2008).
19. Mendelejeff D. (1869) Zeitschrift fur Chemie, **12**, 405-406.
20. Butlerov A.M. (1861) Z. Chem. Pharm., **4**, 549-555.
21. Dimitrov S., Mekenyan O. (2010) An Introduction to Read-Across for the Prediction of the Effects of Chemicals, in "In Silico Toxicology. Principles and Applications". (Cronin M.T.D., Madden J.C., eds.), RSC, Cambridge, pp. 372-384.
22. Diderich R. (2010) Tools for Category Formation and Read-Across: Overview of the OECD (Q)SAR Application Toolbox, *ibid.*, pp. 385-407.
23. Jeliaskova N., Jaworska J., Worth A.P. (2010) Open Source Tools for Read-Across and Category Formation, *ibid.*, pp. 408-445.
24. Raevsky O.A., Grigor'ev V.Yu., Modina E.A., Worth A. (2010) SAR & QSAR in Environmental Research, **21**, 265-275.
25. Раевский О.А., Герасименко В.А., Трепалин С.В. (26.02.1999) Программа MOLDIVS, Роспатент № 990093.
26. Gerasimenko V.A., Trepalin S.V., Raevsky O.A. (2000) MOLDIVS - A New Program for Molecular Similarity and Diversity Calculations, in "Molecular Modelling and Prediction of Bioactivity", (Gundertofe K., Jorgensen F., eds.), Kluwer Academic/Plenum Publ., pp. 423-424.

27. SPSS (<http://www.spss.com>)
28. PharmaAlgorithms. ADME Boxes, Version 3.0; PharmaAlgorithms, Inc.: 591 Indian Road, Toronto, ON M6P 2C4, Canada.
29. *Lipinski C.A.* Drugs structure and properties, past and present. Can we design drugs with beautiful properties? http://www.iaimn.demon.co.uk/spring99/lipins_n.pdf

Поступила: 05. 03. 2011.

**COMPUTER CLASSIFICATION MODELS ON THE RELATIONSHIP BETWEEN
CHEMICAL STRUCTURES OF COMPOUNDS AND DRUGS WITH THEIR
BLOOD BRAIN BARRIER PENETRATION**

O.A. Raevsky¹, S.L. Solodova¹, O.E. Raevskaya¹, Y.V. Liplavskiy¹, R.M. Mannhold²

¹Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds of Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia; tel.: (495)785-70-24; e-mail: raevsky@ipac.ac.ru

²Molecular Drug Research Group, Medical Faculty; Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany

Ability of drugs to cross blood-brain barrier (BBB) (BBB+ for BBB-penetrating and BBB- for non-penetrating compounds) is one of the most important properties of chemicals acting on the central nervous system (CNS). This work presents the results of modelling of the relationship between chemicals structure and BBB-crossing ability. The data set included 1513 compounds BBB+/- (1276 BBB+ and 237 BBB-). Computer modelling of structure-activity relationship was realized by two directions: using the "read-across" method and linear discriminant analysis (LDA) based on physico-chemical descriptors. It was found that a sum of donor-acceptor factors is the principal parameter, which define BBB penetration.

Key words: blood-brain barrier, structural similarity, hydrogen bound, physico-chemical descriptors, HYBOT.