

П. Г. Полищук, А. П. Косинская, В. Б. Ларионов, Л. Н. Огниченко,
В. Е. Кузьмин, Н. Я. Головенко

РАНЖИРОВАННЫЙ РЯД ФРАГМЕНТОВ МОЛЕКУЛ ЛЕКАРСТВЕННЫХ СРЕДСТВ, ОПРЕДЕЛЯЮЩИХ ИХ НЕЙРОДОСТУПНОСТЬ

Физико-химический институт им. А. В. Богатского НАН Украины, Украина, Одесса

Использование методов компьютерного моделирования на основе зависимости “структура — свойство” (QSPR) позволяет рационально планировать как проведение экспериментальных исследований, так и осуществлять эффективное генерирование структур, обладающих способностью проникать через ГЭБ. Целью работы было создание моделей “структура — свойство” на основе симплексного представления молекул для внеэкспериментальной оценки их способности проникать через ГЭБ, и проведение структурной интерпретации построенных моделей, учитывающей взаимное влияние атомов в молекуле. С использованием метода симплексного представления молекулярной структуры был оценен вклад различных молекулярных фрагментов и функциональных групп в величину нейродоступности диффундирующих через ГЭБ соединений. Показано, что наличие сильно полярных и ионогенных групп (карбоксильная, карбонильная, фенольный гидроксил) негативно сказывается на способности веществ проникать в мозг. Наличие атомов галогена и ароматических фрагментов положительно влияет на эффективность переноса молекул через ГЭБ. Построенная QSPR модель характеризуется удовлетворительной прогнозирующей способностью, оцененной согласно внутренней процедуре валидации моделей RandomForest ($R^2 = 0,54$, RMSE = 0,47).

Ключевые слова: гематоэнцефалический барьер; диффузия; симплекс; QSPR.

В процессе разработки (конструирования) лекарственных средств, регулирующих функции центральной нервной системы (ЦНС), необходимо учитывать способность соединений проникать (в количествах, достаточных для создания эффективной концентрации) через гематоэнцефалический барьер (ГЭБ), то есть определять их нейродоступность. К настоящему времени экспериментальное определение проницаемости ГЭБ для различных химических соединений является достаточно разработанной операцией, которая имеет большое количество возможных экспериментальных решений в зависимости от свойств исследуемого вещества и характера этого процесса (межклеточная, чрезклеточная диффузия или транспорт с помощью переносчиков). Однако из-за высокой стоимости и трудоемкости проводимых исследований возникла необходимость, а с накоплением значительного количества экспериментальных данных стало возможным математическое моделирование некоторых (или суммарного) процессов поступления веществ в головной мозг на основании особенностей их химической структуры. Использование методов компьютерной химии и компьютерного моделирования на основе зависимости “структура — свойство” (QSPR) позволяет рационально планировать как проведение экспериментальных исследований, так и осуществлять эффективное конструирование структур, обладающих способностью проникать через ГЭБ.

Для веществ, поступающих в ЦНС в результате простой диффузии, особое значение имеют такие физико-химические свойства как размер молекулы, способность к образованию водородных связей, величина липофильности и полярной поверхности молекулы. Для предварительной оценки и прогнозирования этих показателей был предложен ряд моделей [1, 2], базирующихся на описании конечных свойств через наборы молекулярных дескрипторов, наличие и вклад оп-

ределенных атомов и групп. Однако существующие модели являются по своей сути аддитивными и не учитывают взаимного влияния атомов в молекуле. Кроме того, прогностическая мощь данных моделей ограничена структурно или функционально однородными выборками соединений, которые использовались при построении этих моделей. В связи с этим ощущается недостаток моделей, обладающих удовлетворительной прогнозирующей способностью и которые бы одновременно с этим имели четкую структурную интерпретацию, учитывающую взаимное влияние атомов в молекуле.

Целью данной работы было создание моделей “структура — свойство” на основе симплексного представления молекул для внеэкспериментальной оценки их способности проникать через ГЭБ, и проведение структурной интерпретации построенных моделей, учитывающей взаимное влияние атомов в молекуле.

Экспериментальная часть

В качестве основного инструмента исследования количественной связи структура — свойство был использован метод симплексного представления молекулярной структуры [3]. Сущность метода заключается в описании молекулярной структуры симплексными дескрипторами, отражающими количество идентичных четырехатомных фрагментов (симплексов) фиксированной структуры, симметрии и хиральности. Важными преимуществами симплексного метода является возможность анализа в пределах одной выборки молекул, которые значительно отличаются по структуре. Принципиальной особенностью симплексного подхода является возможность учета как природы атома, так и ряда других характеристик (например, заряда, липофильности, поляризуемости и т.п.), оказывающих влияние на способность веществ к диффузии через биологические мембраны. В данной работе дифферен-

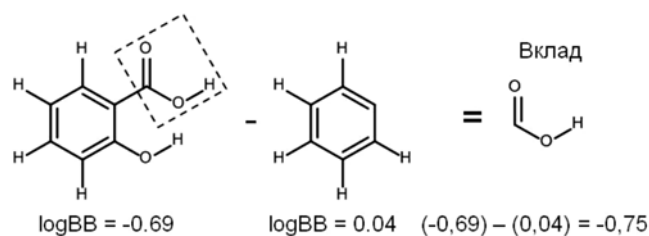


Рис. 1. Схема расчета вкладов молекулярных фрагментов, исходя из QSAR модели.

циация атомов в симплексах проводилась на основе следующих характеристик: 1) частичный заряд на атоме [4]; 2) липофильность [5]; 3) рефракция [6]; 4) возможность атома выступать донором/акцептором водорода при образовании водородной связи. Всего рассчитано 9519 2D симплексных дескрипторов.

Для построения QSAR моделей использован метод RandomForest [7] из пакета randomForest программы R [8]. Этот метод хорошо себя зарекомендовал при решении различных QSAR задач [9, 10] и обладает рядом таких преимуществ, как быстрое построение моделей, устойчивая внутренняя процедура оценки предсказательной способности модели и малое число настроечных параметров. К недостаткам метода можно отнести трудность интерпретации получаемых моделей. Для моделей RandomForest можно оценить лишь важность отдельно взятых дескрипторов. Однако этой информации зачастую недостаточно, чтобы отразить влияние отдельных структурных фрагментов на исследуемое свойство. В связи с этим использован новый методический подход [11], в соответствии с которым для исходного соединения предсказывается величина свойства, после чего из структуры соединения исключается интересующий фрагмент и для оставшегося фрагмента той же QSAR моделью предсказывается значение свойства. Разница между этими значениями составляет вклад интересующего фрагмента в изменение рассматриваемого свойства (рис. 1).

Для количественного описания проницаемости ГЭБ могут быть использованы такие величины, как максимальная скорость материального потока вещества через единицу площади ГЭБ, коэффициент диффузии (удельная скорость) или количество поступившего в ткани мозга исследуемого соединения. Однако наиболее информативным является показатель \log_{BB} — логарифм соотношения концентраций исследуемого соединения в мозге и крови, интегральный показатель способности вещества распределяться между транспортной тканью (кровь) и биофазой действия (головной мозг).

Объектом исследования послужила выборка различных классов химических соединений (свыше 900), для которых имеются данные параметров \log_{BB} [12 – 22]. Однако часть соединений была представлена одновременно в разных статьях и, следовательно, собранная выборка содержала большое число дубликатов. После идентификации и удаления дубликатов с использованием InstantJChem и исключения соединений, проходящих через ГЭБ с помощью Pgp переносчика, в выборке осталось 325 соединений, для которых экспе-

риментальные значения \log_{BB} , измеренные в тесте пассивной диффузии, варьировались в диапазоне от $-2,2$ до $1,6$.

Среди оставшихся 325 соединений присутствовали представители различных классов: ациклические углеводороды (гептан, октан, изогексан, октен-1 и т.д.), карбо- и гетероциклические соединения (кофеин, бромцистин и т.д.), соединения с различными функциональными группами (салициловая кислота, этанол, пентанон и т.д.). Большинство же соединений являются лекарственными препаратами, в основном действующими на ЦНС, обладающие успокаивающим, снотворным, противосудорожным, миорелаксирующим, амнестическим, противозипилептическим (диазепам, карбамазепин, хлорпромазин, габапентин, золпидем), болеутоляющим и жаропонижающим действием (аспирин, ибупрофен и т.д.), противовирусные препараты (зидовудин, ставудин и т.д.).

Результаты и их обсуждение

Первое QSAR моделирование проницаемости ГЭБ осуществлено в 1988 г. [20]. Методом линейной регрессии сделан анализ влияния липофильности на активность H_2 антагонистов центрального действия. С использованием выборки из 6 соединений получено уравнение зависимости \log_{BB} от величины $\Delta \log P$ (разность между величинами распределения вещества в октанол и циклогексане). Авторами сделан вывод о возможности увеличения проникновения через ГЭБ посредством уменьшения количества водородных связей.

Другие авторы [21, 22], используя различные структурные параметры, предложили обобщенное уравнение сольватации. Важным выводом явилось то, что $\log P$ не может быть использован для прямого влияния на распределение между кровью и мозгом, однако наряду с поляризуемостью/дипольностью и основностью/кислотностью является одним из основных факторов, определяющих течение данного процесса. Вместе с тем эффективное прогнозирование способности молекул биологически активных соединений к простой диффузии через ГЭБ может быть реализовано только с учетом комплексного влияния нескольких параметров.

Простая диффузия является одним из основных путей поступления лекарственных веществ в мозг. Гидрофобный характер ГЭБ ограничивает диффузию гидрофильных соединений из крови, однако на поступление липофильных веществ это оказывает меньшее влияние. Также в значительной степени процесс диффузии ограничен специфической структурой ГЭБ — наличием плотных контактов тесно примыкающих друг к другу эндотелиальных клеток. В связи с этим эффективность процесса простой диффузии лекарственных препаратов через ГЭБ обусловлена комплексом физико-химических факторов, определяемых структурой их молекулы.

Анализ показателей \log_{BB} показал (рис. 2), что наименьшей проникающей способностью обладают вещества, в молекулах которых имеются такие полярные группы, как карбоксильная, фрагменты, содержащие

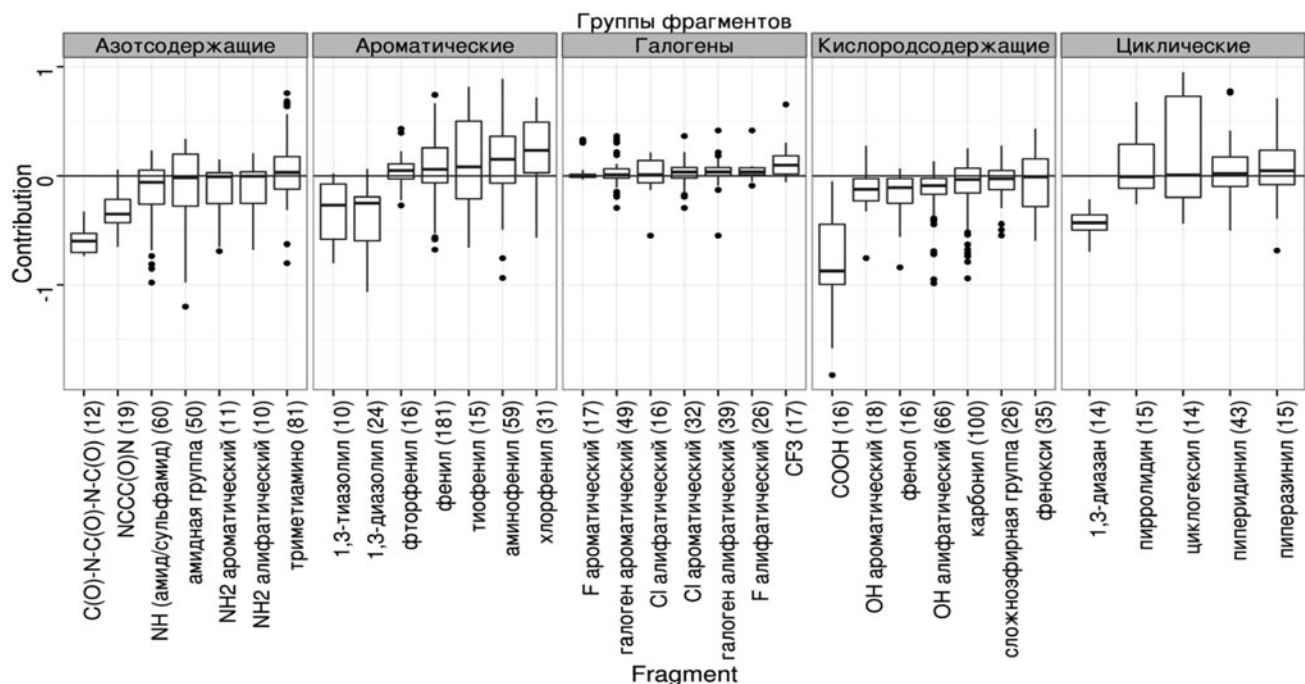


Рис. 2. Ряд структурных фрагментов, сгруппированный по типу входящих атомов и ранжированный по величине их вклада в изменение величины logBB соединений обучающей выборки (в скобках указано количество соединений, содержащих указанный фрагмент).

амидную группу (NCCC(=O)N, и фрагмент барбитуровой кислоты), а также ряд гетероциклических фрагментов (тиазолил, диазолил, фрагмент 1,3-диазана). По-видимому, их отрицательное влияние связано с тем, что атомы данных фрагментов, способны к образованию водородных связей (N, C=O) и к ионизации (COOH).

В целом способность атомов к образованию водородных связей в большинстве случаев приводит к отрицательному влиянию на диффузию соединений, содержащих соответствующие фрагменты, через ГЭБ. Так фрагменты, содержащие атомы, способные выступать акцепторами водородной связи, заметно снижают величину logBB (тиазолил, диазолил и фрагмент барбитуровой кислоты, а также фрагменты, содержащие амидную, сульфонамидную или гидроксильную группы). Аналогичная картина наблюдается и в случае выраженности частичного заряда на атоме анализируемого фрагмента. Положительное влияние на logBB отмечается только для симплексов, включающих трифторметильный, аминофенильный и хлорфенильный фрагменты. Вместе с тем для большинства подобных групп обратимая ионизация не является характерной, вследствие чего даже изменение поляризации молекул не только не приводит к уменьшению способности диффундировать через ГЭБ, но и сочетается с их положительным вкладом в этот процесс.

Группировка анализируемых фрагментов по природе входящих в их состав атомов позволила установить, что наличие атома галогена в структуре, вне зависимости от углеродного скелета, положительно влияет на способность молекул к диффузии (рис. 2). Вместе с тем по абсолютной величине вклада в logBB атомы галогенов можно отнести к фрагментам с низкой степенью влияния. Таким образом, можно прийти к заклю-

чению, что высокая электроотрицательность атомов галогенов и повышенная поляризуемость не оказывают значительного влияния на межмолекулярные взаимодействия, уменьшающие скорость диффузии.

В анализируемой выборке соединений фрагментам, содержащим кислород (в разных функциональных группах), соответствует отрицательное влияние на способность веществ диффундировать через ГЭБ (рис. 2). Если в случае групп кислотного характера, склонных к ионизации в физиологических условиях (-0,87 для карбоксильной и -0,11 для фенольной), это может быть следствием образования анионов, не проникающих через липидные биомембраны, то для алифатической гидроксигруппы, карбонильной (C=O) или сложноэфирной (C(=O)O) группы отрицательный вклад может быть следствием возможности образования водородных связей. Вместе с тем, принимая во внимание незначительное влияние данных функциональных групп на logBB, образование водородных связей посредством атомов кислорода можно рассматривать как второстепенное.

Одним из условий эффективности процесса диффузии лекарственных средств через биомембраны и, как следствие, нейродоступности, является их способность к обратимой ионизации и образованию гидрофильных и липофильных форм, что обеспечивается в данном случае атомами азота (рис. 2). Обращает внимание тот факт, что структурам, содержащим атом азота, способный к протонированию (аминный азот), соответствует положительный вклад в общую нейродоступность соединений, в противоположность амидной или сульфонамидной группам.

В целом, молекулярные фрагменты, содержащие электронодонорные или электроноакцепторные атомы, характеризуются отрицательными вкладами в из-

менение logBB (рис. 2), тогда как присутствие ароматических групп ассоциировано с повышенной нейродоступностью. Следует отметить, что в представленном ряду влияния фрагментов на изменение величины logBB преобладают асимметричные оценки, что может быть обусловлено как гетерогенностью самой выборки, так и полифункциональным характером включенных в ее состав соединений и возможными внутримолекулярными взаимодействиями между группами атомов. Это подтверждается и наличием экспериментальных “выбросов” — значительно отличающихся показателей вклада в общую нейродоступность.

Таким образом, с использованием метода симплексного представления структуры и QSPR-анализа оценен относительный вклад различных молекулярных фрагментов и функциональных групп в величину нейродоступности соединений, проникающих через ГЭБ путем простой диффузии. Установлено, что присутствие сильно полярных групп (карбоксовая, карбонильная, фенольный гидроксил) негативно сказывается на способности веществ поступать в ткань мозга, тогда как наличие атомов галогенов и ароматических структур преимущественно увеличивает эффективность переноса молекул через ГЭБ. Построенная QSAR модель (число деревьев 500, число случайно выбираемых переменных 1350) характеризуется удовлетворительной прогнозирующей способностью, оцененной согласно внутренней процедуре валидации моделей RandomForest ($R_{\text{ооб}}^2 = 0,54$, $\text{RMSE}_{\text{ооб}} = 0,47$).

Коэффициент детерминации ($R_{\text{ооб}}^2$) и среднеквадратичная ошибка прогноза ($\text{RMSE}_{\text{ооб}}$) рассчитывались по соответствующим формулам:

$$R_{\text{ооб}}^2 = 1 - \frac{\sum (y_{\text{pred},i} - y_{\text{exp},i})^2}{\sum (y_{\text{exp},i} - y_{\text{exp,mean}})^2},$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (y_{\text{pred},i} - y_{\text{exp},i})^2},$$

где y_{pred} — предсказанное моделью случайного леса значение logBB для out-of-bag выборки (аналог процедуры скользящего контроля), y_{exp} — экспериментально измеренное значение logBB, $y_{\text{exp,mean}}$ — среднее значение величины logBB для всех соединений в выборке, n — количество соединений в выборке (в данной задаче 325).

ке, n — количество соединений в выборке (в данной задаче 325).

ЛИТЕРАТУРА

1. О. А. Раевский, С. Л. Солодова, О. Е. Раевская и др., *Биомед. химия*, № 3, 246 – 256 (2012).
2. О. А. Раевский, С. Л. Солодова, О. Е. Раевская, Р. Манхольд, *Хим.-фарм. журн.*, **46**(3), 3 – 8 (2012); *Pharm. Chem. J.*, **46**(3), 133 – 138 (2012).
3. V. E. Kuz'min, A. G. Artemenko, P. G. Polishchuk, et al., *Mol. Modelling*, № 11, 457 – 467 (2005).
4. W. L. Jolly and W. B. Perry, *J. Am. Chem. Soc.*, **95**(17), 5442 – 5450 (1973).
5. R. Wang, Y. Fu and L. Lai, *J. Chem. Inform. Computer Sci.*, **37**(3), 615 – 621 (1997).
6. Б. В. Иоффе, *Рефрактометрические методы в химии*, Химия, Ленинград (1983).
7. L. Breiman, *Random Forests*, **45**(1), 5 – 32 (2001).
8. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2012), 1731.
9. N. A. Kovdienko, P. G. Polishchuk, N. E. Muratov, et al., *Molec. Inform.*, **29**(5), 394 – 406 (2010).
10. А. Г. Артеменко, В. Е. Кузьмин, Е. Н. Муратов и др., *Хим.-фарм. журн.*, **43**(8), 27 – 35 (2009); *Pharm. Chem. J.*, **43**(8), 454 – 462 (2009).
11. P. G. Polishchuk, V. E. Kuz'min, A. G. Artemenko, E. N. Muratov, *Molec. Inform.*, **32**(9 – 10), 843 – 853 (2013).
12. M. H. Abraham, A. Ibrahim, et al., *J. Pharmaceut. Sci.*, **95**(10), 2091 – 2100 (2006).
13. P. Garg, J. Verma, *J. Chem. Inform. Model.*, **46**(1), 289 – 297 (2005).
14. A. Guerra, J. A. Páez, N. E. Campillo, *QSAR Combinatorial Sci.*, **27**(5), 586 – 594 (2008).
15. J. Kelder, P. D. Grootenhuys, D. M. Bayada, et al., *Pharmaceut. Res.*, **16**(10), 1514 – 1519 (1999).
16. D. A. Konovalov, D. Coomans, E. Deconinck, Y. V. Heyden, *J. Chem. Inform. Modeling.*, **47**(4), 1648 – 1656 (2007).
17. S. R. Mente, F. Lombardo, *J. Computer-Aided Molec. Design*, **19**(7), 465 – 481 (2005).
18. R. Narayanan, S. B. Gunturi, *Bioorgan. Med. Chem.*, **13**(8), 3017 – 3028 (2005).
19. J. A. Platts, M. H. Abraham, Y. H. Zhao, et al., *Eur. J. Med. Chem.*, **36**(9), 719 – 730 (2001).
20. R. C. Young, R. C. Mitchell, T. H. Brown, et al., *J. Med. Chem.*, **39**, 656 – 671 (1988).
21. M. H. Abraham, H. S. Chadha, R. C. Mitchell, *J. Pharm. Sci.*, **83**, 1257 – 1268 (1994).
22. J. A. Gratton, M. H. Abraham, Bradbury, et al., *J. Pharm. Pharmacol.*, **49**, 1211 – 1216 (1997).

Поступила 09.04.14

RANKED SERIES OF MOLECULAR FRAGMENTS DEFINING NEUROAVAILABILITY OF DRUGS

P. G. Polishchuk, A. P. Kosinskaya, V. B. Larionov, L. N. Ognichenko, V. E. Kuz'min, and N. Ya. Golovenko

Bogatskiy Physico-Chemical Institute, National Academy of Sciences of Ukraine, 65080 Odessa, Ukraine

By using computer modeling methods based on the quantitative structure – property relationships (QSPRs), it is possible both to perform rational planning of experimental researches and carry out the effective generation of structures able to permeate through the blood-brain barrier (BBB). This work was aimed at (i) the creation of QSPR models based on the simplex representation of molecules for nonexperimental estimation of their ability to permeate the BBB and (ii) structural interpretation of the proposed models with allowance for the mutual influence of atoms in molecules. Based on the simplex representation of molecular structure, the contribution of various fragments and functional groups to the neuroavailability was estimated. In particular, it was found that the presence of highly polar groups (carboxyl, carbonyl, phenolic hydroxyl) has negative influence on the ability of molecule to enter the brain. The presence of halogen atoms and aromatic fragments positively influence the efficacy of transfer through the BBB. The formulated QSPR model possesses satisfactory prognostic ability that has been estimated using the RandomForest internal validation procedure ($R^2 = 0.54$, $\text{RMSE} = 0.47$).

Keywords: blood-brain barrier; diffusion; simplex; QSPR.