

К 40-летию Института физиологически активных веществ РАН**ОБЗОРЫ****ВКЛАД ВОДОРОДНОГО СВЯЗЫВАНИЯ В БИОДОСТУПНОСТЬ ЛЕКАРСТВ:
МЕТОДЫ ХЕМОИНФОРМАТИКИ****О.А. Раевский**Институт физиологически активных веществ Российской академии наук,
142432, Черноголовка Московской обл., Северный проезд, 1; эл. почта: raevsky@ipac.ac.ru

Обзор, основанный преимущественно на собственных публикациях, посвящён выявлению количественных связей “структура-биодоступность”. В первой части обзора описывается схема классификации дескрипторов водородных связей, создание оригинальных 2D термодинамических дескрипторов водородных связей, разработка компьютерной программы HYBOT, создание оригинальных трёхмерных потенциалов водородных связей и HYBOT PSA дескрипторов. Во второй части обзора представлены конкретные результаты использования вышеуказанных дескрипторов при создании QSAR моделей предсказания свойств, связанных с биодоступностью: липофильности, растворимости в воде и физиологических средах, абсорбции и проницаемости лекарств через гематоэнцефалический барьер.

Ключевые слова: QSAR; HYBOT; дескрипторы водородных связей; биодоступность**DOI:** 10.18097/BMCRM00060**ОБВЕДЕНИЕ**

Понятие водородной связи описывает взаимодействие атома водорода, ковалентно связанного с каким-либо электроотрицательным атомом (азот, кислород, сера, фтор, хлор), с другим электроотрицательным атомом этой же молекулы (внутримолекулярная водородная связь) или другой молекулы (межмолекулярная водородная связь). При этом указанный атом водорода определяется как “донор водородной связи”, а электроотрицательные атомы как “акцептор водородной связи”. Энергии водородных связей меняются в очень широких пределах (от 1 до 160 кДж/моль) и зависят от пространственной конформации образованного Н-комплекса, типов атомов, непосредственно участвующих в формировании водородной связи, и природы заместителей, расположенных в молекулах в непосредственной близости от доноров и акцепторов водородной связи [1].

Первая публикация, упоминающая водородную связь, вышла в 1912 году – Моор и Винмил использовали понятие о водородной связи, чтобы объяснить, почему гидроксид триметиламмония является слабым основанием по сравнению с гидроксидом тетраметиламмония [2]. В свою очередь, в 1920 году М.Л. Хаггинсом была высказана гипотеза образования “водородносвязанных стержней” [3] в воде, объясняющая аномальные свойства воды.

Физическая природа водородной связи с момента обнаружения и до настоящего времени остаётся предметом повышенного внимания. Однако выяснены далеко не все детали этого сложного взаимодействия атомов. Продолжаются дискуссии о характере взаимодействий при образовании водородной связи (электростатических, ковалентных или их сочетания), о пространственной ориентации атомов и неподелённых электронных пар гетероатомов,

участвующих в образовании Н-комплексов [1]. В то же время, очевидно, что взаимодействия между молекулами, являющимися донорами и акцепторами водородной связи, приводят к образованию разнообразных молекулярных и ионных комплексов, играющих важную роль в химических и биохимических процессах, включая ферментативный катализ [4].

Структуры, образуемые за счёт водородных связей, исключительно важны в биологических системах, поскольку они играют ключевую роль в макромолекулярных структурах и при молекулярном распознавании. Классический пример – формирование 3D структуры нуклеиновых кислот и определяющая роль водородного связывания в процессе репликации ДНК [5], основанные на образовании высокоспецифичных комплексов пуриновых и пиримидиновых оснований. Другой пример – формирование структуры белка, элементы вторичной структуры которого стабилизируются посредством водородных связей (энтальпийные вклады), в то время как третичная структура в большей степени определяется гидрофобными контактами (преимущественно энтропийные вклады) [6].

Данный обзор затрагивает лишь небольшую часть эффектов, связанных с образованием водородных связей, которые имеют прямое отношение к решению проблемы биодоступности лекарственных соединений. В основе понятия биодоступности лежит такой физиологический процесс, как абсорбция, сам по себе достаточно сложный и включающий в себя, например, перенос вещества через апикальную плазматическую мембрану внутрь клеток ЖКТ, внутриклеточный транспорт веществ, опционально метаболическую модификацию, перенос вещества из клетки в кровь или лимфу [7]. По определению FDA, биодоступность – скорость и степень поступления (всасывания) активного фармацевтического ингредиента из соответствующей



лекарственной формы и места введения в системный кровоток, в результате чего он становится доступным в зоне действия [8]. Среди большого числа факторов, влияющих на биодоступность, важную роль играют физические свойства лекарственного средства, в частности, гидрофобность (гидрофильность), степень диссоциации на ионы, растворимость. Для многих органических соединений определяющим фактором при изучении этой группы свойств является способность образовывать водородные связи.

Отдел компьютерного молекулярного дизайна (ОКМД) ИФАВ РАН уделяет первостепенное внимание развитию QSAR моделей с участием водородного связывания, обзору которых и посвящена настоящая работа.

1. ДЕСКРИПТОРЫ ВОДОРОДНЫХ СВЯЗЕЙ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ИХ РАСЧЁТА

1.1. Классификация дескрипторов водородных связей

В литературе описан целый ряд различных дескрипторов водородной связи, включая индикаторные переменные, числа донорных и акцепторных атомов водородной связи, поверхности атомов, участвующих в образовании водородной связи, квантово-химические дескрипторы, термодинамические и сольватохромные параметры [9]. Каждый из этих дескрипторов прямо или косвенно связан со сложным процессом формирования водородных связей на различных уровнях. Классификация дескрипторов водородной связи (рис. 1) по их информационному содержанию предложена нами в [10].

Индикаторные дескрипторы самого низкого уровня, косвенно характеризующие водородную связь, дают возможность определить лишь тенденцию в изменении свойства. Например, чем больше

донорных и акцепторных групп водородной связи, тем лучше растворимость соединений в воде (не бесспорное суждение для кристаллических соединений!). Дескрипторы промежуточного уровня (энтальпийные и свободноэнергетические факторы водородной связи) дают возможность количественно охарактеризовать одноцентровые водородные связи. Дескрипторы следующего уровня (потенциалы водородной связи, термодинамические параметры поверхностей водородной связи) дают количественное описание водородной связи при оптимальном расположении взаимодействующих химических групп. А расчёт термодинамических параметров с учётом зависимости потенциалов водородной связи от расстояний между атомами, образующих водородную связь, и от ориентации атомов и неподелённых электронных пар является самым высоким уровнем количественного описания водородной связи.

1.2. Создание оригинальных 2D термодинамических дескрипторов водородных связей

Разумной альтернативой любым качественным подходам оценки водородносвязанных параметров может быть только путь описания образования водородной связи на основе термодинамических функций комплексообразования согласно уравнению:

$$\Delta G = -RT \ln(K) = \Delta H - T\Delta S \quad (1)$$

где K – константа равновесия, ΔG – разность свободных энергий, ΔH – разность энтальпий, ΔS – разность энтропий, T – температура.

Для описания изменения энтальпии при образовании водородной связи между i -тым и j -тым атомами (ΔH_{ij}) нами было предложено эмпирическое уравнение, в котором использованы донорные (E_d) и акцепторные (E_a) энтальпийные факторы водородной связи, упорядоченные в рамках единой шкалы и имеющие противоположные знаки [11-13]:

$$\Delta H = k_1 E_a E_d \quad (2)$$

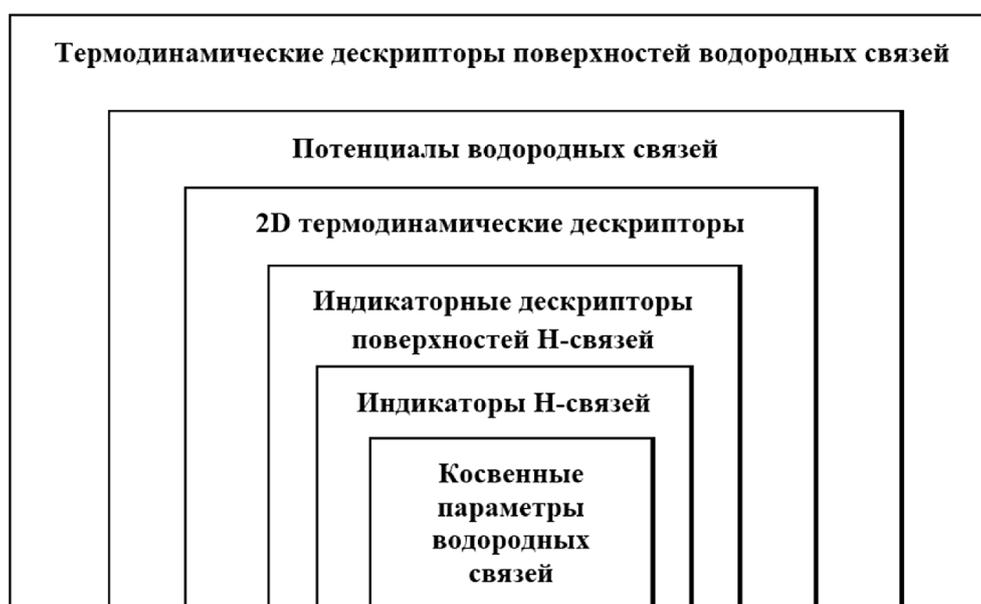


Рисунок 1. Иерархия информационного содержания различных дескрипторов водородной связи [10].

Примеры использования этого подхода для оценки акцепторной способности водородной связи для обширного массива разнообразных по структуре соединений представлены в публикациях [14-21].

Аналогичное уравнение было предложено и для описания изменения свободной энергии при образовании водородной связи [17]:

$$\Delta G = k_1 C_a C_d + k_0 \quad (3),$$

где C_d – донорный, а C_a – акцепторный факторы свободной энергии водородной связи.

Развитие и успешное применение рассматриваемого подхода количественного описания водородной связи возможно лишь на базе обширных экспериментальных данных по термодинамике водородной связи. Использование таких данных в работе [22] позволило получить статистически устойчивые параметры. Из многих тысяч комплексов с водородной связью нами было отобрано 936, удовлетворяющих следующим требованиям: (1) стехиометрия комплексов водородной связи 1:1; (2) растворитель во всех системах CCl_4 ; (3) для комплексов имеются данные как по ΔH , так и по ΔG ; (4) значения ΔH и ΔG определены экспериментальными методами.

Был разработан специальный алгоритм для расчёта энтальпийных и свободной энергии параметров водородной связи с фиксацией гексаметилфосфорида в качестве референсного акцептора (встречался 88 раз среди изученных 936 комплексов) и фенола в качестве референсного донора водородной связи (встречался 119 раз). В результате были рассчитаны донорные факторы водородной связи для 163 соединений и акцепторные факторы водородной связи для 195 соединений. Оценку качества полученных значений факторов водородной связи проводили, сравнивая экспериментальные значения ΔH и ΔG с рассчитанными с использованием уравнений 2 и 3:

$$\Delta H_{\text{calc}} = -0.27(\pm 0.45) + 1.00(\pm 0.02)\Delta H_{\text{exp}} \quad (4),$$

$$n = 936, R = 0.954, s = 2.70, F = 9553;$$

$$\Delta G_{\text{calc}} = -0.07(\pm 0.12) + 1.00(\pm 0.01)\Delta G_{\text{exp}} \quad (5),$$

$$n = 936, R = 0.984, s = 1.11, F = 28556.$$

Приведённые статистические параметры уравнений (4) и (5) показывают, что полученные факторы водородной связи (рис. 2) обеспечивают возможность достаточно точного расчёта энтальпии и свободной энергии водородного связывания в комплексе.

Расширение спектра органических соединений, для которых можно рассчитать свободной энергии факторы, стало возможным за счёт использования экспериментальных данных по термодинамике водородной связи при условии, что известно только ΔG комплекса [23, 24]. Число факторов таким образом было увеличено до 414 для доноров и 1298 для акцепторов водородной связи, что потенциально даёт возможность рассчитать свободную энергию 537372 комплексов с водородной связью. Особенно стоит отметить, что предложенная шкала адекватно описывает влияние заместителей при группах, образующих водородные связи, на значение факторов водородной связи [22]. Этот широко известный факт нередко игнорируется при использовании простых методов учёта водородного связывания. В то же время, подходы, аналогичные предложенному нами, также развивались и другими исследователями [25-27].

1.3. Компьютерная программа НУВОТ

Собранная база экспериментальных данных по термодинамике водородной связи и предложенная шкала факторов водородного связывания стали основой для создания в 90-х годах программного комплекса НУВОТ (Hydrogen Bond Thermodynamics) [28]. Помимо собственно базы данных термодинамических

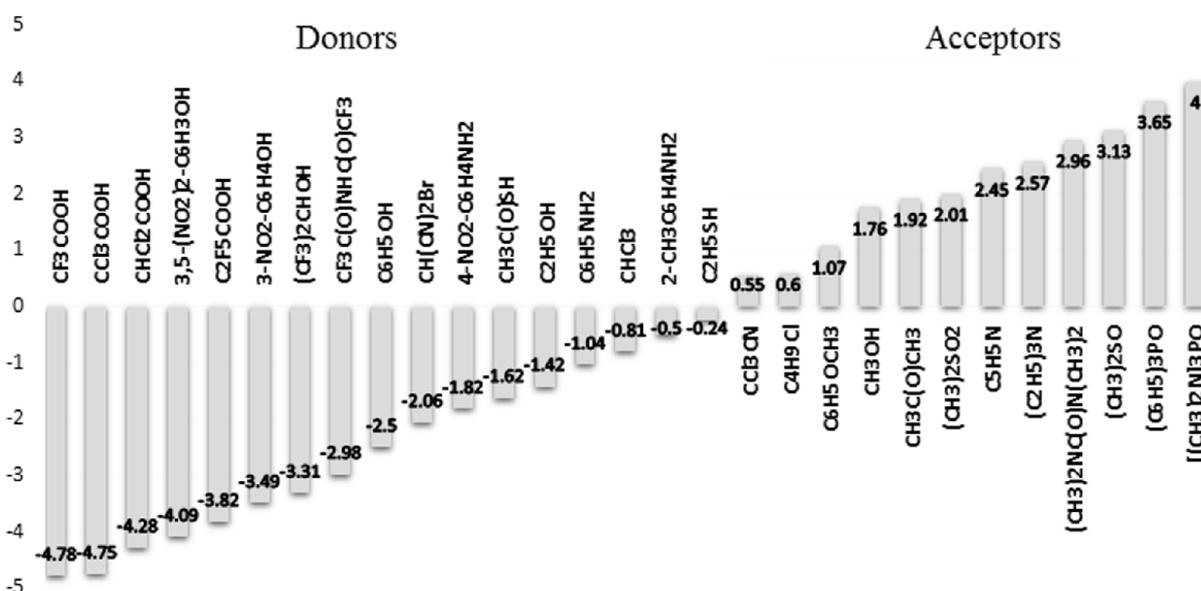


Рисунок 2. Унифицированная шкала донорных и акцепторных факторов водородной связи (пример для простых органических соединений).

параметров ~16000 комплексов с водородной связью программа включает в себя базы данных энтальпийных и свободноеэнергетических факторов водородной связи (более 50000), модуль расчёта фактора водородного связывания для произвольной органической структур методом ближайшего соседа, а также расчёт дополнительных 32 молекулярных дескрипторов, включая молекулярную поляризуемость, суммы положительных и отрицательных зарядов на атомах, суммы энтальпийные и свободноеэнергетические факторы водородной связи. Программа функционирует как самостоятельно в операционных системах Windows и LINUX, так и в составе системы управления химическими базами данных CHED [29]. Примеры использования программного комплекса HYBOT для установления количественных взаимосвязей “структура-свойство” будут представлены в следующей части этого обзора.

1.4. Трёхмерные потенциалы водородной связи и программа HYBOT 3D

Расчёт описанных выше факторов водородного связывания базируется исключительно на структурной формуле органического соединения; таким образом, все рассчитанные с их использованием дескрипторы можно отнести к 2D-дескрипторам. В то же время современная тенденция в молекулярном моделировании и QSAR – использование в работе данных по 3D системам [30]. В той или иной степени учёт водородного связывания есть в большинстве программ молекулярной механики. От наиболее простых, таких как поправочные коэффициенты при расчёте ван-дер-ваальсовых взаимодействий (например, в поле сил TRIPOS [31]), до самостоятельных потенциалов, таких как потенциал “10-12” в старых версиях программы AMBER [32] или “6-8” в программе GRID [33]. Недостаток таких потенциалов, как правило, – ограниченное число фиксированных параметров пары донор/акцептор. Например, в программе GRID фактически зафиксированы три потенциала водородной связи: -4,00 ккал/моль в качестве оптимальной энергии водородной связи для пары

O–H...O, -2,8 ккал/моль для O–H...N и -2 ккал/моль для N–H...N. Эти значения потенциалов не учитывают того факта, что энтальпия водородной связи для одного и того же типа атомов меняется в очень широких интервалах и существенно зависит от природы заместителя. Так, по данным расчёта программой HYBOT для пары O–H...O оптимальная энтальпия водородной связи (при идеальной геометрии) меняется в интервале 0,9-15,9 ккал/моль, для O–H...N – в интервале 0,5-10,5 ккал/моль, для N–H...N – в пределах 0,5-11,5 ккал/моль. Кроме того, при анализе набора низкомолекулярных кристаллических комплексов из Кембриджского банка структурных данных [34] с близкими к “идеальным” значениям углов в геометрии водородной связи (угол донорный – “тяжелый” атом – атом водорода – атом акцептора в интервале 173-187, угол между ядрами акцептора неподеленной пары электронов и атомом водорода в интервале 170-190), было показано [39], что от природы заместителей зависит не только оптимальная энергия водородной связи (E_m), но и оптимальная длина (r_m) водородной связи (рис. 3). Зависимость между ними аппроксимируется уравнением:

$$r_m = k_7 \log \left(\frac{k_4 - E_m}{E_m} \right) + k_8 \quad (6)$$

В результате была предложена модификация потенциала “6-8”, реализованная в программном комплексе HYBOT 3D [35], где расчёт энергии водородного связывания проводится по следующей схеме:

- 1) рассчитываются энтальпийные акцепторные и донорные факторы водородной связи (идентично с модулем расчёта HYBOT);
- 2) рассчитываются значения оптимального расстояния и оптимальной энергии для пары атомов, образующих водородную связь, с помощью уравнений 6 и 3;
- 3) рассчитывается значение потенциала “6-8” с учётом отклонений от идеального расположения атомов, участвующих в образовании водородной связи:

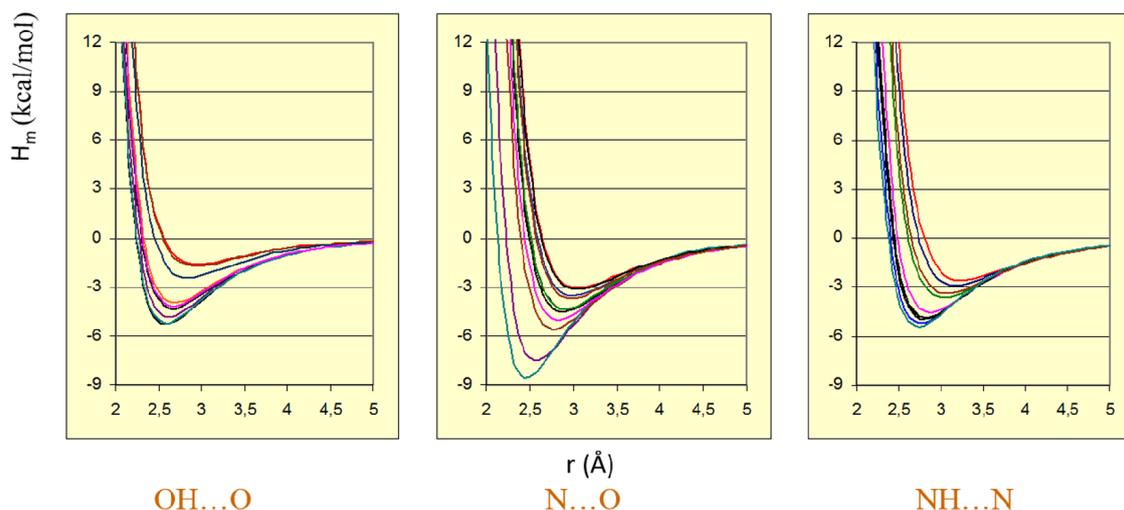


Рисунок 3. Потенциалы водородной связи. Зависимость оптимальной энергии от длины связи.

$$E_r = \frac{-3E_m r_m^8}{r^8} - \frac{-4E_m r_m^6}{r^6} \quad (7),$$

где E_m – оптимальная свободная энергия (ккал/моль), r – расстояние между донором и акцептором водорода (r_m – оптимальное, Å).

Значение оптимальной свободной энергии рассчитывается по уравнению (2) при k_1 равным 1.33, а оптимальное расстояние между донорным и акцепторным атомами рассчитывается по уравнению (6). Коэффициенты k_4 , k_7 и k_8 рассчитаны для каждой пары акцепторный атом/донорный атом по данным из кристаллографической базы данных низкомолекулярных соединений [34].

Сопоставление полученных расчётных данных с результатами работы [35-37] показало, что данный метод более корректно описывает сложные межмолекулярные взаимодействия по сравнению с программой GRID.

1.5. HYBOT PSA дескрипторы

С использованием факторов водородной связи можно рассчитать также 6 новых дескрипторов, идеологически близких такому широко используемому дескриптору как PSA (polar surface area). Они дают возможность определить области молекулярной поверхности, благоприятные для образования межмолекулярных водородных связей и также учесть влияние заместителей при донорных и акцепторных центрах на силу этих связей.

HYBOT PSA дескрипторы можно рассчитать по следующим формулам:

$$PSA_{EA} = \sum_n k_a E_a \quad (8),$$

где PSA_{EA} – вандерваальсова акцепторная поверхность (в Å²), пропорциональная факторам E_a , $k_a = (1/5)(1/3S_O)$, S_O – поверхность сферы с радиусом 1.36 Å (O_{sp3}).

$$PSA_{CA} = \sum_n k_a C_a \quad (9),$$

где PSA_{CA} – вандерваальсова акцепторная поверхность (в Å²), пропорциональная факторам C_a , $k_a = (1/5)(1/3S_O)$, S_O – поверхность сферы с радиусом 1.36 Å (O_{sp3}).

$$PSA_{ED} = \sum_n k_d E_d \quad (10),$$

где PSA_{ED} – вандерваальсова донорная поверхность (в Å²), пропорциональная факторам E_d , $k_d = (1/5)(1/3S_H)$, S_H – поверхность сферы с радиусом 1.08 Å (атом водорода).

$$PSA_{CD} = \sum_n k_d C_d \quad (11),$$

где PSA_{CD} – вандерваальсова донорная поверхность (в Å²), пропорциональная факторам C_d , $k_d = (1/5)(1/3S_H)$, S_H – поверхность сферы с радиусом 1.08 Å (атом водорода).

$$PSA_E = PSA_{ED} + PSA_{EA} \quad (12),$$

где PSA_E – вандерваальсова донорная и акцепторная поверхности (в Å²), пропорциональные сумме факторов E_d и E_a .

$$PSA_C = PSA_{CD} + PSA_{CA} \quad (13),$$

где PSA_C – вандерваальсова донорная и акцепторная поверхности (в Å²), пропорциональные сумме факторов C_d и C_a .

Расчёт HYBOT PSA дескрипторов реализован в программе MOLTRA II [36]. Программа осуществляет также непосредственный расчёт термодинамических параметров водородной связи при "идеальном" расположении донорных и акцепторных атомов и оптимальном расстоянии между этими атомами.

2. ИССЛЕДОВАНИЕ ВЗАИМОСВЯЗЕЙ ДЕСКРИПТОРОВ ВОДОРОДНЫХ СВЯЗЕЙ С БИОДОСТУПНОСТЬЮ ЛЕКАРСТВ

Так как фактически биодоступность лекарств является оценочной характеристикой для работы сложнейшей системы физических, физиологических и отдельных биохимических процессов, часто каскадно организованных, имеет смысл рассматривать каждый элемент по отдельности. К таким элементам можно отнести растворение в физиологической среде, мембранный транспорт, абсорбцию, различные метаболические процессы, распределение по различным органам. Например, для лекарств, предназначенных для воздействия на центральную нервную систему, очень важно преодоление гематоэнцефалического барьера (ГЭБ). В свою очередь, для оценки способности соединения проходить ГЭБ или участвовать в процессе мембранного транспорта большую роль играет такой фактор как липофильность органического соединения. Последний также важен в понимании такого явления как общая токсичность. Для создания предсказательных QSAR моделей, описывающих некоторые из этих явлений, использование понятия водородного связывания имеет принципиальное значение.

2.1. Липофильность органических соединений

Физико-химические модели липофильности, основанные на стерических, электростатических и водородносвязанных взаимодействиях, рассматриваются с начала 1980-х годов [37, 38]. В 1995 году нами было предложено уравнение, в котором в качестве дескрипторов используются молекулярный объём и водородносвязанные факторы для описания коэффициентов распределения химических соединений в системе октанол-вода ($\log P$) [42]. Позднее молекулярный объём был заменён на молекулярную поляризуемость (α). В результате было получено корреляционное уравнение, описывающее с высокой точностью зависимость $\log P$ от двух дескрипторов [39] для 2850 простых монофункциональных органических соединений:

$$\log P = 0.267\alpha - 1.00 \sum C_a \quad (14),$$

$n = 2850$, $R = 0.970$, $S = 0.23$.

Это уравнение было в дальнейшем использовано нами для расчёта липофильности программой SLIPPER-98 [40]. Несмотря на то, что значительная

часть соединений обучающей выборки содержала группы доноров водорода, вклад донорных факторов в липофильность не был показан. В то же время уравнение не содержит постоянного члена, то есть для соединений, не содержащих доноров и акцепторов водородной связи, липофильность прямо пропорциональна молекулярной поляризуемости, которая, в свою очередь, хорошо коррелирует с такими параметрами как молекулярный вес или объём.

Хотя применение уравнения (14) в программе SLIPPER-98 обеспечило хорошую точность оценок $\log P$, для расчётных значений этого параметра в случае сложных химических соединений и лекарств, содержащих в составе молекул несколько функциональных групп, наблюдались существенные отклонения по сравнению с экспериментальными значениями. Это побудило нас в работе [41] модифицировать формулу (14), включив в уравнение экспериментальное значение липофильности для ближайшего структурно-родственного соединения и разности величин дескрипторов для него и тестируемого соединения:

$$\log P_i = \log P_{nm} + 0.267(\alpha_i - \alpha_{nm}) - 1.00(\sum C_{a(i)} - \sum C_{a(nm)}) \quad (15).$$

В дальнейшем формула (15) была применена для расчёта $\log P$ в программном комплексе SLIPPER-2001 [41]. Эффективность расчёта липофильности с помощью уравнения (15) была проверена на массиве, содержащем 10937 органических соединений и лекарств. Коэффициент корреляции между экспериментальными и рассчитанными значениями для этих соединений составил 0.972.

2.2. Растворимость в физиологической среде

Растворимость химического соединения в воде является важным параметром при оценке биодоступности. Плохая растворимость кандидатов – частая причина неудачи проектов по поиску новых лекарственных соединений [42]. Что касается ЦНС, новое соединение вряд ли станет лекарством для лечения её заболеваний, если его растворимость меньше 1 мкм/л [43].

Известно, что растворимость жидких органических соединений хорошо коррелирует с коэффициентом распределения в системе октанол-вода [44]:

$$\log S_{\text{exp}} = 0.98 - 1.34 \log P_{\text{exp}} \quad (16),$$

$$n = 150, R^2 = 0.874, SD = 0.47.$$

Полученная нами корреляция, основанная на данных по растворимости 606 жидких неионизированных органических соединений, оказалась похожей [16]:

$$\log S = 0.41 - 1.02 \log P \quad (17),$$

$$n = 606, R^2 = 0.799, RMSE = 0.89.$$

Поскольку ранее нами уже была установлена (см. выше) зависимость для $\log P$ (уравнение 14), логично было использовать тот же набор дескрипторов. Однако в данном случае значимыми были как молекулярная поляризуемость (α) и сумма акцепторных факторов ($\sum C_a$), так и два

дополнительных дескриптора (сумма частичных отрицательных зарядов ($\sum Q^-$) и сумма донорных факторов ($\sum C_d$) [45]):

$$\log S = 0.434 - 0.298\alpha + 0.487Q^- - 1.09\sum C_a - 0.34\sum C_d \quad (18),$$

$$n = 630, R^2 = 0.897, RMSE = 0.65.$$

В то же время, ситуация с растворимостью в воде кристаллических соединений (чаще встречающихся среди лекарственных соединений) является гораздо более сложной, так как энергия, необходимая для их растворения, зависит от структуры кристаллической решётки. В настоящее время имеется большое число работ [16, 46-51], посвященных предсказанию растворимости таких веществ в воде. Однако, как правило, ошибка предсказания в этих работах от единицы и более в логарифмической шкале $\log S$ [52, 53].

Для решения данной проблемы мы использовали две оригинальные QSAR модели: AMP (средняя арифметическая модель) [54] и LoReP (регрессия локального кластера) [55]. Результаты оценки растворимости 2615 кристаллических соединений представлены в таблице 1. В обоих случаях получены весьма удовлетворительные статистические критерии QSAR моделей с участием параметра $A\log P$ (коэффициент распределения октанол/вода рассчитанный методом Ghose-Crippen [56]) и молекулярной поляризуемости.

Стоит также упомянуть, что в физиологических условиях для лекарственных соединений важное значение имеет pH среды. В работах [57-59] этот вопрос рассматривается более подробно.

2.3. Абсорбция химических соединений и лекарств в живых организмах

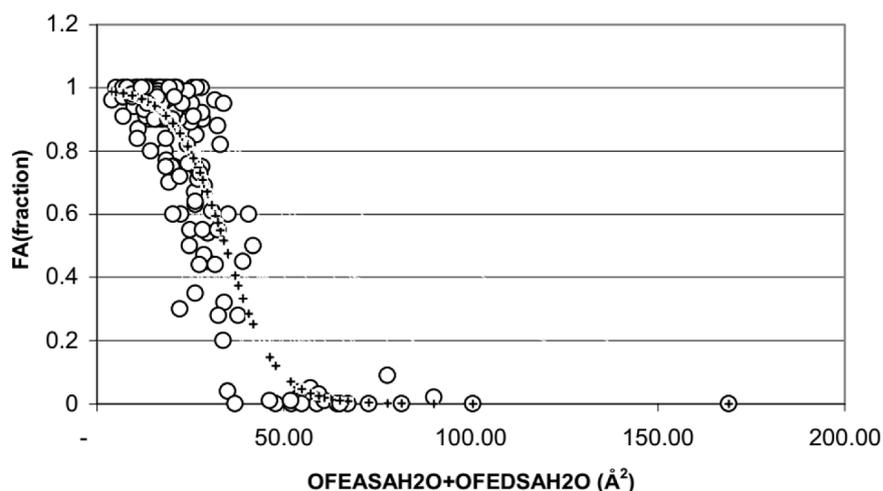
Первое подробное исследование в области предсказания абсорбции нам удалось выполнить на примере модели пассивного транспорта 32 лекарств при пероральном способе введения в организм человека [60].

Величину FA (Fraction Absorption) измеряли в пределах от нуля до единицы как отношение абсорбированного вещества к его общему количеству. Линейные регрессионные модели взаимосвязи $\logit(FA)$ с использованными физико-химическими дескрипторами показали неплохие статистические критерии, но не выдержали тестирования при скользящем контроле. Далее в работе были рассмотрены эмпирические нелинейные модели. Например, уравнение (19) имеет сигмоидную зависимость (рис. 4) величины FA от суммы таких дескрипторов, как молекулярная поляризуемость, сумма частичных зарядов и суммы донорных и акцепторных факторов. Статистические критерии данного набора моделей (табл. 2) хорошие, причём стандартное отклонение рассчитанных значений FA находятся на уровне ошибки их экспериментального определения:

$$FA = \frac{1}{1 - 10^{-[k_0 + k_1\alpha + k_2\sum Q^- + k_3\sum C_{ad} + k_4\sum C_a + k_5\sum C_d]}} \quad (19).$$

Таблица 1. Одиночные и консенсусные AMP и LoReP модели растворимости в воде 2615 кристаллических соединений

Method	T _c	Descriptors	N	R ²	SD	a ₀ ±err ₀	a ₁ ±err ₁
AMP	0	AlogP	2609	0.823	0.96	0.09±0.04	1.03±0.01
		α	2609	0.805	1.00	0.05±0.04	1.03±0.01
		α, AlogP	2613	0.825	0.95	0.11±0.04	1.04±0.01
	0.3	α	1886	0.859	0.85	0.09±0.04	1.02±0.01
		AlogP	1914	0.849	0.89	0.10±0.04	1.02±0.01
		α, AlogP	2069	0.854	0.87	0.13±0.04	1.03±0.01
	0.5	AlogP	907	0.922	0.71	0.19±0.05	1.03±0.01
		α	862	0.921	0.72	0.11±0.05	1.02±0.01
		α, AlogP	1020	0.917	0.72	0.21±0.05	1.03±0.01
LoReP	0	AlogP	2609	0.842	0.90	-0.03±0.04	1.00±0.01
		α	2609	0.827	0.94	-0.07±0.04	1.00±0.01
		α, AlogP	2613	0.855	0.86	0.02±0.03	1.02±0.01
	0.3	α	1886	0.878	0.79	0.00±0.04	1.00±0.01
		AlogP	1914	0.871	0.82	0.02±0.04	1.01±0.01
		α, AlogP	2069	0.867	0.83	-0.05±0.04	1.00±0.01
	0.5	α, AlogP	862	0.935	0.65	0.04±0.05	1.01±0.01
		AlogP	907	0.935	0.65	0.10±0.04	1.02±0.01
		α, AlogP	1020	0.929	0.67	0.03±0.04	1.01±0.01

**Рисунок 4.** Графические зависимости фракций абсорбции (FA, Fraction Absorption) от поверхностных HYBOT дескрипторов водородных связей (OFEASA+OFEDSA).

Позднее сигмоидный характер зависимости пассивной абсорбции лекарств от размера поверхностей дескрипторов водородной связи был нами подтвержден на обширной выборке из 154 лекарств [61].

В работе [62] нами были использованы данные по абсорбции в организме человека 100 разнообразных по структуре лекарств, подверженных не только пассивной диффузии, но и активному транспорту. Учёт этого явления меняет картину зависимости абсорбции от потенциала водородного связывания. При этом предсказание возможно, если данные можно разбить на несколько выборок родственных соединений.

Таким образом, полученные корреляционные зависимости свидетельствуют о важной роли водородного связывания в явлении абсорбции.

По мере усложнения структуры соединений и изучаемых биологических эффектов характер зависимости структура-свойство (активность) может быть все более сложным. В данном случае уже недостаточно использования только линейных моделей, а требуется рассмотреть различные варианты, такие как параболические, билинейные, сигмоидные и другие подобные модели.

2.4. Проницаемость через гематоэнцефалический барьер

Гематоэнцефалический барьер (ГЭБ) – один из важнейших защитных механизмов ЦНС, объединяющий совокупность анатомических и функциональных систем, которые поддерживают постоянство и регулируют состав внутренней среды

Таблица 2. Коэффициенты уравнения (19) и их статистические критерии [60]

Eq.	Drugs (no.)	Regression coefficients										Statistical criteria				
		k_0	k_1	k_{2a}	k_{2b}	k_3	k_4	k_5	n	R	s	Q				
4.1	1-19	5.510 ± 3.211				-0.339 ± 0.196						19	0.925	0.136	0.873	
4.2	1-19	5.980 ± 2.303									-0.428 ± 0.162	0.320 ± 0.134	19	0.973	0.088	0.941
4.3	1-19	6.962 ± 3.896	-0.015 ± 0.058								-0.449 ± 0.174	0.357 ± 0.176	19	0.975	0.090	0.883
4.4	1-19	7.560 ± 2.945		-0.341 ± 0.310							-0.431 ± 0.179	0.308 ± 0.151	19	0.98	0.084	0.899
4.5	1-19	3.608 ± 1.542		1.476 ± 0.631	-0.186 ± 0.079						-0.462 ± 0.125	0.331 ± 0.104	19	0.992	0.053	0.775
4.6	1-19	-0.426 ± 2.037	0.095 ± 0.057	1.390 ± 0.500	-0.210 ± 0.069						-0.419 ± 0.095	0.162 ± 0.080	19	0.995	0.042	0.971
4.7	14-32	4.194 ± 2.288				-0.264 ± 0.144							19	0.913	0.163	0.867
4.8	14-32	3.592 ± 1.761									-0.281 ± 0.128	0.171 ± 0.113	19	0.948	0.131	0.918
4.9	14-32	4.275 ± 3.143	-0.009 ± 0.041								-0.305 ± 0.159	0.198 ± 0.149	19	0.949	0.135	0.808
4.10	14-32	3.409 ± 1.841		0.055 ± 0.236							-0.286 ± 0.135	0.177 ± 0.119	19	0.948	0.134	0.855
4.11	14-32	8.106 ± 7.719	-0.066 ± 0.089	0.215 ± 0.346							-0.513 ± 0.460	0.380 ± 0.331	19	0.958	0.134	0.645
4.12	14-28, 30-32	5.238 ± 2.656				-0.328 ± 0.165							18	0.942	0.137	0.900
4.13	14-28, 30-32	4.430 ± 1.660									-0.337 ± 0.120	0.218 ± 0.100	18	0.977	0.091	0.944
4.14	14-28, 30-32	6.824 ± 2.983	-0.033 ± 0.029								-0.420 ± 0.160	0.305 ± 0.130	18	0.985	0.077	0.866
4.15	14-28, 30-32	4.349 ± 1.901		0.016 ± 0.187							-0.337 ± 0.124	0.217 ± 0.103	18	0.977	0.094	0.863
4.16	14-28, 30-32	3.319 ± 3.958		0.593 ± 2.009	-0.063 ± 0.219						-0.351 ± 0.138	0.234 ± 0.121	18	0.977	0.096	0.855
4.17	14-28, 30-32	6.236 ± 3.272	-0.109 ± 0.047	2.917 ± 1.729	-0.282 ± 0.180						-0.707 ± 0.233	0.558 ± 0.181	18	0.997	0.045	0.930
4.18	1-32	4.371 ± 1.699				-0.268 ± 0.107							32	0.922	0.140	0.895
4.19	1-32	4.475 ± 1.489									-0.323 ± 0.107	0.232 ± 0.091	32	0.952	0.113	0.934
4.20	1-32	2.552 ± 1.424	0.034 ± 0.032								-0.289 ± 0.102	0.162 ± 0.079	32	0.952	0.114	0.893
4.21	1-32	4.313 ± 1.534		0.042 ± 0.166							-0.326 ± 0.114	0.233 ± 0.094	32	0.952	0.115	0.921
4.22	1-32	3.399 ± 2.004		0.680 ± 0.899	-0.077 ± 0.108						-0.350 ± 0.122	0.248 ± 0.100	32	0.956	0.113	0.910
4.23	1-32	5.570 ± 3.392	-0.041 ± 0.051	1.143 ± 1.063	-0.120 ± 0.121						-0.474 ± 0.200	0.386 ± 0.191	32	0.962	0.111	0.795
4.24	1-28, 30-32	5.019 ± 1.809				-0.307 ± 0.112							31	0.943	0.122	0.918
4.25	1-28, 30-32	5.047 ± 1.335									-0.362 ± 0.096	0.263 ± 0.080	31	0.974	0.085	0.959
4.26	1-28, 30-32	6.682 ± 2.408	-0.022 ± 0.027								-0.412 ± 0.116	0.328 ± 0.113	31	0.977	0.083	0.955
4.27	1-28, 30-32	4.945 ± 1.395		0.025 ± 0.138							-0.364 ± 0.100	0.264 ± 0.082	31	0.974	0.087	0.938
4.28	1-28, 30-32	3.694 ± 1.524		0.829 ± 0.676	-0.096 ± 0.081						-0.389 ± 0.101	0.281 ± 0.083	31	0.979	0.080	0.932
4.29	1-28, 30-32	5.838 ± 2.204	-0.045 ± 0.033	1.207 ± 0.686	-0.126 ± 0.078						-0.493 ± 0.132	0.404 ± 0.125	31	0.985	0.070	0.906

головного и спинного мозга [63, 64]. ГЭБ регулирует распределение различных веществ между током крови и мозгом, и обладает высокой избирательностью. Барьерная функция ГЭБ – основная проблема при оценке будущей биодоступности лекарств, действующих на ЦНС.

В настоящее время общепринято, что физико-химические свойства лекарств, такие как размер молекул, заряды, способность к образованию водородных связей, липофильность, являются важными составляющими факторами, влияющими на преодоление ГЭБ [65]. Существует целый набор правил, созданных медицинскими химиками, для оценки возможности проницаемости соединений через ГЭБ. Здесь следует упомянуть о ряде публикаций: об оптимальном для проницаемости $\log P=2.0$ [66]; о том, что проницаемость кровь-мозг ($\log BB$) увеличивается за счёт $\log P$ и уменьшается при увеличении способности образовывать водородные связи [67]; об использовании для предсказания $\log BB$ величины $\log P$ (алкан-вода) и молекулярного объёма или PSA [68]; “правило-5” [69]; наблюдении, что соединения с $O+N \leq 5$ или $clogP > 0$ имеют большие шансы проникать в ЦНС [70]; о возможности избежать

эффлюкс Р-гликопротеином для соединений с $MW < 400$, $N+O < 4$, $pK_a < 8$ [71], и многих других [72-78]. Наши первые работы, посвященные решению этой проблемы, были опубликованы в 1990 году [79, 80]. Затем в совместных работах с зарубежными учёными [81-84] была показана эффективность дескрипторов водородного связывания, рассчитанных программой HUBOT.

В 2015 году Rankovic опубликовал обширный обзор [85], посвящённый важности физико-химических свойств для предсказания проницаемости через ГЭБ с обсуждением различных правил медицинских химиков, упомянутых выше. В обзоре содержатся данные по шести физико-химическим свойствам и концентрации 616 соединений в мозге мышей после введения их в кровь. С использованием этой выборки (дополнив её соединениями, не проходящими ГЭБ) мы провели специальное QSAR исследование возможности уже упомянутых правил [86, 87] и современных методов бинарной классификации для оценки проходимости (“ЦНС”; 500 – обучающая выборка, 50 – тестовая) или не проходимости (“не ЦНС”; также 500 и 50) химическим соединением ГЭБ [88]. Как видно из таблицы 3, ни один подход медицинских химиков не даёт удовлетворительных

Таблица 3. Протокол ЦНС/не ЦНС классификации соединений интуитивными подходами и статистическими методами. TP - правильное распознавание ЦНС, FN - неправильное распознавание ЦНС, TN - правильное распознавание не ЦНС, FP - неправильное распознавание не ЦНС, SE - чувствительность (TP/TP+EN), SP - (TN/TN++EP), ACC - точность (TP+TN)/(TP+EN+TN+EP)

Approach	Descriptors	Sets	TP	FN	TN	FP	SE	SP	ACC
“Rule of 5” 1997 [70]	MW>500, MlogP>4.15, HBD>5, HBA>10	Training	397	103	198	302	0.794	0.396	0.595
		External	47	3	11	39	0.940	0.220	0.580
Waterbeemd [81]	MW<450, PSA<90 Å ² , logD 1-4	Training	219	281	416	84	0.438	0.832	0.635
		External	30	29	43	7	0.600	0.860	0.730
Norinder [71]	N+O≤5	Training	496	4	59	441	0.992	0.118	0.555
		External	50	0	5	45	1.000	0.100	0.550
Norinder [71]	ClogP-(N+O)>0	Training	345	155	304	196	0.690	0.608	0.649
		External	36	14	34	16	0.720	0.680	0.700
Raub [74]	clogP<4, TPSA 40-80 Å ²	Training	178	322	394	106	0.356	0.644	0.572
		External	23	27	37	13	0.460	0.740	0.600
Hitchcock [78]	PSA<90 Å ² , HBD<3, logP 2-5, logD 2-5, W<500	Training	69	431	432	48	0.138	0.864	0.501
		External	15	35	46	4	0.300	0.820	0.610
Wager [75]	MPO score ≥4.0 /<4.0	Training	342	158	256	244	0.684	0.512	0.598
		External	34	16	18	32	0.680	0.360	0.520
LR	maxQ ⁻ , maxC _a maxC _a *maxC _d	Training	370	130	339	161	0.740	0.678	0.709
		CV	369	131	333	167	0.738	0.666	0.702
		external	42	8	39	11	0.840	0.780	0.810
LR	maxQ ⁻ , maxC _a maxC _a *maxC _d MW, HBD, NCC, logD	Training	382	118	340	160	0.764	0.680	0.722
		CV	384	116	336	164	0.768	0.672	0.720
		External	43	7	38	12	0.860	0.760	0.810
RF	*	Training	500	0	500	0	1.000	1.000	1.000
		CV (out-of-bag)	412	88	378	122	0.824	0.756	0.790
		External	47	3	35	15	0.940	0.700	0.820
8SVM	**	Training	448	52	403	97	0.896	0.806	0.851
		External	46	4	38	12	0.920	0.760	0.840

Примечание. * – $\alpha \max Q^+ \max Q^- \sum Q^+ \sum Q^- / \alpha \max E_a \max C_a \max E_d \sum C_a / \alpha \sum C_d / \alpha$ NRB NCC logDD pK_a PSA_{ed} TPSA(N,O) MlogP; ** – $\max Q^+ \max Q^- \sum Q^+ \sum Q^- / \alpha \max E_a \max C_a \max C_d \max E_a * \max E_d \max C_a * \max C_d \sum E_d \sum E_{ad} \sum E_a / \alpha \sum E_d / \alpha \sum C_a / \alpha \sum C_{ad} / \alpha$ MW HBD NRB NCC logPP logDD pK_a PSA_{ed} TPSA(NO) MlogP AlogP

результатов прогноза. Использование методов логистической регрессии (LR), случайного леса (RF) и метода опорных векторов (SVM) позволило создать стабильные QSAR модели с точностью (ACC) 0.81-0.84. В этих моделях дескрипторы водородных связей (прежде всего дескрипторы HYBOT) вносят существенный вклад.

ЗАКЛЮЧЕНИЕ

Водородные связи играют исключительно важную роль в формировании биохимических комплексов и в различных процессах с их участием, в том числе и в процессах, определяющих биодоступность лекарственных соединений. К настоящему времени компьютерное моделирование взаимосвязей между структурой молекул и их комплексов с такими процессами (растворением в физиологической среде, липофильностью, абсорбцией, проницаемостью в ЦНС, распределением по органам живого организма) является неременной составляющей дизайна новых эффективных лекарств. Использование дескрипторов водородной связи позволяет получать адекватные предсказательные модели, особенно, если используются те дескрипторы, которые учитывают природу ближайших заместителей при атомах, обладающих донорной и акцепторной способностью для образования водородной связи, такие как дескрипторы HYBOT.

БЛАГОДАРНОСТИ

Работа выполнена в рамках государственного задания на 2018 год (тема № 0090-2017-0020).

ЛИТЕРАТУРА

- Pimentel, G.C., & McClellan, A.L. (1960) The hydrogen bond. Freeman: San Francisco, CA.
- Moore, T.S., & Winmill, T.F. (1912) CLXXVII.—The state of amines in aqueous solution. *Journal of the Chemical Society, Transactions*, 101, 1635-1676. DOI: 10.1039/CT9120101635
- Latimer, W.M., & Rodebush, W.H. (1920) Polarity and ionization from the standpoint of the Lewis theory of valence. *Journal of the American Chemical Society*, 42(7), 1419-1433. DOI: 10.1021/ja01452a015
- Varfolomeev, S.D., & Pozhitkov, A.E. (2000) Active centers of hydrolases: the main types of structures and the mechanism of catalysis. *Vestnik Moskovskogo universiteta [Bulletin of Moscow University]*. Part 2. Chemistry, 41(3), 147-156.
- Pauling, L., & Pauling, P. (1978) *Chemistry* [Russian translation]. Mir, Moscow, 299.
- Kubinyi H. (2001) Hydrogen bonding, the last mystery in drug design? In: *Pharmacokinetic. Optimization in Drug Research* (Testa B., Van de Waterbeemd H., Folkers G., Guy R., eds.) Wiley - VCH, Weinheim and VHC, Zurich, pp. 513-524.
- Hamelink, J., Landrum, P.F., Bergman, H., & Benson, W.H. (1994) *Bioavailability: physical, chemical, and biological interactions*. CRC Press.
- U.S. Code of Federal Regulations. 21 CFR 320.1.
- Devillers, J., & Balaban, A.T. (Eds.). (2000) *Topological indices and related descriptors in QSAR and QSPAR*. CRC Press.
- Raevskij, O.A. (2015) Modelirovanie sootnoshenij "struktura-svoystva", Dobrosvet, M. (in Russian)
- Raevskii, O.A., Avidon, V.V., & Novikov, V.P. (1982) Use of a unified scale of donor-acceptor interactions for the analysis of the similarity of the structures of biologically active compounds. *Pharmaceutical Chemistry Journal*, 16(8), 633-636. DOI: 10.1007/BF00773099
- Raevskii, O.A., & Novikov, V.P. (1982) Unification of the characteristics of donor-acceptor interactions in the framework of the problem of the structure-activity relationship. *Khim.-Farm. Zh.*, 16(5), 583-586.
- Raevsky, O.A., Grigor'ev, V.Y., & Solov'ev, V.P. (1989) The Estimation of Donor-Acceptor Parameters in Biologically Active Compounds. *Khim. Pharm. Zhurn.(Rus.)*, 23, 1294-1300.
- Martynov, I.V., & Raevsky, O.A. (1983) Physical-chemical approach to purposeful search of biological active substances. *Vestnik akademii nauk SSSR*, (7), 93-101.
- Martynov, I.V., & Raevskii, O.A. (1983) Estimation of electron-donor and acceptor ability of some active-centers in molecules of physiologically active compounds. *Zhurnal vsesoyuznogo khimicheskogo obshchestva imeni D.I. Mendeleeva*, 28(6), 716-717.
- Raevskii, O.A., Grigor'ev, V.Y., & Solov'ev, V.P. (1984) Evaluation of the electron-donor and electron-acceptor functions of ionized atoms and groups in biologically active substances on the basis of thermodynamic data. *Pharmaceutical Chemistry Journal*, 18(5), 327-331. DOI:
- Sapegin, A.M., Raevsky, O., Chistyakov, V.V., & Martynov, I.V. (1987) Donor-acceptor factor-analysis algorithm to depict molecules of biologically-active compounds. *Khimiko-farmatsevticheskii zhurnal*, 21(9), 1098-1102.
- Raevsky, O., & Sapegin, A.M. (1987) Developed physicochemical approach to recognition of physiologically active compound structures. *Khimiko-farmatsevticheskii zhurnal*, 21(11), 1338-1341.
- Sapegin, A.M., Razdolsky, A., Chistyakov, V.V., & Raevsky, O. (1987) Structure recognition-realization of a physicochemical approach to the study of structure-activity-relationships. *Khimiko-farmatsevticheskii zhurnal*, 21(11), 1341-1344.
- Raevsky, O.A., Solotnov, A.F., & Solovyev, V.P. (1987) Electron-donating and electrophilic functions of physiological acting and model compounds. *Journal of common chemistry*, 57(6), 1241-1248.
- Raevskii, O.A., Grigoriev, V., Soloviev, V., & Martynov, I.V. (1988) Electron-acceptor enthalpy factors of phenols. *Doklady Akademii Nauk SSSR*, 298(5), 1166-1169.
- Raevsky, O.A., Grigor'ev, V.Y., Kireev, D.B., & Zefirov, N.S. (1992) Complete thermodynamic description of H-bonding in the framework of multiplicative approach. *Quantitative Structure-Activity Relationships*, 11(1), 49-63. DOI: 10.1002/qsar.19920110109
- Raevsky, O.A. (1997) Quantification of non-covalent interactions on the basis of the thermodynamic hydrogen bond parameters. *Journal of physical organic chemistry*, 10(5), 405-413. DOI: 10.1002/(SICI)1099-1395(199705)10:5<405::AID-POC922>3.0.CO;2-A
- Raevsky, O.A. (1997) Hydrogen bond strength estimation by means of the HYBOT program package. *Computer-Assisted Lead Finding and Optimization: Current Tools for Medicinal Chemistry*, 367-378.
- Abraham, M.H., Ibrahim, A., Zissimos, A.M., Zhao, Y.H., Comer, J., & Reynolds, D.P. (2002) Application of hydrogen bonding calculations in property based drug

- design. *Drug Discovery Today*, 7(20), 1056-1063. DOI: 10.1016/S1359-6446(02)02478-9
26. Ruelle, P. (1999) Towards a comprehensive non-ergodic treatment of H-bonds and hydrophobicity in real solutions: The mobile order and disorder theory. *Perspectives in Drug Discovery and Design*, 17(1), 61-96.
27. Laurence, C., Brameld, K.A., Graton, J., Le Questel, J.Y., & Renault, E. (2009) The p K BHX database: toward a better understanding of hydrogen-bond basicity for medicinal chemists. *Journal of medicinal chemistry*, 52(14), 4073-4086. DOI: 10.1021/jm801331y
28. Raevsky, O.A., Grigor'ev, V.J., & Trepalin, S.V. (1999) HYBOT program package. Registration by Russian State Patent Agency, (990090).
29. Trepalin, S.V., Yarkov, A.V. (2001) CheD: Chemical Database Compilation Tool, Internet Server, and client for SQL Servers. *J. Chem. Inf. Comput. Sci.*, 41, 100-107. DOI: 10.1021/ci000039n
30. Verma, J., Khedkar, V.M., & Coutinho, E.C. (2010) 3D-QSAR in drug design-a review. *Current topics in medicinal chemistry*, 10(1), 95-115. DOI: 10.2174/156802610790232260
31. SYBYL-X 2.1. Certara, Princeton, NJ, USA
32. Pearlman, D.A., Case, D.A., Caldwell, J.W., Seibel, G.L., Singh, U.C., Weiner, P., & Kollman, P.A. (1991) AMBER 4.0, University of California, San Francisco.
33. Boobbyer, D.N., Goodford, P.J., McWhinnie, P.M., & Wade, R.C. (1989) New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure. *Journal of medicinal chemistry*, 32(5), 1083-1094. DOI: 10.1021/jm00125a025
34. Groom, C.R., Bruno, I.J., Lightfoot, M.P., & Ward, S.C. (2016) The Cambridge structural database. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 72(2), 171-179. DOI: 10.1107/S2052520616003954
35. Raevsky, O.A., & Skvortsov, V.S. 3D HYBOT program. Russian State Patent Department, (004612207).
36. Trepalin, S.V., Yarkov, A.V., & Raevsky, O.A. (2018) MOLTRA-II. New three dimensional descriptors of the hydrogen bond. *Biomedical Chemistry: Research and Methods*, 1(3), e00069. DOI: 10.18097/bmcrm00069
37. Van de Waterbeemd, H., & Testa, B. (1987) The parametrization of lipophilicity and other structural properties in drug design. *Advances in drug research*, 16, 85-225.
38. Raevsky, O.A., Schaper, K.J., & Seydel, J.K. (1995) H-Bond Contribution to Octanol-Water Partition Coefficients of Polar Compounds. *Quantitative Structure-Activity Relationships*, 14(5), 433-436. DOI: 10.1002/qsar.19950140504
39. Raevsky O.A., Trepalina E.P., Trepalin S.V. (2000) in: *Molecular Modelling and Prediction of Bioactivity* (Gundertofe K. and Jorgensen F., eds.), Kluwer Academic/Plenum Publ.; p. 489-490.
40. Raevsky, O.A. (2001) Molecular lipophilicity calculations of chemically heterogeneous chemicals and drugs on the basis of structural similarity and physicochemical parameters. *SAR and QSAR in Environmental Research*, 12(4), 367-381. DOI: 10.1080/10629360108033245
41. Raevsky, O.A., Trepalin, S.V., Trepalina, H.P., Gerasimenko, V.A., & Raevskaja, O.E. (2002) SLIPPER-2001 - Software for predicting molecular properties on the basis of physicochemical descriptors and structural similarity. *Journal of chemical information and computer sciences*, 42(3), 540-549. DOI: 10.1021/ci010097o
42. Lipinski, C.A. (2000) Drug-like properties and the causes of poor solubility and poor permeability. *Journal of pharmacological and toxicological methods*, 44(1), 235-249. DOI: 10.1016/S1056-8719(00)00107-6
43. Alelyunas, Y.W., Empfield, J.R., McCarthy, D., Spreen, R.C., Bui, K., Pelosi-Kilby, L., & Shen, C. (2010) Experimental solubility profiling of marketed CNS drugs, exploring solubility limit of CNS discovery candidate. *Bioorganic & medicinal chemistry letters*, 20(24), 7312-7316. DOI: 10.1016/j.bmcl.2010.10.068
44. Hansch, C., Quinlan, J.E., & Lawrence, G. L. (1968) Linear free-energy relationship between partition coefficients and the aqueous solubility of organic liquids. *The Journal of Organic Chemistry*, 33(1), 347-350. DOI: 10.1021/jo01265a071
45. Schaper, K.J., Kunz, B., & Raevsky, O.A. (2003) Analysis of water solubility data on the basis of HYBOT descriptors: Part 2. Solubility of liquid chemicals and drugs. *QSAR & Combinatorial Science*, 22(9-10), 943-958. DOI: 10.1002/qsar.200330840
46. Dearden, J.C. (2006) *In silico* prediction of aqueous solubility. *Expert opinion on drug discovery*, 1(1), 31-52. DOI: 10.1517/17460441.1.1.31
47. Faller, B., & Ertl, P. (2007) Computational approaches to determine drug solubility. *Advanced drug delivery reviews*, 59(7), 533-545. DOI: 10.1016/j.addr.2007.05.005
48. Johnson, S.R., & Zheng, W. (2006) Recent progress in the computational prediction of aqueous solubility and absorption. *The AAPS journal*, 8(1), E27-E40. DOI: 10.1208/aapsj080104
49. Sugano, K., Okazaki, A., Sugimoto, S., Tavornvivas, S., & Omura, A. (2007) Solubility and dissolution profile assessment in drug discovery. *Drug metabolism and pharmacokinetics*, 22(4), 225-254. DOI: 10.2133/dmpk.22.225
50. Wang, J., & Hou, T. (2011) Recent advances on aqueous solubility prediction. *Combinatorial chemistry & high throughput screening*, 14(5), 328-338. DOI:
51. Skyner, R.E., McDonagh, J.L., Groom, C.R., Van Mourik, T., & Mitchell, J.B.O. (2015) A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Physical Chemistry Chemical Physics*, 17(9), 6174-6191. DOI: 10.1039/C5CP00288E
52. Chevillard, F., Lagorce, D., Reyne's, C., Villoutreix, B.O., Vayer, P., & Miteva, M.A. (2012) *In silico* prediction of aqueous solubility: a multimodel protocol based on chemical similarity. *Molecular pharmaceutics*, 9(11), 3127-3135. DOI: 10.1021/mp300234q
53. Yalkowsky, S.H., & Valvani, S.C. (1980) Solubility and partitioning I: solubility of nonelectrolytes in water. *Journal of pharmaceutical sciences*, 69(8), 912-922. DOI: 10.1002/jps.2600690814
54. Raevsky, O.A., Polianczyk, D.E., Grigorev, V.Y., Raevskaja, O.E., & Dearden, J.C. (2015) *In silico* prediction of aqueous solubility: A comparative study of local and global predictive models. *Molecular informatics*, 34(6-7), 417-430. DOI: 10.1002/minf.201400144
55. Raevsky, O.A., Grigor'ev, V.Y., Polianczyk, D.E., Raevskaja, O.E., & Dearden, J.C. (2014) Calculation of aqueous solubility of crystalline un-ionized organic chemicals and drugs based on structural similarity and physicochemical descriptors. *Journal of chemical information and modeling*, 54(2), 683-691. DOI: 10.1021/ci400692n
56. DRAGON, version 5.5; Talete srl, Milano, Italy, 2011.
57. Obrezanova, O., Csányi, G., Gola, J.M., & Segall, M.D. (2007) Gaussian processes: a method for automatic QSAR modeling of ADME properties. *Journal of chemical information and modeling*, 47(5), 1847-1857. DOI: 10.1021/ci7000633

58. Obrezanova, O., Gola, J.M., Champness, E.J., & Segall, M.D. (2008) Automatic QSAR modeling of ADME properties: blood-brain barrier penetration and aqueous solubility. *Journal of computer-aided molecular design*, 22(6-7), 431-440. DOI 10.1007/s10822-008-9193-8
59. Raevsky, O.A., Grigorev, V.Y., Polianczyk, D.E., Raevskaja, O.E., & Dearden, J.C. (2017) Six global and local QSPR models of aqueous solubility at pH=7.4 based on structural similarity and physicochemical descriptors. *SAR and QSAR in Environmental Research*, 28(8), 661-676. DOI: 10.1080/1062936X.2017.1368704
60. Raevsky, O.A., Fetisov, V.I., Trepalina, E.P., McFarland, J.W., & Schaper, K.J. (2000) Quantitative estimation of drug absorption in humans for passively transported compounds on the basis of their physico-chemical parameters. *Quantitative Structure-Activity Relationships*, 19(4), 366-374. DOI: 10.1002/1521-3838(200010)19:4<366::AID-QSAR366>3.0.CO;2-E
61. Raevsky, O.A., & Skvortsov, V.S. (2005) Quantifying hydrogen bonding in QSAR and molecular modeling. *SAR and QSAR in Environmental Research*, 16(3), 287-300. DOI: 10.1080/10659360500036893
62. Raevsky, O.A., Schaper, K.J., Artursson, P., & McFarland, J.W. (2001) A novel approach for prediction of intestinal absorption of drugs in humans based on hydrogen bond descriptors and structural similarity. *Quantitative Structure-Activity Relationships*, 20(5-6), 402-413. DOI: 10.1002/1521-3838(200112)20:5/6<402::AID-QSAR402>3.0.CO;2-6
63. Bradbury, M.W.B. (1979) *The concept of a blood-brain barrier*. John Wiley & Sons.
64. Wolf, S., Seehaus, B., Minol, K., & Gassen, H.G. (1996) The blood-brain barrier: a specialty of cerebral microcirculation systems. *Die Naturwissenschaften*, 83(7), 302-311.
65. Pauletti, G.M., Okumu, F.W., & Borchardt, R.T. (1997) Effect of size and charge on the passive diffusion of peptides across Caco-2 cell monolayers via the paracellular pathway. *Pharmaceutical research*, 14(2), 164-168. DOI: 10.1023/A:1012040425146
66. Hansch, C., Steward, A.R., Anderson, S.M., & Bentley, D.L. (1968) Parabolic dependence of drug action upon lipophilic character as revealed by a study of hypnotics. *Journal of medicinal chemistry*, 11(1), 1-11. DOI: 10.1021/jm00307a001
67. Young, R.C., Mitchell, R.C., Brown, T.H., Ganellin, C.R., Griffiths, R., Jones, M., ... & Smith, I.R. (1988) Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H2 receptor histamine antagonists. *Journal of medicinal chemistry*, 31(3), 656-671. DOI: 10.1021/jm00398a028
68. van de Waterbeemd, H., & Kansy, M. (1992) Hydrogen-bonding capacity and brain penetration. *CHIMIA International Journal for Chemistry*, 46(7-8), 299-303.
69. Lipinski, C.A., Lombardo, F., Dominy, B.W., & Feeney, P.J. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, 23(1-3), 3-25. DOI: 10.1016/S0169-409X(96)00423-1
70. Norinder, U., & Haeberlein, M. (2002) Computational approaches to the prediction of the blood-brain distribution. *Advanced drug delivery reviews*, 54(3), 291-313. DOI: 10.1016/S0169-409X(02)00005-4
71. Didziapetris, R., Japertas, P., Avdeef, A., & Petrauskas, A. (2003) Classification analysis of P-glycoprotein substrate specificity. *Journal of drug targeting*, 11(7), 391-406. DOI: 10.1080/10611860310001648248
72. Pajouhesh, H., & Lenz, G.R. (2005) Medicinal chemical properties of successful central nervous system drugs. *NeuroRx*, 2(4), 541-553. DOI: 10.1602/neurorx.2.4.541
73. Borchardt, R., Kerns, E., Hageman, M., Thakker, D., & Stevens, J. (Eds.) (2007) *Optimizing the "drug-like" Properties of Leads in Drug Discovery*. Springer Science & Business Media.
74. Wager, T.T., Chandrasekaran, R.Y., Hou, X., Troutman, M.D., Verhoest, P.R., Villalobos, A., & Will, Y. (2010) Defining desirable central nervous system drug space through the alignment of molecular properties, *in vitro* ADME, and safety attributes. *ACS chemical neuroscience*, 1(6), 420-434. DOI: 10.1021/cn100007x
75. Ghose, A.K., Herbertz, T., Hudkins, R.L., Dorsey, B.D., & Mallamo, J.P. (2011) Knowledge-based, central nervous system (CNS) lead selection and lead optimization for CNS drug discovery. *ACS chemical neuroscience*, 3(1), 50-68. DOI: 10.1021/cn200100h
76. Desai, P.V., Sawada, G.A., Watson, I.A., & Raub, T.J. (2013) Integration of *in silico* and *in vitro* tools for scaffold optimization during drug discovery: predicting P-glycoprotein efflux. *Molecular pharmaceutics*, 10(4), 1249-1261. DOI: 10.1021/mp300555n
77. Panarin, V.A., Kondratyev, V.A., & Rayevsky, O.A. (1990) Some characteristics of the functioning of membrane receptor-channel complexes of *Limnaea stagnalis* neurones. *The Journal of physiology*, 423(1), 363-380. DOI: 10.1113/jphysiol.1990.sp018027
78. Hitchcock, S.A., & Pennington, L.D. (2006) Structure-brain exposure relationships. *Journal of medicinal chemistry*, 49(26), 7559-7583. DOI: 10.1021/jm060642i
79. Raevsky, O.A. (1990) The structure and properties of complexes simulating molecular recognition. *Uspekhi Khimii*, 59(3), 375-400.
80. Raevsky, O.A. (2004) Physicochemical descriptors in property-based drug design. *Mini reviews in medicinal chemistry*, 4(10), 1041-1052. DOI: 10.2174/1389557043402964
81. van de Waterbeemd, H., Camenisch, G., Folkers, G., Chretien, J.R., & Raevsky, O.A. (1998) Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors. *Journal of drug targeting*, 6(2), 151-165. DOI: 10.3109/10611869808997889
82. Raevsky, O.A., & Schaper, K. J. (1998) Quantitative estimation of hydrogen bond contribution to permeability and absorption processes of some chemicals and drugs. *European journal of medicinal chemistry*, 33(10), 799-807. DOI: 10.1016/S0223-5234(99)80031-2
83. Raevsky, O.A., Solodova, S.L., Raevskaya, O.E., Liplavskiy, Y.V., & Mannhold, R. (2012) The computer classification models on the relationship between chemical structures of compounds and drugs with their blood brain barrier penetration ability. *Biochemistry (Moscow) Supplement Series B: Biomedical Chemistry*, 6(1), 31-38. DOI: 10.1134/S1990750812010131
84. Raevsky, O.A., Solodova, S.L., Raevskaya, O.E., & Mannhold, R. (2012) Quantitative interaction between the structures of organic compounds and their abilities to penetrate the blood-brain barrier. *Pharmaceutical Chemistry Journal*, 46(3), 133-138. DOI: 10.1007/s11094-012-0748-6
85. Rankovic, Z. (2015) CNS drug design: balancing physicochemical properties for optimal brain exposure. *Journal of medicinal chemistry*, 58(6), 2584-2608. DOI: 10.1021/jm501535r

86. Raevsky, O.A. (2016) CNS Multiparameter Optimization Approach: Is it in Accordance with Occam's Razor Principle? *Molecular informatics*, 35(3-4), 94-98. DOI: 10.1002/minf.201500109
87. Raevsky, O.A., Polianczyk, D.E., Mukhametov, A., & Grigorev, V.Y. (2016) Assessment of the classification abilities of the CNS multi-parametric optimization approach by the method of logistic regression. *SAR and QSAR in Environmental Research*, 27(8), 629-635. DOI: 10.1080/1062936X.2016.1212922
88. Raevsky, O.A., Grigorev, V.Y., Polianczyk, D.E., Sandakov, G.I., Solodova, S.L., Yarkov, A.V., Bachurin, S.O. & Dearden, J.C. (2016) Physicochemical property profile for brain permeability: comparative study by different approaches. *Journal of drug targeting*, 24(7), 655-662. DOI: 10.3109/1061186X.2015.1132224

Поступила: 13. 08. 2018.

Принята к публикации: 13. 09. 2018.

HYDROGEN BOND CONTRIBUTION TO DRUG BIOAVAILABILITY: CHEMINFORMATICS APPROACH

O.A. Raevsky

Institute of Physiologically Active Compounds of the Russian Academy of Sciences,
1 Severny proezd, Moscow region, Chernogolovka, 142432 Russia; *e-mail: raevsky@ipac.ac.ru

A review, based mainly on own publications, is devoted to methods of investigation of "structure-bioavailability" relationships. The first part of this review contains information about classification of hydrogen bond descriptors, original 2D hydrogen bond thermodynamic descriptors, program HYBOT, original 3D hydrogen bonding potentials, original hydrogen bond surface area descriptors. The second part includes the results of applications of the above mentioned of hydrogen bond descriptors for prediction of bioavailability components such as lipophilicity, solubility in water and in physiological fluids, absorption and blood-brain barrier permeability.

Key words: QSAR; HYBOT; hydrogen bond descriptors; bioavailability

ACKNOWLEDGMENTS

The work was performed within the framework of the state task for 2018 (the topic number 0090-2017-0020).