

ТЕОРИЯ  
И МОДЕЛИРОВАНИЕ

УДК 541.64:539.199:532.73

ПРЕДСКАЗАНИЕ ПАРАМЕТРОВ РАСТВОРИМОСТИ ЛЕГКИХ ГАЗОВ  
В СТЕКЛООБРАЗНЫХ ПОЛИМЕРАХ НА ОСНОВЕ МОДЕЛИРОВАНИЯ  
КОРОТКОГО ОТРЕЗКА ПОЛИМЕРНОЙ ЦЕПИ

© 2019 г. М. В. Губко<sup>а,\*</sup>, О. А. Милосердов<sup>а,б</sup>, Ю. П. Ямпольский<sup>б</sup>, В. Е. Рыжих<sup>б</sup>

<sup>а</sup> Институт проблем управления им. В.А. Трапезникова Российской академии наук  
117997 Москва, ул. Профсоюзная, 65, Россия

<sup>б</sup> Институт нефтехимического синтеза им А.В. Топчиева Российской академии наук  
119991 Москва, Ленинский пр., 29, Россия

\*e-mail: mgoubko@mail.ru

Поступила в редакцию 02.03.2019 г.

После доработки 26.04.2019 г.

Принята к публикации 13.05.2019 г.

Разработан метод предсказания сорбционных параметров полимеров (коэффициента растворимости газов при бесконечном растворении  $S$  и константы равновесия закона Генри  $k_D$  модели двойной сорбции) большинства легких газов в стеклообразных полимерах на основе компьютерной модели отрезка полимерной цепи. Из молекулярно-динамической модели конформаций короткого (порядка 200–600 атомов) отрезка полимерной цепи вычисляются зависимости ряда геометрических индексов, таких как площадь доступной поверхности, площадь доступной поверхности с положительным частичным зарядом и т.д., от радиуса “обкатки” (фактически от эффективного радиуса пенетранта или сорбата). Коэффициенты линейной аппроксимации этих зависимостей используются как объясняющие переменные во множественной линейной регрессии. Значимые переменные и их веса в регрессии находятся на основе экспериментальных измерений из Базы данных физико-химических свойств полимеров Лаборатории мембранного газоразделения Института нефтехимического синтеза Российской академии наук. За счет использования более точного алгоритма расчета площадей макромолекул удалось устранить недостатки первой версии данного метода. Для  $S$  коэффициент детерминации (скорректированный  $R^2$ ) регрессии на тестовой выборке равен 0.8 (соответствует средней относительной ошибке 77% предсказания  $S$ , что считается неплохой точностью), для  $k_D$  коэффициент детерминации составляет 0.81, средняя относительная ошибка 47%, что также является хорошим результатом с учетом больших (до десятичного порядка) погрешностей экспериментального измерения  $S$  и  $k_D$ .

DOI: 10.1134/S2308112019050067

ВВЕДЕНИЕ

Мембранное разделение – сравнительно новое направление промышленного разделения газовых смесей, быстро развивающееся благодаря ряду технологических преимуществ по сравнению с традиционными методами разделения газов (ректификацией, абсорбцией и адсорбцией) [1–4]. К достоинствам мембранного газоразделения относятся низкое энергопотребление, отсутствие фазовых переходов, малая материалоемкость, гибкость управления и модульность структуры технологических установок. Наибольшее распространение получили промышленные процессы мембранного газоразделения для выделения азота из воздуха, удаления углекислого газа из “кислых” природных газов и извлечение водорода из различных потоков химии и нефтепереработки.

Но в основе всего лежит квалифицированный выбор материала мембран и возможность приготовления мембран, которые были бы при этом и высокопроизводительными и селективными. В подавляющем большинстве процессов мембранного газоразделения используются полимеры. Синтез или выбор полимера и изучение его свойств и поведения в процессе газоразделения является предметом мембранного материаловедения. Важнейшая задача мембранного материаловедения – поиск новых высокоэффективных полимеров для решения разнообразных задач мембранной технологии. Для мембранного газоразделения решающую роль играют транспортные свойства полимерного материала – проницаемость  $P_i$  и селективность  $P_i/P_j$  для пары разделяемых газов  $i$  и  $j$ . Возможность хотя бы приблизительно предсказать транспортные свойства гипоте-

тического полимера еще до его синтеза существенно сокращает усилия и экономит время, позволяя отбросить заведомо неудачные варианты и сконцентрироваться на наиболее перспективных соединениях.

Традиционно важнейшим фактором, влияющим на транспортные свойства аморфного полимера, считается его химическое строение. Химическая природа мономерных звеньев и способы их объединения в полимерную цепь определяют взаимное расположение цепей и геометрию свободного объема внутри образца. Именно структура свободного объема полимерной матрицы определяет коэффициент проницаемости  $P_i$  молекул газа  $i$  через мембрану.

Как известно, коэффициент проницаемости  $P$  можно записать в виде произведения коэффициента растворимости  $S$  и коэффициента диффузии  $D$ :

$$P = S \cdot D \quad (1)$$

Первый описывает движущую силу процесса переноса молекул газа, второй соответствует кинетической компоненте процесса. Во многих случаях именно коэффициент растворимости  $S$ , а также селективность растворимости  $S_i/S_j$  по паре газов  $i$  и  $j$  определяют транспортные свойства и ценность полимера для мембранной технологии. Это обуславливает актуальность задачи предсказания  $S$  для вновь создаваемых полимерных материалов.

Согласно модели двойной сорбции [5], общее количество  $C$  газа, растворенного в объеме аморфного стеклообразного полимера в зависимости от давления  $p$ , описывается комбинацией закона Генри и сорбции лэнгмюровского типа по формуле

$$C = k_D p + \frac{C'_H b p}{1 + b p}, \quad (2)$$

где  $k_D$  – константа равновесия закона Генри,  $b$  – константа равновесия лэнгмюровской сорбции,  $C'_H$  – лэнгмюровская сорбционная емкость. Коэффициент растворимости при бесконечном разбавлении  $S$  вычисляется как отношение  $C/p$  при  $p \rightarrow 0$ , т.е. из изотермы сорбции (2),  $S$  можно представить в виде суммы двух слагаемых:  $S = k_D + C'_H b$ .

В настоящей статье развивается подход [1] к предсказанию коэффициента растворимости  $S$  легких газов в стеклообразных аморфных полимерах, основанный на компьютерном моделировании конформаций короткой полимерной цепи. При той же точности и сравнимом числе объясняющих переменных усовершенствованный метод позволил устранить ряд недостатков подхода [1], а также предсказать константу закона Генри с относительной точностью 47%.

Как и в работе [1], для предсказания транспортных свойств генерируются несколько конформаций короткого отрезка цепи рассматриваемого полимера (обычно включающего несколько мономерных звеньев и от 200 до 600 атомов) путем релаксации в эмпирическом поле Дрейдинга из случайных начальных позиций атомов.

Виртуальные шаровые модели полученных конформеров “обкатываются” пенетрантом различного радиуса в диапазоне от 0 до 3 Å и строится зависимость площади доступной поверхности (площади контакта макромолекулы с пенетрантом) от радиуса обкатки. Параллельно вычисляются еще несколько геометрических химических индексов (см. детали в разделе 4). В отличие от [1], вместо эмпирического алгоритма [6], реализованного в Molecular Surface Area Plugin продукта ChemAxon, для “обкатки” используется (более точный и прозрачно реализованный) классический алгоритм Ли–Ричардса [7].

Полученные зависимости линейризуются на диапазоне радиусов обкатки, который в отличие от работы [1] был адаптивно подобран для получения наилучших результатов. Линеаризованная зависимость значения  $i$ -го индекса от радиуса обкатки  $R$  имеет вид  $a_i + b_i R$ , и пары коэффициентов  $a_i$ ,  $b_i$  используются в качестве основных объясняющих переменных для предсказания логарифма как растворимости  $S$ , так и константы равновесия закона Генри  $k_D$ .

Вес объясняющих переменных для предсказания  $\lg S$  был подобран на основе того же, что и в работе [1] массива из 344 экспериментальных измерений (плюс еще 134 измерения были отобраны для тестовой выборки) 64 полимеров различных классов из Базы данных физико-химических свойств полимеров Лаборатории мембранного газоразделения ИНХС РАН. Регрессия для  $\lg k_D$  настраивалась на основе 76 экспериментальных измерений (плюс 17 измерений, входящих в тестовую выборку) 40 полимеров.

Анализ построенных регрессий позволяет выделить ключевые аспекты структуры полимера, влияющие на растворимость и, главное, на селективность растворимости: это чувствительность площади поляризованной поверхности молекулы к радиусу “зонда” и площадь доступной поверхности с учетом заряда. Понимание определяющих структурных факторов помогает направить поиск новых полимерных материалов за счет их отбора по результатам компьютерного моделирования. В то же время предлагаемые процедуры вычислительно достаточно дешевые, что позволяет быстро в автоматическом режиме перебирать сотни потенциальных формул стеклообразных полимеров с различными, в том числе экстремальными, транспортными свойствами.

## ОБЗОР ЛИТЕРАТУРЫ

В литературе известно несколько подходов к предсказанию транспортных свойств полимеров. Наиболее хорошо разработанные и популярные основаны на атомистическом компьютерном моделировании (см. обзоры в [8, 9]). В частности, для предсказания коэффициентов проницаемости, диффузии и растворимости газов в полимерах использовались методы молекулярной динамики и большого канонического ансамбля [10, 11].

В методе молекулярной динамики [12] атомы перемещаются в соответствии с уравнениями движения Ньютона в эмпирическом силовом поле, моделирующим взаимодействия между молекулами газа и атомами полимерной цепи. Однако прямое применение метода молекулярной динамики к предсказанию диффузии невозможно, поскольку требует моделирования на колоссальном временном отрезке. Теория переходного состояния [13] позволила преодолеть эти трудности, что обусловило популярность данного подхода в моделировании процессов диффузии. В теории переходного состояния и ее многочисленных вариантах [14, 15] используется иерархия взаимосвязанных моделей процесса диффузии, причем каждая модель действует на своем временном и пространственном уровне [8], от поправочных членов к эмпирическому силовому полю до геометрического анализа границ “ям” потенциальной энергии (так называемых стабильных состояний) и марковской сетевой модели случайных переходов частицы газа между этими стабильными состояниями.

В отличие от метода молекулярной динамики широко используемого при моделировании диффузии, методы большого канонического ансамбля показали свою эффективность при моделировании растворимости. В тесте Видома [8] коэффициент растворимости вычисляется на основе потенциальной энергии молекулы газа, помещенной в случайное место полимерной фазы.

Как метод молекулярной динамики, так и метод большого канонического ансамбля требуют в качестве входных данных компьютерную пространственную модель полимерной матрицы, которая обычно получается с помощью методов молекулярной механики. Последние сводятся к минимизации потенциальной энергии системы выбором положений атомов одной или несколько полимерных цепочек, помещенных в куб размера порядка нескольких нанометров с периодическими граничными условиями (для ускорения релаксации положений атомов здесь также используется метод большого канонического ансамбля).

Стоит отметить, что в большинстве работ (за исключением, пожалуй, недавней статьи [16]), предполагается заранее известной плотность стеклообразной полимерной фазы, что делает

проблематичным применение такого моделирования в рамках молекулярного дизайна, когда требуется моделировать гипотетические, ранее не синтезированные или не изученные полимеры.

Кроме того, детальное моделирование таких сложных объектов, как полимерная матрица, требует весьма существенных даже по современным меркам вычислительных мощностей, применение суперкомпьютеров, что ограничивает возможность массового применения метода молекулярной механики для быстрого моделирования десятков и сотен перспективных полимеров (в том числе не синтезированных ранее) с целью отбора материалов с экстремальными значениями заданных транспортных характеристик.

Совершенно другой подход к предсказанию транспортных характеристик полимерных материалов, основанный на использовании топологических индексов и метода эмпирических групповых вкладов был предложен в [17, 18]. В методе эмпирических групповых вкладов каждому уникальному подграфу молекулярного графа (“группе”) присваивается такой статистический вес, чтобы взвешенная сумма числа вхождений в полимерное звено всевозможных молекулярных групп наилучшим образом коррелировала со значением интересующего физико-химического показателя на некоторой выборке экспериментальных данных. В частности, была показана эффективность метода для предсказания коэффициентов проницаемости и диффузии широкого круга газов в аморфных полимерах. К недостаткам стоит отнести относительно слабую предсказательную силу (способность к обобщению за пределы обучающей выборки), что является следствием сложности и нелинейности связи между топологией полимерного звена и структурой свободного объема полимерного образца.

Отметим еще несколько работ, в которых рассматривались родственные проблемы [19–24]. Монографию J. Visegano [19], в которой декларируется применение теории графов, но в действительности при вычислении конкретных параметров использовались подгоночные параметры, строго говоря, можно было бы исключить из рассмотрения. Важный вклад в вычисление многочисленных физико-химических параметров полимеров был внесен А. Аскадским [20, 21]. Однако среди большого числа предсказанных параметров (даже коэффициентов проницаемости для некоторых газов) коэффициенты растворимости газов отсутствуют. Метод, родственник используемому в настоящей работе, был применен И. Роновой с сотрудниками [22–24]. Он включает обкатку полимерных цепей сферическими зондами с размерами, моделирующими размеры молекул газовых пенетрантов. Но в этих работах рассматривались только корреляции экспериментальных значений  $P$  и  $D$  с доступным свободным

объемом у макромолекулярных цепей и никаких предсказаний не делалось. Такие корреляции были проанализированы в том числе и в работе [23], где они были построены и для коэффициентов растворимости, найденных как отношения экспериментальных параметров P/D.

В настоящей статье развивается относительно новый подход, впервые представленный в работе [1], который обладает следующими преимуществами: не требует экспериментальных данных о полимерном материале; применим к широкому кругу полимеров, в том числе еще не синтезированным; предъявляет весьма скромные требования к вычислительным мощностям.

Подход [1] основан на молекулярно-механическом моделировании относительно короткого отрезка полимерной цепи и использовании геометрических индексов, таких как ван-дер-ваальсов объем и площадь доступной поверхности макромолекулы для построения (на основе некоторого объема экспериментальных данных) многомерной регрессии для предсказания транспортных свойств полимеров, в частности, коэффициента растворимости при бесконечном разбавлении и константы равновесия закона Генри в модели двойной сорбции.

Первый этап – моделирование короткого отрезка полимерной цепи и использование площади доступной поверхности идейно наиболее близок подходу [25]. Второй этап – использование статистических инструментов для обучения линейной регрессии ближе к подходу [17, 18]. В то же время есть и многочисленные отличия, обуславливающие научную новизну подхода настоящей статьи: использование большого числа (а именно, девяти) различных геометрических индексов, идея о том, что важными объясняющими переменными могут стать параметры зависимости площади доступной поверхности (и родственных ей индексов) от радиуса “обкатки”, уточнен-

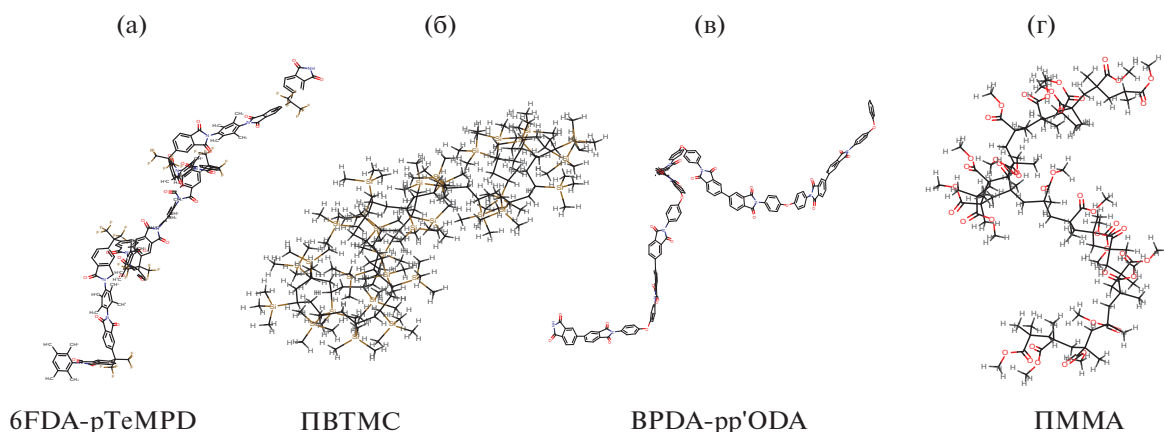
ный (по сравнению с работой [26]) общий вид зависимости транспортных характеристик от параметров пенетранта и полимера, статистически корректная процедура отбора значимых переменных для линейной регрессии, позволяющая избежать эффекта переобучения.

В целом описываемая ниже методика предсказания транспортных характеристик аморфных полимеров является попыткой соединить в себе достоинства метода эмпирических групповых вкладов и трудоемких, но аккуратных методов молекулярной динамики и большого канонического ансамбля.

### МОЛЕКУЛЯРНО-ДИНАМИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Подход к моделированию полимерной цепи в настоящей работе не претерпел изменений по сравнению с опубликованным в работе [1]. Как и ранее, для каждого полимера в среде молекулярного моделирования Instant JChem создается цепочка из нескольких мономерных звеньев, включающая от 200 до 600 атомов. Для этой цепочки средствами Conformer Plugin [27] пакета Instant JChem ChemAxon [28] генерируется 6 случайных конформаций различной геометрии, получаемой оптимизацией в эмпирическом силовом поле Дрейдинга из различных начальных позиций атомов. В следующем разделе обосновывается достаточность выбранной длины цепочки и количества генерируемых конформеров для устойчивости получаемых далее результатов.

Ниже приведено несколько полученных таким образом типичных конформаций для ряда “популярных” стеклообразных полимеров из Базы данных физико-химических свойств полимеров Лаборатории мембранного газоразделения ИНХС им. А.В. Топчиева [29].



ChemAxon Instant JChem считается довольно “всеядным” программным пакетом, не имеющим

узкой специализации по классам моделируемых веществ. В то же время ряд построенных конфор-

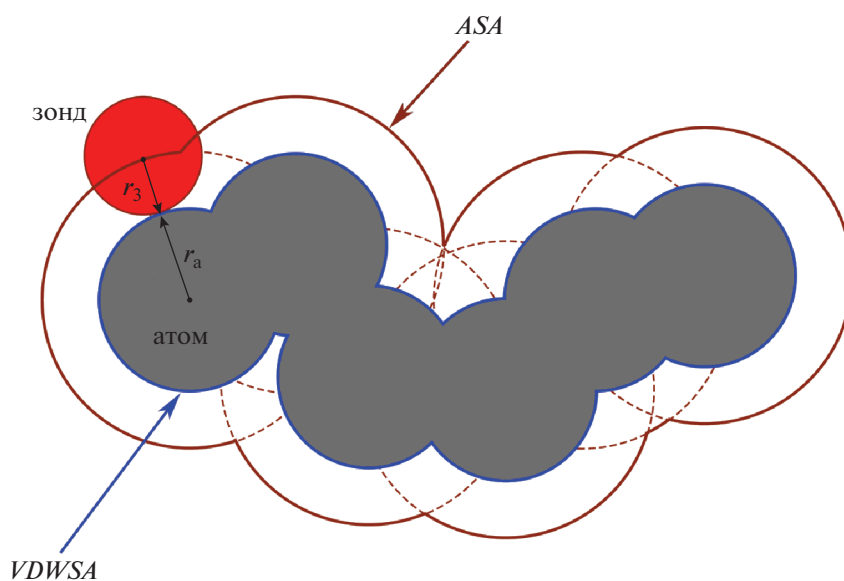
маций выглядел весьма нереалистично. Эти полимеры были пересчитаны в пакете PerkinElmer Chem3D версии 15.1.0.144 [30] по аналогичной методике; для задания случайного начального положения атомов к моделируемой полимерной цепи проводилось молекулярно-динамическое моделирование при температуре от 300 до 3000 К в течение 1000 итераций, после чего полученная структура оптимизировалась по свободной энергии в эмпирическом поле MM2.

Описанный подход позволяет достаточно точно моделировать локальную конфигурацию полимерной цепи, порождаемую взаимодействием соседних мономерных звеньев, но игнорирует влияние на форму молекулы далеко расположенных частей цепи и, что существенно, других полимерных цепей. Следовательно, от такого подхода не стоит ожидать точного предсказания ни свободного объема, ни распределения объема полостей по размерам, ни связности полостей — факторов, которые, как считается [31], сильно влияют на коэффициент проницаемости, но скорее через его кинетическую, диффузионную составляющую (см. формулу (1)).

## ВЫЧИСЛЕНИЕ ГЕОМЕТРИЧЕСКИХ ИНДЕКСОВ

В основе нашего метода лежит гипотеза о том, что коэффициент растворимости газа в полимере и связанные с этим характеристики в той или иной мере зависят от параметров поверхности контакта между макромолекулой полимера и молекулой газа. Например, в работе [32] была обоснована линейная зависимость между логарифмом коэффициента растворимости и площадью доступной поверхности молекулы газа-пенетранта. Коэффициенты такой линейной зависимости различны для разных полимеров, и логично предположить, что они как-то зависят от аналогичных параметров макромолекулы полимера.

Площадь доступной поверхности (accessible surface area, *ASA*) молекулы — это площадь поверхности, описываемой центром шарообразного “зонда” заданного радиуса во всевозможных положениях его касания с ван-дер-ваальсовой поверхностью (van der Waals surface area *VDWSA*) этой молекулы. Иллюстрация к построению этой поверхности для простой молекулы из нескольких атомов приведена ниже.



*ASA* является геометрическим индексом [33], т.е. функцией, которая каждой конформации молекулы (пространственному расположению ее атомов) ставит в соответствие число, зависящее только от взаимного расположения атомов, но не от положения молекулы в пространстве. По сравнению с похожим индексом — площадью поверхности контакта (площадью Конолли) — *ASA* макромолекулы имеет преимущество: именно площадь доступной поверхности пропорциональна вероятности контакта между макромолекулой

(полимера) и хаотически движущимся “зондом”, представляющим молекулу газа-пенетранта.

*ASA* для радиуса “зонда” 1.4 Å (соответствующего воде) широко используется в биохимии для изучения взаимодействия белковой макромолекулы с растворителями. В мембранной технологии газы-пенетранты отличаются своим эффективным радиусом, поэтому для целей мембранного газоразделения важно знать *ASA* макромолекул аморфных полимеров для различных радиусов “обкатки” *R*. Обкатка полимерных цепей сферическими зондами радиуса *R*, модели-

**Таблица 1.** Площадь доступной поверхности и родственные геометрические индексы

<i>ASA</i>	Удельная (на один см <sup>3</sup> образца) площадь доступной поверхности (Å <sup>2</sup> моль/см <sup>3</sup> )
<i>ASA<sup>+</sup></i>	Удельная площадь доступной поверхности, где контакт происходит в точке поверхности с частичным положительным ( <i>q</i> > 0) зарядом* (Å <sup>2</sup> моль/см <sup>3</sup> )
<i>ASA<sup>-</sup></i>	Удельная площадь доступной поверхности, где контакт происходит в точке поверхности с частичным отрицательным ( <i>q</i> < 0) зарядом* (Å <sup>2</sup> моль/см <sup>3</sup> )
<i>ASA<sub>H</sub></i>	Удельная площадь доступной гидрофобной (с низким, $ q  < 0,125$ , уровнем частичного заряда*) поверхности (Å <sup>2</sup> моль/см <sup>3</sup> )
<i>ASA<sub>P</sub></i>	Удельная площадь доступной поляризованной (с высоким, $ q  \geq 0,125$ , уровнем частичного заряда*) поверхности (Å <sup>2</sup> моль/см <sup>3</sup> )
<i>DPSA<sub>3</sub></i>	$DPSA_3 = \sum_i asa_i \cdot q_i$ , где <i>asa<sub>i</sub></i> – вклад <i>i</i> -го атома в удельную площадь доступной поверхности молекулы, <i>q<sub>i</sub></i> – это частичный заряд* <i>i</i> -го атома (Å <sup>2</sup> е моль/см <sup>3</sup> )*
<i>PPSA<sub>3</sub></i>	$PPSA_3 = \sum_i asa_i \cdot q_i$ , где суммирование ограничено атомами с положительным частичным зарядом*: $asa_i \cdot q_i > 0$ (Å <sup>2</sup> е моль/см <sup>3</sup> **)
<i>PNSA<sub>3</sub></i>	$PNSA_3 = \sum_i asa_i \cdot q_i$ , где суммирование ограничено атомами с отрицательным частичным зарядом*: $asa_i \cdot q_i < 0$ (Å <sup>2</sup> е моль/см <sup>3</sup> **)
<i>VDWV</i>	Удельный (на один см <sup>3</sup> ) ван-дер-ваальсов объем макромолекулы (Å <sup>3</sup> · моль/см <sup>3</sup> ): объем, очерчиваемый ван-дер-ваальсовой поверхностью, то есть доступной поверхностью при нулевом радиусе “обкатки”

\* Частичные заряды по Гастайгеру–Марсели вычислялись процедурой `rdPartialCharges` библиотеки `rdkit.Chem`.

\*\* Здесь под “е” понимается единица электронного заряда.

рующего размер молекулы газового пенетранта, применялась И. Роновой с сотрудниками [22–24] для вычисления доступного свободного объема (*FAV*) макромолекулярных цепей, однако, насколько нам известно, кривые зависимости *ASA(R)* полимерных материалов в задачах мембранной науки и технологии никогда не исследовались. В настоящей статье, как и в работе [1], именно параметры этой кривой и аналогичных ей позволяют предсказать коэффициент растворимости газа и константу равновесия закона Генри. В биохимии также широко используются несколько родственных *ASA* геометрических индексов (табл. 1) [33].

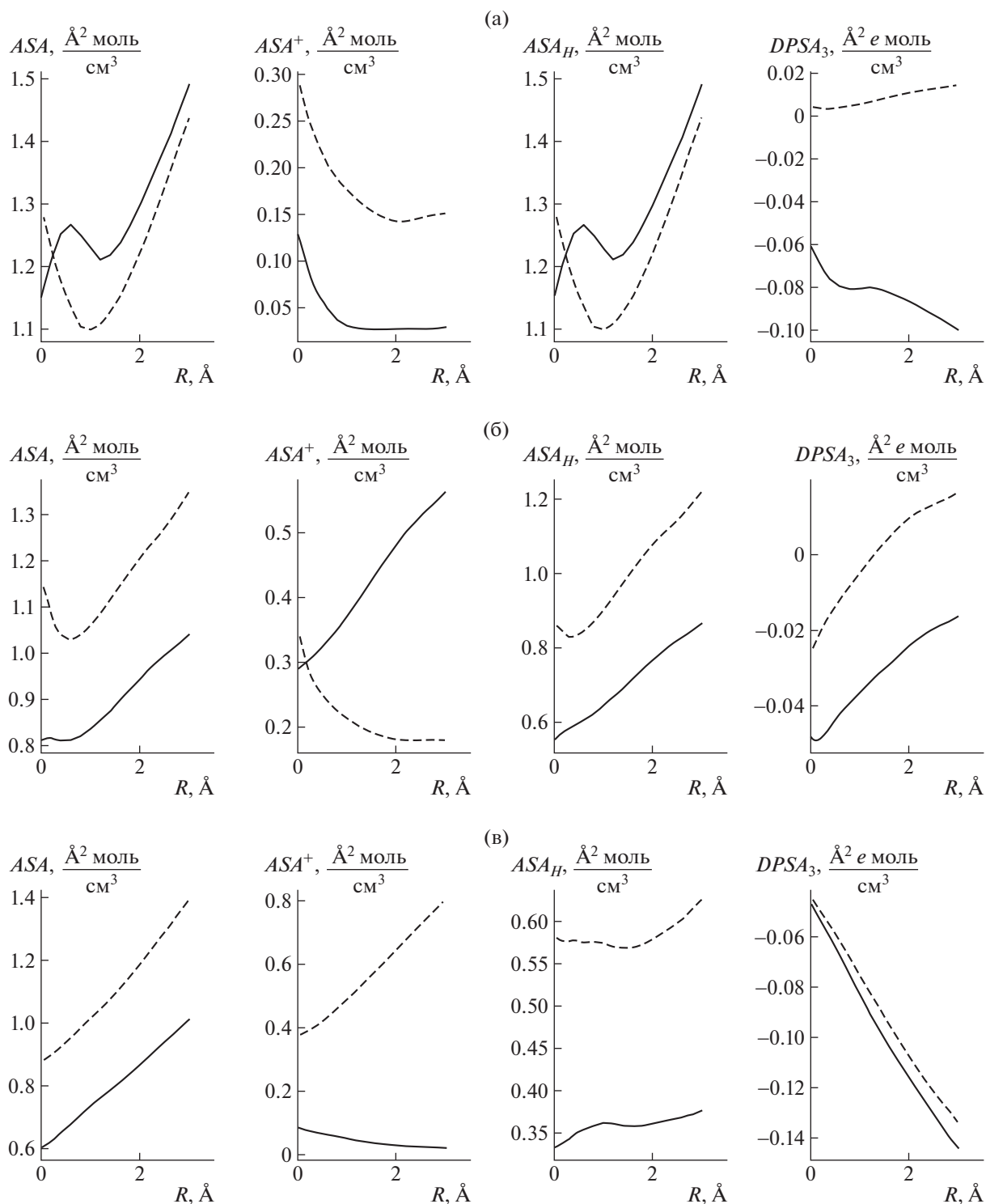
Перечисленные в табл. 1 индексы являются функциями радиуса “обкатки” *R*. Для построения

зависимостей *ASA(R)*, *ASA<sup>+</sup>(R)*, и т.д. значение индекса вычислялось для каждого *R* от 0 до 3 Å с шагом 0.05 Å. Значение геометрического индекса зависит от конформации макромолекулы. Для получения стабильной характеристики полимерного материала “вообще”, не зависящей от конкретной конформации, значения каждого индекса для каждого радиуса “обкатки” усреднялись по шести различным конформациям макромолекулы. В работе [1] для десяти случайным образом выбранных полимеров из исследуемой выборки была построена зависимость средней удельной площади доступной поверхности *ASA* от числа усредняемых конформаций. Было показано (см. рис. 4 в работе [1]), что вариации значения *ASA* не превышают 1% при усреднении шести конформаций и более, на основании чего принимается, что шесть конформаций – достаточно большая выборка для обеспечения инвариантности результата усреднения.

Аналогичным образом в работе [1] проводилось исследование зависимости *ASA* от длины моделируемой “олигомерной” цепи. Было установлено (см. рис. 5 в работе [1]), что для 10 случайным образом выбранных полимеров значение *ASA* (а также и остальных вычисляемых индексов) варьируется в пределах нескольких процентов при длине олигомерной цепи более 200 атомов. Отдельным интересным направлением будущих исследований является изучение зависимости минимальной допустимой длины олигомерной цепи от длины сегмента Куна.

В первой версии нашего метода [1] для вычисления площади доступной поверхности использовался эмпирический алгоритм [6], реализованный в `Molecular Surface Area Plugin` продукта `Chemaxon Instant JChem`. Опыт его использования, однако, показал, что помимо преимуществ (высокой скорости расчета и корректного учета групповых пересечений) этот алгоритм имеет и ряд недостатков. В частности, оказалось, что радиусы атомов и стандартных остатков специальным образом откалиброваны, в первую очередь, для белковых молекул в водной среде, что делает вычисленные значения индекса менее предсказуемыми и объяснимыми.

В настоящей статье описывается модифицированный подход, в котором мы используем собственную реализацию (на языке Python) классического алгоритма Ли–Ричардса. Помимо большей точности (достигаемой, впрочем, за счет увеличения трудоемкости) такая реализация алгоритма позволяет находить вклад каждого атома в площадь доступной поверхности для вычисления индексов табл. 1, зависящих от частичных зарядов. На рис. 1 приведены несколько типичных зависимостей геометрических индексов от радиу-



**Рис. 1.** Типичные зависимости геометрических индексов из табл. 1 от радиуса “обкатки” для нескольких полимеров различных классов (пунктиром даны индексы, рассчитанные с помощью алгоритма [6], сплошной линией – с помощью [7]). а – поливинилтриметилсилан, б – ПММА, в – полиимид 6FDA-pTeMPD.

са обкатки и их сравнение с индексами, рассчитанными с помощью эмпирического алгоритма [6].

Помимо отличия между алгоритмами, рисунок иллюстрирует разнообразие кривых. Имея общую тенденцию к возрастанию с увеличением  $R$

(что естественным образом следует из определения), зависимость площади доступной поверхности  $ASA$  от радиуса “обкатки”  $R$  может быть монотонной, иметь точку минимума, перегиб. Вид кривых для родственных индексов, зависящих от



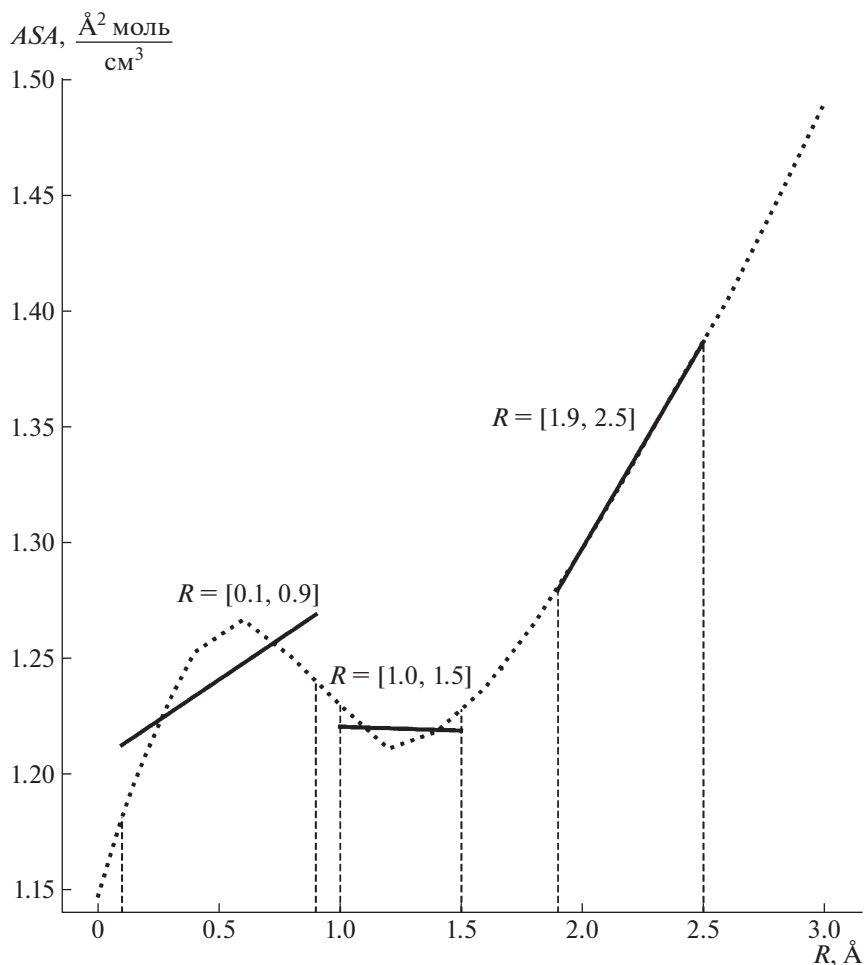


Рис. 2. Пример построения линейной аппроксимации ПВТМС *ASA* на отрезке  $[R_-, R_+]$ .

частичных зарядов, может быть и существенно более сложным, определяясь доступностью положительно/отрицательно заряженных участков молекулярной поверхности для “зондов” разного радиуса.

Совокупность этих восьми кривых дает довольно информативное описание поверхности макромолекулы с учетом частичных зарядов. Детальное исследование связи их вида с физико-химическими свойствами вещества является предметом будущих исследований. В настоящей работе мы будем пользоваться лишь их (достаточно грубой) линейной аппроксимацией. А именно, зададимся “эффективным” промежуточком радиусов  $[R^-, R^+]$ , на котором построим методом наименьших квадратов лучшую аппроксимацию кривых для всех восьми индексов из табл. 1. Процесс выбора оптимального диапазона аппроксимации будет подробно описан в следующем разделе.

Линейная аппроксимация *i*-го индекса описывается парой числовых коэффициентов (рис. 2): константа  $c_i$  соответствует значению линеаризо-

ванного индекса при нулевом радиусе, тангенс угла наклона  $d_i$  описывает чувствительность (или селективность) индекса к радиусу пенетранта (имея, как позже будет показано, прямое отношение к селективности растворимости полимера). Именно эти числовые коэффициенты, описывающие характер взаимодействия макромолекулы полимера с пенетрантами различного радиуса, используются далее в качестве объясняющих переменных для предсказания транспортных характеристик полимера наряду с удельным (на единицу объема полимера) ван-дер-ваальсовым объемом *VDWV* (табл. 1). Последний описывает объемные характеристики макромолекулы и считается одним из важных детерминантов свободного объема полимерной матрицы [34].

### ПРЕДСКАЗАНИЕ КОЭФФИЦИЕНТА РАСТВОРИМОСТИ

В работе [31] была обоснована линейная зависимость логарифма коэффициента растворимости (при бесконечном разбавлении) газа в поли-



мере от площади доступной поверхности молекулы газа-пенетранта. В работе [1] было показано, что  $MaxPA$  – максимальная площадь проекции молекулы газа (где атомы представлены своими ван-дер-ваальсовыми сферами) – позволяет еще лучше описать линейную зависимость логарифма коэффициента растворимости газа  $g$  в полимере  $p$  для большинства популярных в мембранной технологии легких газов (за исключением окиси углерода, демонстрирующей девиантное поведение):

$$\lg S(g, p) = a(p) + b(p)MaxPA(g) \quad (3)$$

При этом коэффициенты  $a$  и  $b$  линейной зависимости (3) являются функциями полимера  $p$ , отличаясь для различных полимеров. Задача настоящей статьи состоит в выявлении зависимости коэффициентов  $a$  и  $b$  от введенных в предыдущем разделе характеристик полимера – коэффициентов  $c_i$ ,  $d_i$  линейных аппроксимаций геометрических индексов макромолекулы полимера (см. табл. 1), а также от ван-дер-ваальсового объема  $VDWV$  макромолекулы.

Как и в работе [1], зависимость  $a$  и  $b$  от  $c_i$ ,  $d_i$  ( $i = ASA, \dots, PNSA_3$ ) предполагается линейной:

$$a(p) = a_0 + a_1c_{ASA} + a_2d_{ASA} + \dots + a_{15}c_{PNSA_3} + a_{16}d_{PNSA_3} + a_{17}VDWV, \quad (4)$$

$$b(p) = b_0 + b_1c_{ASA} + b_2d_{ASA} + \dots + b_{15}c_{PNSA_3} + b_{16}d_{PNSA_3} + b_{17}VDWV \quad (5)$$

Числовые коэффициенты  $a_0, \dots, a_{17}, b_0, \dots, b_{17}$  подбирали с помощью регрессионного анализа на основе экспериментальных данных из упомянутой Базы данных. Для обеспечения сопоставимости результатов использовали ту же выборку, что в предыдущей версии метода (см. работу [1]) с учетом удаления трех обнаруженных дублирующихся измерений. В результате осталось 478 уникальных измерений (344 в обучающей и 134 в тестовой выборке) по растворимости 18 различных легких газов (самым тяжелым из которых был бутан, СО в выборку не входил) в 64 полимерах различных классов (в основном полиимидов, полистиролов, полиакрилатов, виниловых полимеров). Коэффициент растворимости вычисляли по формуле (1) как отношение  $S = P/D$  экспериментально измеренных коэффициентов проницаемости и диффузии. Коэффициенты, полученные при разной температуре, приводили к единой температуре (308 К) по методике, описанной в работе [35]. Для расчета удельных (на единицу объема) значений геометрические индексы нормализовали на экспериментальные плотности (средние плотности, при наличии более одного источника) соответствующих образцов полимеров.

Подробное описание набора данных приведено в Дополнительных материалах, а критический анализ методики их подготовки описан в работе [1].

В связи с большим числом объясняющих переменных при небольшом объеме выборки для предотвращения переобучения использовали процедуру отбора переменных [36] на языке Rython, реализующую алгоритм шаговой регрессии [37] на основе статистики Фишера  $F$ : переменная с наименьшей вероятностью  $p(F)$  включается в регрессию, если  $p(F)$  меньше заданного порога  $f_{in}$ , переменная исключается, если для нее  $p(F)$  больше заданного порога  $f_{out}$ .

Выбор оптимального диапазона  $[R^-, R^+]$  линеаризации кривых  $ASA, \dots, PNSA_3$  и оптимальных параметров шаговой регрессии  $f_{in}, f_{out}$  производился 5-блочной кросс-валидацией; а именно, обучающая выборка случайным образом разбивалась на пять частей и для каждой комбинации  $R^- < R^+$  в диапазоне от 0 до 3 Å, а также  $f_{in} < f_{out}$  в диапазоне от 0.01 до 0.1 строились пять шаговых регрессий на основе обучающей выборки с исключенным первым, вторым, ..., пятым блоком и вычислялась средняя корреляция по исключенным блокам. В результате наилучшим оказался диапазон [2.2, 2.6] Å и параметры  $f_{in} = 0.06$ ,  $f_{out} = 0.09$ , давшие в результате кросс-валидации максимальную среднюю корреляцию.

Данным значениям параметров соответствовала регрессия (настроенная уже на полной обучающей выборке):

$$\begin{aligned} \lg S = & -8.48 + 16.50d_{ASA_p} - 55.36c_{DPSA_3} + \\ & + 0.74VDWV - 0.75c_{ASA_H} + 2.93d_{ASA_H} + \\ & + MaxPA(0.45 - 0.63d_{ASA_p} + \\ & + 3.99c_{DPSA_3} + 1.24d_{PNSA_3}), \end{aligned} \quad (6)$$

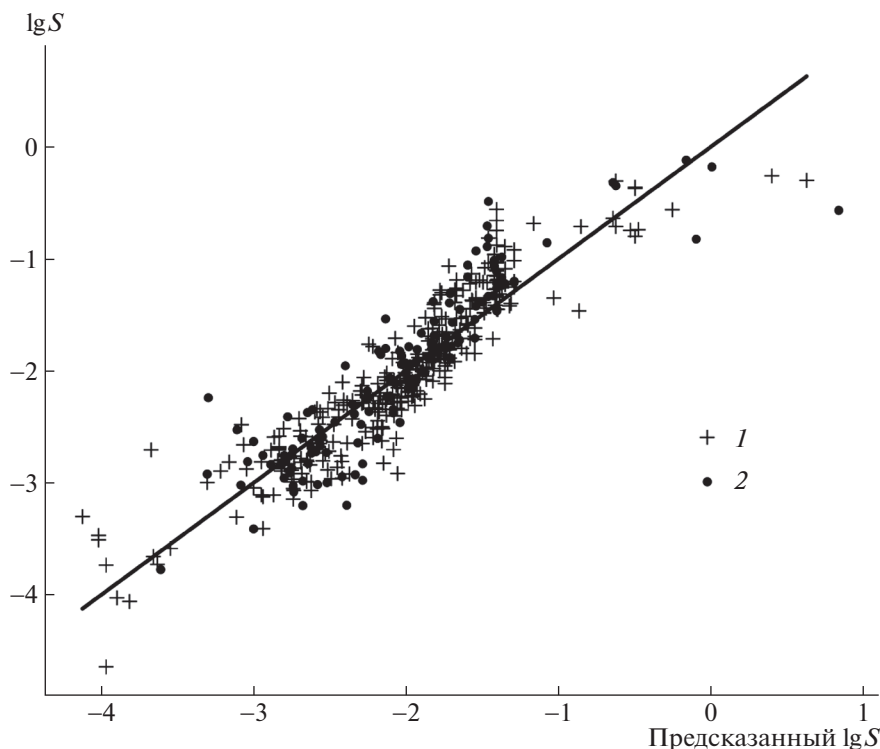
задействующая девять объясняющих переменных, в том числе, значимыми оказались шесть различных параметров полимера. (Обучающая выборка:  $N = 344$ ;  $R^2 = 0.85$ ; скорректированный коэффициент детерминации  $q^2 = 0.84$ ,  $RMSE = 0.11$ ,  $F = 207$ . Тестовая выборка:  $N = 134$ ,  $R^2 = 0.89$ ,  $RMSE = 0.11$ ). Значения  $\lg S$  в обучающей выборке, а также остатки распределены нормально.

Таким образом,

$$a(p) = -8.48 + 16.50d_{ASA_p} - 55.36c_{DPSA_3} + 0.74VDWV - 0.75c_{ASA_H} + 2.93d_{ASA_H} \quad (7)$$

$$b(p) = 0.45 - 0.63d_{ASA_p} + 3.99c_{DPSA_3} + 1.24d_{PNSA_3} \quad (8)$$

Переменные в формулах (7) и (8) следуют в порядке убывания значимости, при этом наиболее значимыми оказались  $d_{ASA_p}$  и  $c_{DPSA_3}$ , присутствующие, вдобавок, в обоих коэффициентах.



**Рис. 3.** Диаграмма рассеяния для предсказания  $\lg S$ ,  $\lg[\text{см}^3(\text{н.у.})/\text{см}^3(\text{см.рт.ст.})]$ . 1 – обучающая выборка, 2 – тестовая выборка.

Коэффициент детерминации регрессии составил 0.85 (скорректированный  $R^2 = 0.84$ ) на обучающей выборке (при корреляции 0.92) и 0.80 на тестовой выборке (при корреляции 0.89), что соответствует средней относительной ошибке предсказания  $S$  56% на обучающей выборке и 77% на тестовой выборке. В результате модифицированный алгоритм дает в среднем ту же точность

что и первая версия метода (70% на тестовой выборке), что считается неплохой точностью с учетом большой (до десятичного порядка) погрешности экспериментального измерения  $S$ .

На рис. 3 представлена диаграмма рассеяния для регрессии (6); в табл. 2 и 3 приведена средняя относительная ошибка предсказания  $S$  для раз-

**Таблица 2.** Распределение полимеров по классам в наборе данных и средняя относительная ошибка предсказания коэффициента растворимости  $S$

Класс полимеров	Обучающая выборка		Тестовая выборка	
	Количество измерений	Средняя относительная ошибка предсказания $S$ , %	Количество измерений	Средняя относительная ошибка предсказания $S$ , %
Полиацетилены	3	65	4	664
Полиакрилаты	25	59	7	36
Поликарбонаты	4	50	1	46
Полиамиды	241	49	89	57
Полинорборнены	4	285	2	130
Полистиролы	41	53	19	63
Виниловые полимеры	26	79	12	64
ИТОГО	344	56	134	77

**Таблица 3.** Газы-пенетранты в наборе данных и средняя относительная ошибка предсказания коэффициента растворимости  $S$ 

Газ-пенетрант	Обучающая выборка		Тестовая выборка	
	Количество измерений	Средняя относительная ошибка предсказания $S$ , %	Количество измерений	Средняя относительная ошибка предсказания $S$ , %
Ar	9	34	3	22
C <sub>2</sub> H <sub>2</sub>	1	19	—	—
C <sub>2</sub> H <sub>4</sub>	3	89	—	—
C <sub>2</sub> H <sub>6</sub>	2	191	—	—
C <sub>3</sub> H <sub>6</sub>	4	45	2	50
C <sub>3</sub> H <sub>8</sub>	4	54	2	1440
C <sub>4</sub> H <sub>10</sub>	1	748	1	53
C <sub>4</sub> H <sub>6</sub>	1	357	1	9
CH <sub>4</sub>	70	36	19	44
CO <sub>2</sub>	75	40	31	44
H <sub>2</sub>	24	46	10	65
He	8	85	3	248
Kr	3	149	1	8
N <sub>2</sub>	70	90	32	82
O <sub>2</sub>	67	38	27	31
SO <sub>2</sub>	1	28	—	—
Xe	1	20	—	—
Ne	—	—	1	45
ИТОГО	344	56	134	77

личных классов полимеров и газов-пенетрантов соответственно.

### ПРЕДСКАЗАНИЕ КОНСТАНТЫ ЗАКОНА ГЕНРИ

Как показано в предыдущем разделе, усовершенствованный подход [1] позволяет с удовлетворительной точностью предсказать коэффициенты растворимости около 20 газов в примерно 60 стеклообразных полимерах разной структуры. Между тем, в этом подходе почти полностью игнорируется влияние свободного объема на коэффициенты растворимости, хотя известно, что многие полимеры, в том числе представляющие интерес в качестве мембранных материалов, отличаются повышенным свободным объемом и большими коэффициентами растворимости, чем обычные стеклообразные полимеры [38–40].

Мерой свободного объема может служить ленгмюровская сорбционная емкость в модели двойной сорбции  $C'_H$  [32]. В рамках данной модели, по формуле (2) коэффициент растворимости при бесконечном разбавлении (низких давлениях

сорбата) может быть представлен как  $S = k_D + C'_H b$ , где  $k_D$  – коэффициент растворимости “популяции Генри” молекул сорбата, т.е. параметр, характеризующий растворимость в более плотных областях полимерной матрицы, подчиняющийся линейной изотерме Генри, а параметр  $b$  является константой равновесия для молекул сорбата в элементах свободного объема. Для высокопроницаемых полимеров выполняется неравенство  $C'_H b \gg k_D$ . Таким образом, параметр  $k_D$  не должен зависеть от свободного объема в полимере. В связи с этим представляет интерес применение описанного в статье подхода для предсказания константы закона Генри  $k_D$ .

Был собран набор экспериментальных значений  $k_D$  из 94 измерений. Набор построен на основе анализа литературы и включает в себя данные по растворимости 13 газов в 40 полимерах. В табл. 2 и 3 приведена статистика по классам полимеров и газам-пенетрантам. Проблемой является то, что эксперименты проводились при разной температуре от 278 до 308 К. В большинстве работ отсутствовала информация о температур-

**Таблица 4.** Распределение полимеров по классам в наборе данных и средняя относительная ошибка предсказания константы закона Генри  $k_D$

Класс полимеров	Обучающая выборка		Тестовая выборка	
	Количество измерений	Средняя относительная ошибка предсказания $k_D$ , %	Количество измерений	Средняя относительная ошибка предсказания $k_D$ , %
Другие гетероцепные полимеры	1	95	2	82
Полиацетилены	15	55	—	—
Полиакрилаты	4	45	—	—
Полиамидоимиды	4	31	—	—
Поликарбонаты	6	42	1	9
Полимеры на основе простых эфиров	3	86	2	21
Полиимиды	14	65	3	66
Полимеры на основе сложных эфиров	—	—	1	12
Полинорборнены	24	52	—	—
Полисульфоны	—	—	1	9
Виниловые полимеры	5	279	7	53
ИТОГО	76	69	17	47

ной зависимости  $k_D$ , и использованные в работе значения  $k_D$  являются усредненными в указанном температурном интервале.

К такому набору данных был применен описанный выше подход. Для каждого полимера из набора данных был построен короткий отрезок цепи размером от 200 до 700 атомов и рассчитано 6 конформаций средствами ChemAxon Instant JChem и PerkinElmer Chem3D версии 15.1.0.144. Для каждой конформации были вычислены значения геометрических индексов из табл. 1 для радиуса обкатки от 0 до 3 Å (за исключением  $VDWV$ , к которому понятие радиуса “обкатки” неприменимо). В связи с отсутствием экспериментальных величин плотности образцов значения индексов не нормализовались на плотность (что равнозначно предположению о средней плотности исследуемых стеклообразных полимеров 1 г/см<sup>3</sup>). В табл. 4 показано распределение числа образцов по классам полимеров, а в табл. 5 – по газам пентрантам. Отметим отсутствие среди исследуемых газов CO (как и для растворимости), а также CO<sub>2</sub>, который, предположительно, также демонстрирует индивидуальный характер.

Набор данных был разбит на обучающую (76 измерений) и тестовую (17 измерений) выборки. Предполагалось, что искомая регрессия имеет вид, аналогичный выражению (3):

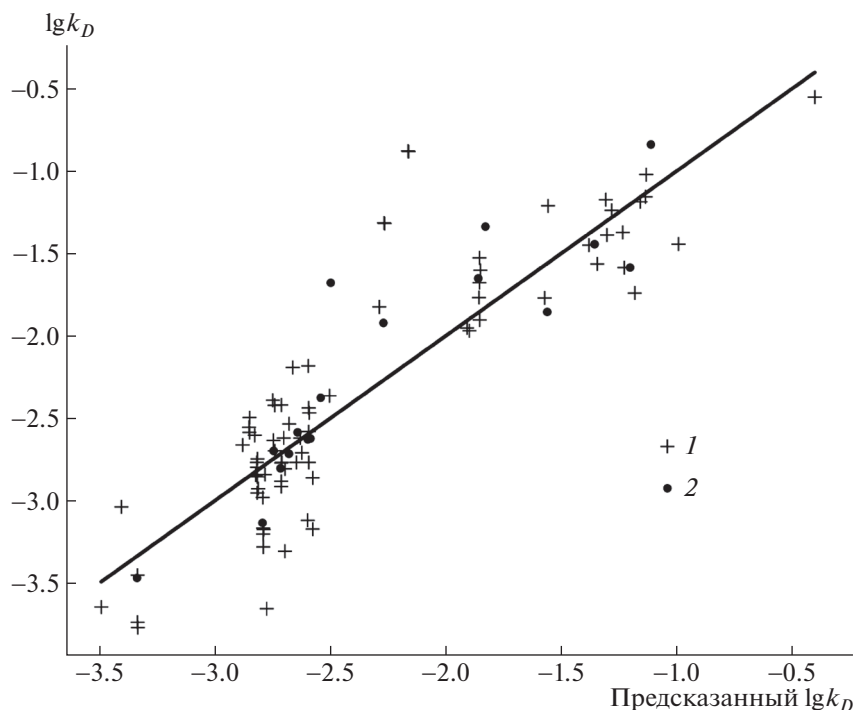
$$\lg k_D(g, p) = a(p) + b(p)MaxPA(g), \quad (9)$$

где коэффициенты  $a$  и  $b$  зависят от полимера. Как и ранее, они подбирались в виде взвешенной суммы коэффициентов линейной аппроксимации кривых  $ASA(R)$ , ...,  $PNSA_3(R)$  на том же что для растворимости  $S$  отрезке [2.2 Å; 2.6 Å]. Также в качестве объясняющей переменной использовался коррелянт свободного объема, ван-дер-ваальсов объем  $VDWV$  макромолекулы (для проверки высказанного выше предположения о том, что для предсказания  $k_D$  показатели свободного объема менее значимы).

В результате применения шаговой регрессии (с теми же параметрами что и в предыдущем разделе) была получена формула

$$\lg k_D = MaxPA(0.12 - 2.96d_{PPSA_3}) - 4.11, \quad (10)$$

задействующая единственную переменную, характеризующую полимер – это чувствительность индекса  $PNSA_3$  (площади отрицательно заряженной доступной поверхности с учетом величины частичного заряда) к радиусу “обкатки” (Обучающая выборка:  $N = 76$ ,  $R^2 = 0,79$ ; скорректированный коэффициент детерминации  $q^2 = 0.79$ ,  $RMSE = 0.12$ ,  $F = 155$ . Тестовая выборка:  $N = 17$ ,  $R^2 = 0.81$ ,  $RMSE = 0.09$ ). Значения  $\lg k_D$  в обучающей выборке, а также остатки распределены нормально.



**Рис. 4.** Диаграмма рассеяния для предсказания  $\lg k_D$ ,  $\lg[\text{см}^3(\text{н.у.})/\text{см}^3(\text{см.рт.ст.})]$ . 1 – обучающая выборка, 2 – тестовая выборка.

Меньшее число значимых переменных обусловлено небольшим объемом выборки, не позволяющим выделить тонкие эффекты. Отметим, что ван-дер-ваальсов объем оказался незначимой переменной; это подтверждает предположение о том, что константа закона Генри  $k_D$  в первом приближении не зависит от свободного объема полимерной матрицы.

Коэффициент детерминации регрессии составил 0.79 (скорректированный  $R^2 = 0.78$ ) на обучающей выборке (при корреляции 0.89) и 0.81 на тестовой выборке (при корреляции 0.9), что соответствует средней относительной ошибке предсказания  $k_D$  69% на обучающей выборке и 47% на тестовой выборке. Таким образом, несмотря на меньший объем выборки, отсутствие нормировки объясняющих переменных на экспериментальную плотность образца, а также на отсутствие нормализации значений  $k_D$  по температуре, удалось получить даже лучшую точность предсказания, чем для коэффициента растворимости  $S$ .

На рис. 4 приведена диаграмма рассеяния для регрессии (9), в табл. 4 и 5 – средняя относительная ошибка предсказания  $k_D$  для различных классов полимеров и газов-пенетрантов соответственно.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Изучение формулы (6) позволяет сделать несколько общих выводов относительно связи между геометрией полимерной цепи и коэффициентом растворимости  $S$ . Важен состав значимых характеристик полимера, а также знак и абсолютная величина коэффициентов при этих характеристиках, определяющие их влияние на процессы сорбции. Из формулы (3) легко видно, что поскольку для двух газов,  $g$  и  $g'$ ,  $\ln[S(p, g)/S(p, g')] = b(p)[\text{MaxPA}(g) - \text{MaxPA}(g')]$ , лишь коэффициент  $b(p)$  влияет на селективность растворимости.

Тогда из формулы (6) следует, что селективность растворимости определяется тремя значимыми параметрами – чувствительностью  $d_{ASA_p}$  площади поляризованной доступной поверхности к радиусу обкатки, чувствительностью  $d_{PNSA_3}$  площади отрицательно заряженной поверхности, а также площадью доступной поверхностью с учетом знака  $c_{DPSA_3}$  (все это при линеаризации кривых зависимостей индекса от радиуса обкатки на диапазоне 2.2–2.6 Å). Интересно, что веса всех трех показателей устройства таким образом, что они всегда вносят отрицательный вклад в значение коэффициента  $b(p)$ , который, тем не менее, всегда остается положительным (на рассматриваемом наборе данных он меняется от 0.1 до

**Таблица 5.** Газы-пенетранты в наборе данных и средняя относительная ошибка предсказания константы закона Генри  $k_D$

Газ-пенетрант	Обучающая выборка		Тестовая выборка	
	Количество измерений	Средняя относительная ошибка предсказания $k_D$ , %	Количество измерений	Средняя относительная ошибка предсказания $k_D$ , %
Ar	8	63	1	22
C <sub>2</sub> H <sub>4</sub>	1	66	1	68
C <sub>2</sub> H <sub>6</sub>	7	27	1	38
C <sub>3</sub> H <sub>6</sub>	7	54	2	35
C <sub>3</sub> H <sub>8</sub>	5	81	1	141
C <sub>4</sub> H <sub>10</sub>	2	49		
CH <sub>4</sub>	15	63	6	25
H <sub>2</sub>	2	50		
He	3	119	1	35
N <sub>2</sub>	15	106	2	65
N <sub>2</sub> O	1	55	1	96
O <sub>2</sub>	6	43		
Xe	4	84	1	55
ИТОГО	76	69	17	47

0.3 lg[см<sup>3</sup>(н.у.)/см<sup>3</sup>(см.рт.ст.)]/Å). Так, полиимиды имеют относительно большую чувствительность  $d_{ASA_p}$  поляризованной поверхности и соответственно меньшую селективность растворимости. Это довольно неожиданный эффект. Наоборот, ожидаемым является то, что переменные  $d_{ASA_p}$  и  $d_{PNSA_3}$ , влияющие также и на коэффициент  $a(p)$ , входят в формулу (7) с весами противоположного знака.

Стоит отметить, что ван-дер-ваальсов объем макромолекулы, коррелирует свободному объему, не оказался значимым фактором, влияющим на селективность растворимости. В то же время он оказался среди факторов, влияющих на коэффициент  $a(p)$  – “коэффициент растворимости при нулевом радиусе пенетранта”, влияющий на абсолютную величину коэффициента растворимости.

Интересно, что наилучший диапазон линеаризации [2.2; 2.6] Å оказался за пределами диапазона изменения эффективных (по Теплякову–Мирсу) радиусов пенетрантов, что может быть признаком того что с точки зрения процесса сорбции молекулы газа имеют больший эффективный радиус. Впрочем, потеря качества предсказания при использовании других диапазонов линеаризации невелика, и все они более или менее информативны.

Регрессия (10) задействует чувствительность индекса  $PNSA_3$  (площади отрицательно заряжен-

ной доступной поверхности с учетом величины частичного заряда) к радиусу “обкатки”. При этом увеличение чувствительности индекса (в выборке присутствуют в основном положительные значения  $d_{PNSA_3}$ ) приводит к уменьшению значения  $k_D$ .

Методика предсказания коэффициента растворимости и константы закона Генри на основе только компьютерных моделей конформаций макромолекул, описанная в настоящей статье, может быть использована при поиске и синтезе полимерных материалов с экстремальными транспортными характеристиками.

Для нового (в том числе еще не синтезированного) полимера в программной среде ChemAxon Instant JChem необходимо создать полимерную цепочку размером 200–600 атомов, рассчитать не менее шести конформаций средствами пакета Conformers Plugin среды Instant JChem или соответствующими средствами приложения ChemOffice Pro и вычислить геометрические индексы табл. 1. При этом достаточно ограничиться индексами, входящими в формулу (6) для радиусов обкатки от 2.2 до 2.6 Å и найти для них коэффициенты  $c_{ASA}$ ,  $d_{ASA}$ , ... наилучшей линейной аппроксимации по радиусу обкатки  $R$ .

После этого для вычисления предсказанного значения  $\lg S$  достаточно подставить вычисленные значения индексов  $c_{ASA}$ ,  $d_{ASA}$ , ...,  $VDWV$  в формулу (6) и использовать полученное значение

растворимости при отборе потенциально интересных для синтеза полимерных материалов с требуемыми транспортными характеристиками. Аналогично используется и формула для предсказания константы равновесия закона Генри  $k_D$ .

Исследование частично выполнено при финансовой поддержке Российского фонда фундаментальных исследований в рамках научного проекта 17-08-00164.

О.А. Милосердов выражает благодарность за финансовую поддержку Российскому фонду фундаментальных исследований в рамках научного проекта 18-37-00265.

Работа частично выполнена в рамках государственного задания ИНХС РАН.

### СПИСОК ЛИТЕРАТУРЫ

- Goubko M., Miloserdov O., Yampolskii Yu., Alentiev A., Ryzhikh V. // *J. Polym. Sci., Polym. Phys.* 2016. V. 55. № 3. P. 228.
- Baker R. // *W. Ind. Eng. Chem. Res.* 2002. V. 41. P. 1393.
- Bernardo P., Drioli E., Golemme G. // *Ind. Eng. Chem. Res.* 2009. V. 48. P. 4638.
- Yampolskii Yu. // *Macromolecules.* 2012. V. 45. P. 3298.
- Barrer R.M., Barrie J.A., Slater J. // *J. Polym. Sci.* 1958. V. 27. № 115. P. 177.
- Ferrara P., Apostolakis J., Caflisch A. // *Proteins Struct. Funct. Bioinform.* 2002. V. 46. № 1. P. 24.
- Lee B., Richards F. // *J. Molec. Biol.* 1971. V. 55. P. 379.
- Theodorou D. // *Materials Science of Membranes for Gas and Vapor Separation* / Ed. by Yu. Yampolskii, I. Pinnau, B.D. Freeman. Wiley: Chichester, 2006. P. 49.
- Fried J. // *Materials Science of Membranes for Gas and Vapor Separation* / Ed. By Yu. Yampolskii, I. Pinnau, B.D. Freeman. Wiley: Chichester, 2006. P. 95.
- Norman G.E., Filinov V.S. // *High Temp.* 1969. V. 7. P. 216.
- Neyertz S. // *Encyclopedia of Membrane Science and Technology* / Ed. By Eric M.V. Hoek, Volodymyr V. Tarabara. New York: Wiley, 2013. P. 1.
- Alder B.J., Wainwright T.E. // *J. Chem. Phys.* 1959. V. 31. P. 459.
- Gusev A.A., Suter U.W. // *J. Chem. Phys.* 1993. V. 99. P. 2228.
- Greenfield M.L., Theodorou D.N. // *Macromolecules* 2001. V. 34. P. 8541.
- Neyertz S., Brown D., Pandiyan S., A. van der Vegt N.F. // *Macromolecules.* 2008. V. 41. P. 2711–2721.
- Mazo M., Balabaev N., Alentiev A., Yampolskii Y. // *Macromolecules.* 2018. V. 51. № 4. P. 1398.
- Yampolskii Yu., Shishatskii S., Alentiev A., Loza K. // *J. Membr. Sci.* 1998. V. 149. P. 203.
- Ryzhikh V., Tsarev D., Alentiev A., Yampolskii Yu. // *J. Membr. Sci.* 2015. V. 487. P. 189.
- Bicerano J. *Prediction of Polymer Properties*, third edition / Ed. by D.E. Hudgin. New-York: Marcel Dekker, 2002.
- Аскадский А.А. // *Обзор. журн. по химии.* 2015. Т. 5. № 2. С. 83.
- Askadskii A.A. // *Computational Materials Science of Polymers.* Cambridge: Cambridge Intern. Sci. Publ. 2001.
- Ronova I.A., Rozhkov E.M., Alentiev A.Yu., Yampolskii Yu.P. // *Macromol. Theory Simul.* 2003. V. 12. № 6. P. 425.
- Ronova I.A., Sokolova E.A., Bruma M. // *J. Polym. Sci., Polym. Phys.* 2008. V. 46. P. 1868.
- Alentev A.Yu., Ronova I.A., Shchukin B.V., Yampolskii Yu.P. // *Polymer Science A.* 2007. V. 49. № 2. P. 217.
- Leay L., Siperstein F.R. // *Adsorption Science & Technology.* 2013. V. 31. № 1. P. 99.
- Yampolskii Yu., Wiley D., Maher C. // *J. Appl. Polym. Sci.* 2000. V. 76. P. 552–560.
- <https://docs.chemaxon.com/display/docs/Conformer+Plugin> (Дата обращения: 30.01.2019).
- <https://www.chemaxon.com> (Дата обращения: 30.01.2019).
- Alentiev A., Yampolskii Yu., Ryzhikh V., Tsarev D. // *Petroleum Chem.* 2013. V. 53. P. 554.
- <http://www.perkinelmer.com/Product/chemoffice-professional-chemofficepro> (Дата обращения: 30.01.2019).
- Materials science of membranes for gas and vapor separation* / Ed. by Yu. Yampolskii, I. Pinnau, B.D. Freeman. Wiley: Chichester, 2006.
- Yampolskii Yu., Wiley D., Maher C. // *J. Appl. Polym. Sci.* 2000. V. 76. P. 552–560.
- Todeschini R., Consonni V. *Handbook of Molecular Descriptors.* Weinheim: Wiley-VCH, 2000.
- Ямпольский Ю.П. // *Успехи химии.* 2007. Т. 76. № 1. С. 66–87.
- Yampolskii Y. et al. // *J. Membr. Sci.* 1998. T. 149. № 2. С. 203.
- <https://datascience.stackexchange.com/questions/24405/how-to-do-stepwise-regression-using-sklearn> (Дата обращения: 30.01.2019).
- Efroymsen M.A. // *Mathematical Methods for Digital Computers* / Ed. by A. Ralston, H.S. Wilf. Wiley: New York, 1960. P. 191.
- Hart K.E., Colina C.M. // *J. Membr. Sci.* 2014. V. 468. P. 259.
- Budd P.M., McKeown N.B., Ghanem B.S., Msayib K.J., Fritsch D., Starannikova L., Belov N., Sanfirova O., Yampolskii Yu., Shantarovich V. // *J. Membr. Sci.* 2008. V. 325. P. 851.
- Nagai K., Masuda T., Nakagawa T., Freeman B.D., Pinnau I. // *Progr. Polym. Sci.* 2001. V. 26. P. 721.
- Kanehashi S., Nagai K. // *J. Membr. Sci.* 2005. V. 253. P. 117.