

УДК 577.1, 577.3, 54.03

## Включение важнейших многочастичных взаимодействий в силовое поле AMBER и оптимизация энергетических параметров обновленного поля

©2015 Гарбузинский С.А.<sup>1</sup>, Маткаримов Б.Т.<sup>2</sup>, Финкельштейн А.В.\*<sup>1</sup>

<sup>1</sup>Институт белка Российской академии наук, Пуцино, Московская обл., Россия

<sup>2</sup>National Laboratory Astana, Nazarbayev University, Астана, Казахстан

**Аннотация.** Разработана новая функциональная форма для силового поля невалентных взаимодействий, учитывающая, наряду с традиционными членами, как наведенную парциальными зарядами поляризацию всех атомов, так и трехчастичные дисперсионные взаимодействия «атом – валентная связь». Соответствующие новые члены добавлены к стандартному силовому полю AMBER. В рамках этой единой функциональной формы силового поля оптимизированы, с использованием данных по молекулярным кристаллам, параметры всех типов невалентных взаимодействий. Показано заметное увеличение коэффициента корреляции между рассчитанной и экспериментально определенной энергией когезии молекул в молекулярных кристаллах.

**Ключевые слова:** невалентные атом-атомные взаимодействия, трехчастичные взаимодействия, поляризуемость, парциальные заряды, взаимодействие «атом-ковалентная связь», молекулярные кристаллы, энергия когезии.

### ВВЕДЕНИЕ

Компьютерное моделирование поведения макромолекул, например, белков требует очень высокой точности применяемых при этом потенциалов взаимодействий. Так, недавний прогресс в моделировании динамики белков [1] стал возможным только после очередного уточнения потенциалов ковалентных сил [2].

Стандартные потенциалы [3–7] невалентных взаимодействий учитывают только попарные атом-атомные взаимодействия, обычно в форме типа

$$U_{\text{noncov}}\{\mathbf{r}\} = \frac{1}{2} \sum_i \sum_{j \in (i,j)} \left[ \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right] + \frac{1}{2\epsilon} \sum_i \sum_{j \in (i,j)} \frac{q_i q_j}{r_{ij}}, \quad (1)$$

где  $\{\mathbf{r}\}$  – совокупность всех трехмерных координат атомов; сумма по  $i$  берется по всем атомам,  $(i, j)$  – совокупность атомов  $j$ , ковалентно не взаимодействующих с  $i$  (т.е. отделенных от  $i$  тремя и более валентными связями или принадлежащих другим молекулам [3, 7]);  $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ , где  $\mathbf{r}_i, \mathbf{r}_j$  – координаты атомов  $i, j$ ;  $r_{ij} = |\mathbf{r}_{ij}|$ ;  $A_{ij}, B_{ij}$  – параметры ван-дер-ваальсового взаимодействия;  $\epsilon$  – диэлектрическая проницаемость среды;  $q_i, q_j$  – парциальные заряды атомов; множители  $1/2$  перед суммами компенсируют то, что в них каждая пара атомов  $(i, j)$  учитывается дважды – как  $i, j$  и как  $j, i$ .

\*afinkel@vega.protres.ru

Однако недавно было показано, что многочастичные взаимодействия (дисперсионные, для взаимодействия пары валентно-связанных атомов с прочими атомами [8, 9] и электростатические взаимодействия двух заряженных атомов с поляризацией третьего [10, 11]) вовсе не являются несущественными. Поэтому разумно ввести потенциалы таких многочастичных взаимодействий (и соответствующие им силы) в какое-либо из хорошо зарекомендовавших себя «стандартных» силовых полей, разработанных с учетом только попарных атом-атомных потенциалов.

Недавно в ходе молекулярно-динамического моделирования с использованием силового поля AMBER [7] удалось проследить сворачивание развернутых цепей нескольких маленьких белков, показав, что они достигают (с точностью 1–5 Å) окрестностей своих нативных структур [1, 12]. Однако погрешность в оценке энергии денатурации тех же белков при этом достигала порой десятков ккал/моль, что показывает необходимость [12] усовершенствования даже такого хорошего силового поля как AMBER.

Поэтому в нашей работе именно AMBER (а точнее, AMBER99) [7, 13] выбран в качестве совершенствуемого силового поля.

В AMBER величины  $A_{ij}$ ,  $B_{ij}$  выражаются через  $\epsilon_i$  и  $\epsilon_j$  – максимальные энергии ван-дер-ваальсового притяжения атомов,  $i$  к  $i$  и  $j$  к  $j$  соответственно, и  $r_{0,i} + r_{0,j}$  – сумму ван-дер-ваальсовых радиусов атомов  $i$  и  $j$ :

$$A_{ij} = \sqrt{\epsilon_i \epsilon_j} \cdot (r_{0,i} + r_{0,j})^{12}, \quad (2)$$

$$B_{ij} = 2\sqrt{\epsilon_i \epsilon_j} \cdot (r_{0,i} + r_{0,j})^6. \quad (3)$$

Величины атомных парциальных зарядов  $q_i$  рассчитываются (как и в AMBER) с помощью процедуры RESP [13], а диэлектрическая проницаемость среды  $\epsilon$  в AMBER полагается равной 1 (как в вакууме).

Наша цель в данной работе – расширить силовое поле AMBER, включив в него важнейшие трехчастичные взаимодействия, и оптимизировать параметры невалентных взаимодействий для наилучшего согласования «новых» многочастичных взаимодействий со «старыми» двухчастичными.

## РАСШИРЕНИЕ СИЛОВОГО ПОЛЯ AMBER; РАСЧЕТ ЭНЕРГИЙ МНОГОЧАСТИЧНЫХ НЕВАЛЕНТНЫХ ВЗАИМОДЕЙСТВИЙ

### 1. Энергия силового поля, включающего трехчастичные невалентные взаимодействия

К взаимодействиям, описываемым силовым полем AMBER, мы добавляем следующие:

1) Взаимодействия зарядов с поляризуемыми ими атомами. Взаимодействия эти вытекают из классической электростатики и (при  $\epsilon = 1$ ) их энергия имеет вид [10, 11]:

$$-\frac{1}{2} \sum_i \alpha_i \mathbf{D}_i^2, \quad (4)$$

где

$$\mathbf{D}_i\{\mathbf{r}\} = \sum_{j \in (i,j)} \frac{q_j}{r_{ji}^2} \mathbf{n}_{ji} - \quad (5)$$

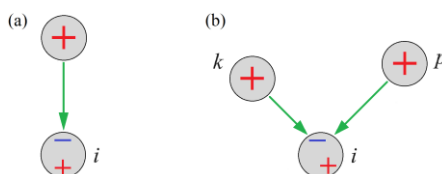
вектор электрической индукции, создаваемой в точке  $\mathbf{r}_i$  зарядами  $q$  других атомов (отделенных от  $i$  тремя и более валентными связями или принадлежащих другим

молекулам),  $\mathbf{n}_{ji} = (\mathbf{r}_i - \mathbf{r}_j)/r_{ij}$ , а  $\alpha_i$  – поляризуемость атома  $i$  (величины  $\alpha$  берутся из AMBER). Формулы (4), (5) соответствуют расширению обобщенной формулы Борна [14] на корпускулярную среду.

Хотя выражение

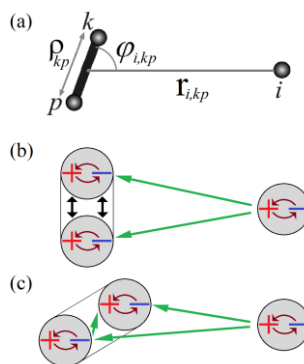
$$-\frac{1}{2} \sum_i \alpha_i \mathbf{D}_i^2 = -\frac{1}{2} \sum_i \alpha_i \cdot \left( \sum_{j \in (i,j)} \frac{q_j}{r_{ji}^2} \cdot \mathbf{n}_{ji} \right) \cdot \left( \sum_{k \in (i,k)} \frac{q_k}{r_{ki}^2} \cdot \mathbf{n}_{ki} \right) \quad (6)$$

определяется в основном трехчастичными взаимодействиями (см. формулу (5) и рисунок 1), время его вычисления растет пропорционально не третьей, а второй степени числа вовлеченных атомов (это следует из формул (4), (5)), то есть оно может быть вычислено достаточно оперативно.



**Рис. 1.** Электростатическое притяжение одного заряженного атома к вызванной им же поляризации третьего атома  $i$  (а) и притяжение двух заряженных атомов к вызванной ими совместно поляризации третьего атома  $i$  (б); сверх притяжения к этой, совместно вызванной зарядами  $k$  и  $p$ , поляризации атома  $i$ , каждый из зарядов  $k$ ,  $p$  притягивается к вызванной им же индивидуально поляризации атома  $i$ , как показано на панели (а).

2) Дисперсионные взаимодействия валентно-связанных атомных пар с прочими атомами. Это трехчастичные, иллюстрируемые рисунком 2, взаимодействия.



**Рис. 2.** Схема дисперсионного взаимодействия одного атома с парой других, валентно-связанных (т.е. очень близко друг от друга расположенных) атомов. **а)** Геометрические параметры, от которых зависит энергия трехчастичного дисперсионного взаимодействия одного атома  $i$  с парой валентно-связанных атомов  $k$  и  $p$ . **б)** Иллюстрация причины повышения энергии трехчастичного взаимодействия при перпендикулярном расположении связанных валентно-связанных атомов к линии, соединяющей их валентную связь с еще одним атомом  $i$ : отталкивание осциллирующих поляризаций (+ -), наводимых осциллирующей поляризацией (+ -) атома  $i$  на каждый из валентно-связанных атомов. **в)** Иллюстрация причины понижения энергии трехчастичного взаимодействия при расположении связанных атомов вдоль линии, соединяющей их с еще одним атомом  $i$ : притяжение осциллирующих поляризаций (+ -), наводимых осциллирующей поляризацией (+ -) атома  $i$  на каждый из валентно-связанных атомов.

Энергия этих взаимодействий имеет вид [8, 9, 15]

$$+ \frac{1}{2} \sum_i \sum_{k \in (i,k)} \sum_{\substack{p \in [k,p], \\ p \in (i,p)}} \frac{X_{ikp}}{\rho_{kp}^3 r_{i,kp}^6} \cdot (1 - 3 \cos^2 \varphi_{i,kp}) \quad (7)$$

Здесь  $[k, p]$  означает пару валентно-связанных атомов;  $\rho_{kp}$  – длина валентной связи  $k$ - $p$ , которую, ввиду жесткости валентной связи, можно считать неизменной;  $r_{i,kp} = |(\mathbf{r}_k + \mathbf{r}_p)/2 - \mathbf{r}_i|$  – расстояние от центра валентной связи  $k$ - $p$  до атома  $i$  (рис. 2);

$$\cos \varphi_{i,kp} = \frac{((\mathbf{r}_k + \mathbf{r}_p)/2 - \mathbf{r}_i) \cdot (\mathbf{r}_p - \mathbf{r}_k)}{|(\mathbf{r}_k + \mathbf{r}_p)/2 - \mathbf{r}_i| \cdot |\mathbf{r}_p - \mathbf{r}_k|}; \quad X_{ikp} - \text{энергетический параметр трехчастичного}$$

взаимодействия атомов  $i, k, p$ , определение величины которого дано ниже.

Число таких трехчастичных взаимодействий тоже растет лишь квадратично с ростом числа атомов, т.к. число валентных связей пропорционально числу атомов.

Согласно теории [10], рассматривающей дисперсионные взаимодействия как взаимодействия квантовых осцилляторов,

$$X_{ikp} = \frac{3\hbar(\omega_i + \omega_k + \omega_p)}{2} \cdot \gamma_{ik} \gamma_{kp} \gamma_{pi}, \quad (8)$$

где  $\gamma_{ik} = \frac{\sqrt{\omega_i \omega_k \alpha_i \alpha_k}}{\omega_i + \omega_k}$ ,  $\omega_i, \omega_k, \omega_p$  – частоты колебаний квантовых осцилляторов,

моделирующих соответствующие атомы,  $\alpha_i, \alpha_k, \alpha_p$  – их поляризуемости, а  $\hbar$  – постоянная Планка.

Величина  $B_{ik}$ , в той же модели квантовых осцилляторов, есть  $\frac{3}{2} \hbar \frac{\omega_i \omega_k \alpha_i \alpha_k}{\omega_i + \omega_k}$ . Хотя в стандартном силовом поле AMBER частот  $\omega$  нет,  $X_{ikp}$  приближенно выражается через величины  $B$  и  $\alpha$ , используя примерные равенства  $(\omega_i + \omega_k)/2 \approx \sqrt{\omega_i \omega_k}$ ,  $(\omega_i + \omega_k + \omega_p)/3 \approx \sqrt[3]{\omega_i \omega_k \omega_p}$ :

$$X_{ikp} \approx \frac{3}{4} \cdot \sqrt[3]{B_{ik} B_{kp} B_{pi} \alpha_i \alpha_k \alpha_p}. \quad (9)$$

При молекулярно-динамических расчетах больших систем (белков с окружающим их растворителем и т.д.) практически всегда используются гладкие «обрезающие функции», позволяющие гладко обнулить на больших расстояниях как энергии атом-атомных взаимодействий, так и соответствующие им силы, и таким образом избавиться от множества несущественных расчетов. Само по себе силовое поле AMBER не содержит таких «обрезающих функций», они вводятся каждым исследователем самостоятельно.

В используемой нами программе PUMA [16, 17] применяются следующие обрезающие функции:

$$f_B(r) = f_A(r) = \begin{cases} 1, & r \leq R_{on} \\ \frac{(R_{off}^2 - r^2)^2 (R_{off}^2 - 3R_{on}^2 + 2r^2)}{(R_{off}^2 - R_{on}^2)^3}, & R_{on} < r < R_{off} \\ 0, & r \geq R_{off} \end{cases}, \quad (10)$$

$$f_{qq}(r) = \begin{cases} \left(1 - \frac{r}{R_q}\right)^2, & r < R_q; \\ 0, & r \geq R_q \end{cases}$$

$f_B(r)$ ,  $f_A(r)$  с  $R_{on} = 9,0 \text{ \AA}$ ,  $R_{off} = 10,5 \text{ \AA}$  применяются в ван-дер-ваальсовом притяжении и отталкивании,  $f_{qq}(r)$  с  $R_q = 10,5 \text{ \AA}$  – в электростатическом взаимодействии.

Те же обрезаящие функции применяются и в данной работе. В результате, энергия всех (двух- и трехчастичных) невалентных взаимодействий приобретает вид

$$U_{\text{calc}}\{\mathbf{r}\} = \frac{\beta}{2} \sum_i \sum_{j \in (i,j)} \left[ A_{ij} \frac{f_A(r_{ij})}{r_{ij}^{12}} - B_{ij} \frac{f_B(r_{ij})}{r_{ij}^6} \right] + \frac{1}{2\epsilon} \sum_i \sum_{j \in (i,j)} \frac{q_i q_j}{r_{ij}} f_{qq}(r_{ij}) - \frac{\gamma}{2} \sum_i \alpha_i \mathbf{D}_i^2 + \\ + \frac{\eta}{2} \sum_i \sum_{k \in (i,k)} \sum_{\substack{p \in [k,p], \\ p \in (i,p)}} \frac{X_{ikp}}{\rho_{kp}^3} \cdot \frac{f_X(r_{i,kp})}{r_{i,kp}^6} \cdot (1 - 3 \cos^2 \varphi_{i,kp}), \quad (11)$$

где

$$\mathbf{D}_i\{\mathbf{r}\} = \sum_{j \in (i,j)} q_j \frac{f_D(r)}{r_{ji}^2} \mathbf{n}_{ji}; \quad (12)$$

при этом  $f_X(r) = f_B(r) = f_A(r)$ ,  $f_D(r) = f_{qq}(r)$ , а множители  $\beta$ ,  $1/\epsilon$ ,  $\gamma$ ,  $\eta$  играют роль оптимизируемых параметров, применяемых для того, чтобы наилучшим образом согласовать новые трехчастичные потенциалы со старыми двухчастичными.

## 2. Оптимизация параметров силового поля невалентных взаимодействий с помощью расчетов молекулярных кристаллов

Цель оптимизации – минимизировать разницу между вычисляемой и экспериментально найденной энергией невалентных взаимодействий (энергией когезии) молекул в кристаллах (считая рассматриваемые молекулы жесткими, мы пренебрегаем изменениями их внутренних энергий при сублимации):

$$U_{\text{calc}}^{(\xi)}\{\mathbf{r}\} \Rightarrow U_{\text{exp}}^{(\xi)} \text{ для всех рассматриваемых кристаллов } \xi, \quad (13)$$

или более конкретно

$$\sum_{\xi} \left( \frac{U_{\text{calc}}^{(\xi)}\{\mathbf{r}\}}{U_{\text{exp}}^{(\xi)}} - 1 \right)^2 \Rightarrow 0. \quad (14)$$

Здесь

$$U_{\text{exp}}^{(\xi)} = -\Delta H_{\text{subl}}^{(\xi)} - kT(2 + n_{\text{rot}}^{(\xi)} / 2) - \quad (15)$$

энергия взаимодействий молекулы в кристалле [10, 11], причем  $\Delta H_{\text{subl}}^{(\xi)} > 0$  – энергия сублимации кристалла  $\xi$  в расчете на одну его молекулу,  $n_{\text{rot}}^{(\xi)}$  – число внутренних свободных вращательных степеней свободы молекулы в газе, а слагаемое 2 в скобках вычисляется из числа вращательных и трансляционных степеней свободы молекулы как целого в газе, которое равно 6 ( $2 = 6 / 2 - 1$ ).

Для целей оптимизации удобно записать вычисленную для каждого кристалла  $\xi$  (точнее, для одной молекулы этого кристалла) энергию в виде

$$U_{\text{calc}}^{(\xi)}\{\mathbf{r}\} = \sum_{s=1}^4 \lambda_s U_s^{(\xi)}; \quad (16)$$

при этом  $\lambda_1 \equiv \beta$ ,  $\lambda_2 \equiv \frac{1}{\varepsilon}$ ,  $\lambda_3 \equiv \gamma$ ,  $\lambda_4 \equiv \eta$  – искомые оптимизирующие коэффициенты.

Величины  $U_s^{(\xi)}$  (т.е. энергии, соответствующие единичным коэффициентам  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ ) вычисляются по формулам:

$$U_1^{(\xi)} = \frac{1}{2} \sum_{i \in \text{mol.M}} \sum_{j \notin \text{mol.M}} \left[ A_{ij} \frac{f_A(r_{ij})}{r_{ij}^{12}} - B_{ij} \frac{f_B(r_{ij})}{r_{ij}^6} \right]; \quad (17)$$

$$U_2^{(\xi)} = \frac{1}{2} \sum_{i \in \text{mol.M}} \sum_{j \notin \text{mol.M}} \left[ \frac{q_i q_j}{r_{ij}} f_{qq}(r_{ij}) \right]; \quad (18)$$

$$U_3^{(\xi)} = - \sum_{i \in \text{mol.M}} \left[ \alpha_i \left( \frac{1}{2} \mathbf{D}_i^2 - \frac{1}{2} {}^0 \mathbf{D}_i^2 \right) \right], \quad (19)$$

где

$$\mathbf{D}_{i \in \text{mol.M}} = \sum_{j \in (i,j)} q_j \frac{f_D(r)}{r_{ji}^2} \mathbf{n}_{ji}, \quad {}^0 \mathbf{D}_{i \in \text{mol.M}} = \sum_{\substack{j \in \text{mol.M} \\ j \in (i,j)}} q_j \frac{f_D(r)}{r_{ji}^2} \mathbf{n}_{ji} \quad (\text{см. (12)});$$

$$U_4^{(\xi)} = \sum_{i \in \text{mol.M}} \frac{1}{2} \sum_{k \notin \text{mol.M}} \sum_{\substack{p \notin \text{mol.M} \\ p \in [k,p]}} \left[ \frac{X_{ikp}}{\rho_{kp}^3} \cdot \frac{f_X(r_{i,kp})}{r_{i,kp}^6} \cdot (1 - 3 \cos^2 \varphi_{i,kp}) \right]. \quad (20)$$

Здесь  $M$  – та молекула кристалла  $\xi$ , для которой ведется расчет энергии.

Минимизируемая функция невязки определяется как

$$\Phi = \sum_{\xi} \left( \frac{\sum_s \lambda_s U_s^{(\xi)} - U_{\text{exp}}^{(\xi)}}{U_{\text{exp}}^{(\xi)}} \right)^2. \quad (21)$$

В результате минимизации выражений (14), (16) относительно коэффициентов  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  получаем четыре уравнения типа  $\frac{\partial}{\partial \lambda_i} \Phi = 0$  (для  $i = 1, 2, 3, 4$ ), то есть:

$$\sum_{s=1}^4 \lambda_s \left[ \sum_{\xi} \frac{U_s^{(\xi)} U_i^{(\xi)}}{[U_{\text{exp}}^{(\xi)}]^2} \right] = \sum_{\xi} \frac{U_i^{(\xi)}}{U_{\text{exp}}^{(\xi)}} \quad (\text{при } i = 1, 2, 3, 4). \quad (22)$$

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Оптимизация параметров силового поля невалентных взаимодействий с помощью расчетов молекулярных кристаллов

Расчеты энергий с помощью нового силового поля и оптимизация параметров силового поля была сделаны на базе данных [11], содержащей 61 кристалл (состоящие из одного вида молекул каждый) небольших органических соединений, а также льда, для которых известны как пространственные структуры молекул в кристалле (Cambridge Structural Database, CSD) [18], так и экспериментальные данные по энтальпии сублимации этих

кристаллов (ссылки на соответствующие экспериментальные данные даны в работе [11]). Из пространственных структур мы вычисляем энергии каждого из 4 типов взаимодействий, а экспериментальные данные по сублимации кристаллов используются для оптимизации коэффициентов при вычислении суммарной энергии невалентных взаимодействий всех рассмотренных типов. Список использованных кристаллов и необходимые нам для расчётов данные для этих кристаллов приведены в таблице 1.

**Таблица 1.** База данных по молекулам, используемым для оптимизации параметров силового поля

	Название молекулы	Брутто-формула	Код в CSD [18]	$\Delta H_{\text{subl}}^{(\xi)}$ , ккал/моль	$n_{\text{rot}}^{(\xi)}$	$U_{\text{exp}}^{(\xi)}$ , ккал/моль
1	1,4-диоксан	C <sub>4</sub> H <sub>8</sub> O <sub>2</sub>	CUKCIU01	11,81	0	-13,01
2	1 <i>H</i> -бензимидазол	C <sub>7</sub> H <sub>6</sub> N <sub>2</sub>	BZDMAZ02	23,52	0	-24,72
3	1 <i>H</i> -индазол	C <sub>7</sub> H <sub>6</sub> N <sub>2</sub>	INDAZL	21,73	0	-22,93
4	2,3-ксиленол	C <sub>8</sub> H <sub>10</sub> O <sub>1</sub>	DIMPHE12	20,06	2	-22,00
5	2,5-ксиленол	C <sub>8</sub> H <sub>10</sub> O <sub>1</sub>	DMPHOL11	20,3	2	-22,09
6	2,6-ксиленол	C <sub>8</sub> H <sub>10</sub> O <sub>1</sub>	DMEPOL10	18,06	2	-19,77
7	2-нитрофенол	C <sub>6</sub> H <sub>5</sub> N <sub>1</sub> O <sub>3</sub>	ONITPH	17,4	0	-18,69
8	3,4-ксиленол	C <sub>8</sub> H <sub>10</sub> O <sub>1</sub>	DPHNOL10 <sup>a</sup>	20,47	2	-22,26
9	аценафтен	C <sub>12</sub> H <sub>10</sub>	ACENAP03	20,12	0	-21,31
10	ацетон	C <sub>3</sub> H <sub>6</sub> O <sub>1</sub>	HIXHIF01	9,52	2	-11,30
11	ацетофенон	C <sub>8</sub> H <sub>8</sub> O <sub>1</sub>	ACETPH	16,98	1	-18,76
12	анилин	C <sub>6</sub> H <sub>7</sub> N <sub>1</sub>	BAZGOY	15,58	0	-16,78
13	антрацен	C <sub>14</sub> H <sub>10</sub>	ANTCEN	24,43	0	-25,63
14	бензол	C <sub>6</sub> H <sub>6</sub>	BENZEN	10,65	0	-11,84
15	дифенил	C <sub>12</sub> H <sub>10</sub>	BIPHEN04	19,15	0	-20,34
16	бутан	C <sub>4</sub> H <sub>10</sub>	DUCKOB04	7,59	0	-8,78
17	дибензотиофен	C <sub>12</sub> H <sub>8</sub> S <sub>1</sub>	DBZTHP	20,33	0	-21,52
18	10 <i>H</i> -фенотиазин	C <sub>12</sub> H <sub>9</sub> N <sub>1</sub> S <sub>1</sub>	PHESAZ01	20,3	0	-27,82
19	диметилсульфид	C <sub>2</sub> H <sub>6</sub> S <sub>1</sub>	PADQAN	9,45	0	-10,64
20	диметилдисульфид	C <sub>2</sub> H <sub>6</sub> S <sub>2</sub>	DEMQIW	12,9	2	-14,69
21	капролактам	C <sub>6</sub> H <sub>11</sub> N <sub>1</sub> O <sub>1</sub>	CAPLAC	20,37	0	-21,57
22	карбазол	C <sub>12</sub> H <sub>9</sub> N <sub>1</sub>	CRBZOL03	24	0	-25,20
23	циклогексан	C <sub>6</sub> H <sub>12</sub>	CYCHEX	8,92	0	-10,11
24	циклопропан	C <sub>3</sub> H <sub>6</sub>	QQQCIS01	6,27	0	-7,46
25	декан	C <sub>10</sub> H <sub>22</sub>	QQQFBG01	20,04	0	-20,63
26	дибензофуран	C <sub>12</sub> H <sub>8</sub> O <sub>1</sub>	DBZFUL11	19,22	0	-20,41
27	диэтиловый эфир	C <sub>4</sub> H <sub>10</sub> O <sub>1</sub>	DETYLE	9,79	0	-10,98
28	диметилоксалат	C <sub>4</sub> H <sub>6</sub> O <sub>4</sub>	DMEOXA01	17,9	0	-19,95
29	этанол	C <sub>2</sub> H <sub>6</sub> O <sub>1</sub>	ETANOL	12,85	0	-14,05
30	флуорантен	C <sub>16</sub> H <sub>10</sub>	FLUANT02	21,95	0	-24,89
31	флуорен	C <sub>13</sub> H <sub>10</sub>	FLUREN01	20,46	0	-21,66
32	гептан	C <sub>7</sub> H <sub>16</sub>	HEPTAN01	13,41	0	-14,60
33	гексан	C <sub>6</sub> H <sub>14</sub>	HEXANE01	11,6	0	-12,79
34	изохинолин	C <sub>9</sub> H <sub>7</sub> N <sub>1</sub>	BERXOL <sup>b</sup>	19,2	0	-18,80
35	метиламин	C <sub>1</sub> H <sub>5</sub> N <sub>1</sub>	METAMI	8,77	1	-10,25
36	м-нитроанилин	C <sub>6</sub> H <sub>6</sub> N <sub>2</sub> O <sub>2</sub>	MNIANL03	23,49	0	-24,25
37	нафталин	C <sub>10</sub> H <sub>8</sub>	NAPHTA10	17,24	0	-18,44
38	нитробензол	C <sub>6</sub> H <sub>5</sub> N <sub>1</sub> O <sub>2</sub>	NITRBE01	16,6	0	-17,79
39	нонан	C <sub>9</sub> H <sub>20</sub>	QQQFAY01	15,68	0	-18,25
40	о-крезол	C <sub>7</sub> H <sub>8</sub> O <sub>1</sub>	OCRSOL <sup>b</sup>	18,01	1	-19,50

**Продолжение таблицы 1.** База данных по молекулам, используемым для оптимизации параметров силового поля

	Название молекулы	Брутто-формула	Код в CSD [18]	$\Delta H_{\text{subl}}^{(\xi)}$ , ккал/моль	$n_{\text{rot}}^{(\xi)}$	$U_{\text{exp}}^{(\xi)}$ , ккал/моль
41	октан	C <sub>8</sub> H <sub>18</sub>	OCTANE10	15,85	0	-17,05
42	<i>o</i> -нитроанилин	C <sub>6</sub> H <sub>6</sub> N <sub>2</sub> O <sub>2</sub>	ONITAN	20,47	0	-22,46
43	<i>n</i> -бензохинон	C <sub>6</sub> H <sub>4</sub> O <sub>2</sub>	BNZQUI03	16,4	0	-17,13
44	<i>n</i> -крезол	C <sub>7</sub> H <sub>8</sub> O <sub>1</sub>	CRESOL01 <sup>†</sup>	17,65	1	-19,14
45	пентан	C <sub>5</sub> H <sub>12</sub>	PENTAN10	9,55	0	-10,74
46	фенантрен	C <sub>14</sub> H <sub>10</sub>	PHENAN08	21,45	0	-22,65
47	фенол	C <sub>6</sub> H <sub>6</sub> O <sub>1</sub>	PHENOL03	16,53	0	-17,72
48	<i>n</i> -гидрохинон	C <sub>6</sub> H <sub>6</sub> O <sub>2</sub>	HYQUIN05	23,63	0	-23,68
49	<i>n</i> -нитроанилин	C <sub>6</sub> H <sub>6</sub> N <sub>2</sub> O <sub>2</sub>	NANIL02	24,05	0	-25,25
50	<i>n</i> -нитротолуол	C <sub>7</sub> H <sub>7</sub> N <sub>1</sub> O <sub>2</sub>	NITOLU	18,89	1	-20,39
51	<i>n</i> -ксилол	C <sub>8</sub> H <sub>10</sub>	ZZZITY01	14,32	2	-16,10
52	пирен	C <sub>16</sub> H <sub>10</sub>	PYRENE02	21,78	0	-24,73
53	пиридин	C <sub>5</sub> H <sub>5</sub> N <sub>1</sub>	PYRDNA01	12,52	0	-13,72
54	пиррол	C <sub>4</sub> H <sub>5</sub> N <sub>1</sub>	RUVQII	13,22	0	-14,42
55	янтарный ангидрид	C <sub>4</sub> H <sub>4</sub> O <sub>3</sub>	SUCANH13	19,37	0	-20,57
56	тетрагидрофуран	C <sub>4</sub> H <sub>8</sub> O <sub>1</sub>	BUNJAV01	11,18	0	-12,37
57	толуол	C <sub>7</sub> H <sub>8</sub>	TOLUEN	11,07	1	-11,78
58	<i>транс</i> -стильбен	C <sub>14</sub> H <sub>12</sub>	TSTILB10 <sup>‡</sup>	24,07	0	-25,27
59	триметиламин	C <sub>3</sub> H <sub>9</sub> N <sub>1</sub>	CEKGUU01	8,27	2	-10,35
60	триоксан	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub>	TROXAN11	13,4	0	-14,59
61	вода (лёд)	H <sub>2</sub> O <sub>1</sub>	- <sup>с</sup>	12,24	0	-13,43

<sup>а</sup> Для 3,4-ксиленола 3 атома водорода, отсутствующие в пространственной структуре его кристалла, добавлены (в работе [11]), основываясь на ковалентных параметрах использованного в работе [11] силового поля ENCAD [19].

<sup>б</sup> Для изохинолина была выбрана (в работе [11]) одна из двух одинаково населенных конформаций.

<sup>в</sup> В структуре кристалла *o*-крезола отсутствуют все атомы водорода; они были добавлены (в работе [11]), основываясь на ковалентных параметрах силового поля ENCAD.

<sup>г</sup> Для *n*-крезола 2 атома водорода, отсутствующих в пространственной структуре кристалла, добавлены (в работе [11]) согласно данным из оригинальной работы [20].

<sup>д</sup> Для *транс*-стильбена из двух встречающихся в кристалле конформаций была выбрана (в работе [11]) та, встречаемость которой составляет 94 %.

<sup>е</sup> Структура кристалла H<sub>2</sub>O (лед I<sub>h</sub>), отсутствующая в CSD, была однозначно восстановлена (в работе [11]) согласно <http://www.ibiblio.org/water/wsn-archive/msg00037.html>.

Примечание:  $\Delta H_{\text{subl}}^{(\xi)}$  – энергия сублимации кристалла  $\xi$  в расчете на одну его молекулу (измеренная при 25 °C или экстраполированная к 25 °C),  $n_{\text{rot}}^{(\xi)}$  – оцененное в работе [11] число внутренних свободных вращательных степеней свободы молекулы в газе (число торсионных углов с потенциальным барьером менее  $RT = 0,6$  ккал/моль согласно силовому полю ENCAD [19]. В приложении к [11] (на сайте <http://pubs.acs.org>) есть все необходимые для данной работы данные, структурные (см. также на <http://phys.protres.ru/resources/FFS/A1en.doc>) и термодинамические (см. также на <http://phys.protres.ru/resources/FFS/A2.pdf> с дополнениями на <http://phys.protres.ru/resources/FFS/Addition%20to%20A2.pdf>).

Каждому атому в молекулах мы приписали тип согласно типизации силового поля AMBER99 [7]. Типы атомов, имеющих в нашей базе; количество кристаллов, имеющих атомы данного типа, и параметры невалентных взаимодействий силового поля AMBER99 для каждого из этих типов представлены в таблице 2.

Заряды для каждого атома, рассчитанные с помощью процедуры RESP [13], взяты из предыдущей работы [11].

**Таблица 2.** Типы атомов, представленные в нашей базе данных, и параметры невалентных взаимодействий для этих типов атомов (согласно AMBER99)

Тип атома (AMBER99)	Описание	$\Xi$ , число кристаллов с этим типом атомов в нашей базе данных	Параметры ван-дер-ваальсова взаимодействия		$\alpha$ , $\text{\AA}^3$
			$r_{0,i}$ , $\text{\AA}$	$\epsilon_i$ , ккал/моль	
<b>Водород</b>					
H	додород, присоединённый к атому азота	11	0,6000	0,0157	0,161
HO	Водород гидроксильной группы	10	0,0000	0,0000	0,135
HC	Алифатический водород, присоединённый к углероду без электроотрицательных групп	28	1,4870	0,0157	0,135
H1	Алифатический водород, присоединённый к углероду с 1 электроотрицательной группой	7	1,3870	0,0157	0,135
H2	Алифатический водород, присоединённый к углероду с 2 электроотрицательными группами	1	1,2870	0,0157	0,135
HA	Ароматический водород, присоединённый к углероду без электроотрицательных групп	38	1,4590	0,0150	0,167
H4	Ароматический водород, присоединённый к углероду с 1 электроотрицательной группой	3	1,4090	0,0150	0,167
H5	Ароматический водород, присоединённый к углероду с 2 электроотрицательными группами	1	1,3590	0,0150	0,167
HW	Водород в молекулах воды	1	0,0000	0,0000	0
<b>Кислород</b>					
O	Кислород карбонильной группы	12	1,6612	0,2100	0,434
OW	Кислород в молекулах воды	1	1,7683	0,1520	0
OH	Кислород в карбоксильных группах	10	1,7210	0,2104	0,465
OS	Кислород в эфирных группах	7	1,6837	0,1700	0,465
<b>Углерод</b>					
C*	$sp^2$ -гибридизованный ароматический углерод	38	1,9080	0,0860	0,360
CT	$sp^3$ -гибридизованный алифатический углерод	30	1,9080	0,1094	0,878
C	$sp^2$ -гибридизованный углерод карбонильной группы	6	1,9080	0,0860	0,616
<b>Азот</b>					
N	$sp^2$ -гибридизованный азот	9	1,8240	0,1700	0,530
N3	$sp^3$ -гибридизованный азот	9	1,8240	0,1700	0,530
<b>Сера</b>					
S	Сера	4	2,0000	0,2500	2,900

Примечания:  $\Xi$  – количество кристаллов в нашей базе, имеющих данный тип атомов;  $r_0$  (в  $\text{\AA}$ ) и  $\epsilon$  (в ккал/моль) – параметры ван-дер-ваальсова взаимодействия в AMBER99;  $\alpha$  – поляризуемость ( $\text{\AA}^3$ ) по AMBER99.

Для одной молекулы в составе каждого кристалла были рассчитаны по формулам (17)–(20) величины  $U$  энергий невалентных взаимодействий с единичными коэффициентами  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ . Они были рассчитаны как для традиционно используемых в силовых полях парных (ван-дер-ваальсовых и электростатических) взаимодействий (члены  $U_1^{(\xi)}$  и  $U_2^{(\xi)}$  соответственно), так и новых многочастичных взаимодействий (член  $U_3^{(\xi)}$  для поляризационного и член  $U_4^{(\xi)}$  для дисперсионного).

Далее с помощью вышеописанных методов оптимизации были вычислены оптимальные значения всех коэффициентов ( $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ ), и с учётом этих значений для каждого кристалла вычислена суммарная энергия невалентных взаимодействий. Коэффициенты были оптимизированы с использованием экспериментально измеренной энергии сублимации. Оптимизация путём минимизации невязки дала следующие оптимальные значения коэффициентов:

$$\begin{aligned}\lambda_1 &= \beta = 0,92; \\ \lambda_2 &= \frac{1}{\varepsilon} = 0,90; \\ \lambda_3 &= \gamma = 2,90; \\ \lambda_4 &= \eta = -1,05.\end{aligned}$$

Коэффициент корреляции предсказанных энергий с экспериментальными при этом оптимальном значении коэффициентов  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  составил 89,9 % (рис. 3), в то время как без оптимизации он составлял 86,5 % при учёте только парных взаимодействий, как в исходном стандартном силовом поле AMBER (без оптимизации при  $\beta = \frac{1}{\varepsilon} = 1$  и  $\lambda_3 = \gamma = \lambda_4 = \eta = 0$ ).

Выше мы рассматривали оптимизацию добавления к стандартному силовому полю AMBER обоих новых членов, описывающих многочастичные (а) поляризационные и (б) дисперсионные взаимодействия. При добавлении к стандартному силовому полю AMBER только многочастичных поляризационных взаимодействий (т.е. при оптимизации  $\lambda_1, \lambda_2, \lambda_3$  при  $\lambda_4 \equiv 0$ ), получаем следующие оптимальные значения коэффициентов:

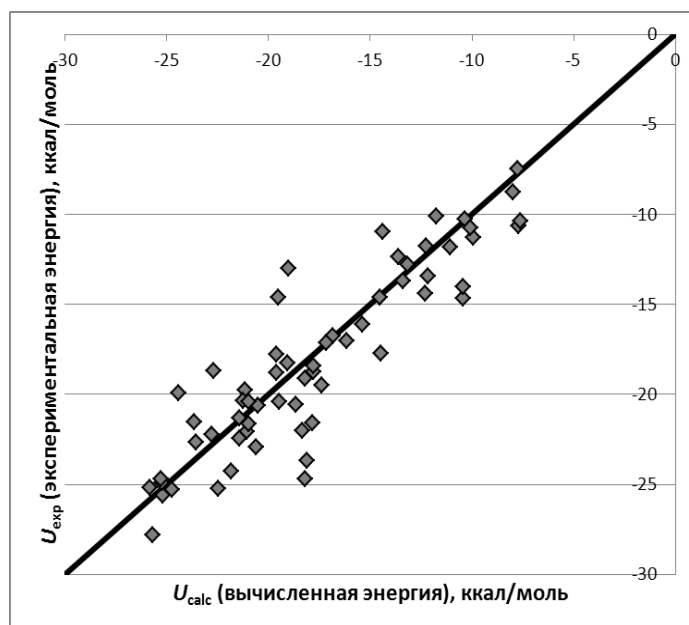
$$\begin{aligned}\lambda_1 &= \beta = 1,04; \\ \lambda_2 &= \frac{1}{\varepsilon} = 0,94; \\ \lambda_3 &= \gamma = 2,89.\end{aligned}$$

При этом коэффициент корреляции предсказанных энергий с экспериментальными составляет 87,8 %.

При добавлении к стандартному силовому полю AMBER только многочастичных дисперсионных взаимодействий (т.е. при оптимизации  $\lambda_1, \lambda_2, \lambda_4$  при  $\lambda_3 \equiv 0$ ), получаем следующие оптимальные значения коэффициентов:

$$\begin{aligned}\lambda_1 &= \beta = 0,92; \\ \lambda_2 &= \frac{1}{\varepsilon} = 0,99; \\ \lambda_4 &= \eta = -1,05.\end{aligned}$$

Коэффициент корреляции предсказанных энергий с экспериментальными составляет 88,9 %.



**Рис. 3.** Сравнение вычисленной (с оптимальными коэффициентами  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ ) и экспериментальной энергии для каждого из 61 рассмотренных в данной работе молекулярных кристаллов. Коэффициент корреляции предсказания с экспериментом составляет 89,9 %. Диагональная линия соответствует полному совпадению вычисленных энергий с экспериментальными.

## ЗАКЛЮЧЕНИЕ

В этой работе разработана новая функциональная форма для силового поля невалентных взаимодействий, включающая важнейшие многочастичные взаимодействия. Разработан метод оптимизации коэффициентов при различных типах взаимодействий с использованием экспериментальных данных по сублимации молекулярных кристаллов. На базе данных, включающей 61 кристалл, проведен расчет энергий каждого типа взаимодействий, а на основе экспериментальных данных по сублимации кристаллов проведена оптимизация коэффициентов при энергиях каждого типа, и сделан расчет суммарной энергии нековалентных взаимодействий для молекулы в кристалле. Показано, что разработанное нами силовое поле позволяет предсказывать энергию невалентных взаимодействий в молекулярных кристаллах с хорошей точностью. Разработанное силовое поле и оптимизированные коэффициенты далее были использованы [17] в молекулярно-динамических расчетах с помощью программы PUMA.

Мы благодарны М.Ю. Лобанову за помощь в работе и Н.К. Балабаеву, О.В. Галзитской, А.В. Гляжиной за ценные консультации.

Работа выполнена при финансовой поддержке гранта 2835/ГФЗ Министерства образования и науки Республики Казахстан, грантов РФФИ № 13-04-00253-а и № 14-04-01554-а, и гранта № 01201358029 программы «Молекулярная и клеточная биология» Российской академии наук.

## СПИСОК ЛИТЕРАТУРЫ

1. Salmon J.K., Shah Y., Wriggers W. Atom-level characterization of structural dynamics of proteins. *Science*. 2010. V. 330. P. 341–346.
2. Lindorff-Larsen K., Piana S., Palmo K., Maragakis P., Klepeis J.L., Dror R.O., Shaw D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct. Funct. Bioinf.* 2010. V. 78. P. 1950–1958.
3. Levitt M., Hirshberg M., Sharon R., Daggett V. Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. *Comput. Phys. Commun.* 1995. V. 91. P. 215–231.
4. MacKerell A.D., Jr., Bashford D., Bellott M., Dunbrack R.L., Jr., Evanseck J.D., Field M.J., Fischer S., Gao J., Guo H., Ha S., Joseph-McCarthy D., Kuchnir L., Kuczera K., Lau F.T.K., Mattos C., Michnick S., Ngo T., Nguyen D.T., Prodhom B., Reiher W.E., III, Roux B., Schlenkrich M., Smith J.C., Stote R., Straub J., Watanabe M., Wiorkiewicz-Kuczera J., Yin D., Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*. 1998. V. 102. P. 3586–3616.
5. Jorgensen W.L., Maxwell D.S., Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 1996. V. 118. P. 11225–11236.
6. Halgren T.A. Merck Molecular Force Field. I. Basis, form, parameterization and performance of MMFF94. *J. Comput. Chem.* 1995. V. 17. P. 490–519.
7. Wang J., Wolf R.M., Caldwell J.W., Kollman P.A., Case D.A. Development and testing of a general Amber force fields. *J. Comput. Chem.* 2004. V. 25. P. 1157–1174.
8. Finkelstein A.V. Average and extreme multi-atom Van der Waals interactions: Strong coupling of multi-atom Van der Waals interactions with covalent bonding. *Chem. Central J.* 2007. V. 1. P. 21.
9. Finkelstein A.V., Lobanov M.Y., Dovidchenko N.V., Bogatyreva N.S. Many-atom Van Der Waals interactions lead to direction-sensitive interactions of covalent bonds. *J. Bioinform. Comput. Biol.* 2008. V. 6. P. 693–707
10. Переяславец Л.Б., Финкельштейн А.В. Силовое поле FFSol для расчета взаимодействий молекул в водном окружении. *Молекулярная биология*. 2010. Т. 44. С. 340–354.
11. Pereyaslavets L.B., Finkelstein A.V. Development and testing of PFFsol\_1, a new polarizable atomic force field for calculation of molecular interactions in implicit water environment. *J. Phys. Chem. B*. 2012. V. 116. P. 4646–4654.
12. Piana S., Klepeis J.L., Shaw D.E. Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations. *Curr. Opin. Struct. Biol.* 2014. V. 24. P. 98–105.
13. Wang J., Cieplak P., Kollman P.A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* 2000. V. 21. P. 1049–1074.
14. Still W.C., Tempczyk A., Hawley R.C., Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 1990. V. 112. P. 6127–6129.
15. Axilrod B.M., Teller E. Interaction of the van der Waals' type between three atoms. *J. Chem. Phys.* 1943. V. 11. P. 299–300.
16. Glyakina A.V., Balabaev N.K., Galzitskaya O.V. Two-, three-, and four-state events occur in the mechanical unfolding of small protein L using molecular dynamics simulation. *Protein Pept. Lett.* 2010. V. 17. P. 92–103.

17. Балабаев Н.К., Гарбузинский С.А, Галзитская О.В., Глякина А.В., Маткаримов Б.Т., Финкельштейн А.В. Включение важнейших многочастичных взаимодействий в силовое поле AMBER и применение обновленного поля к молекулярно-динамическим расчетам. *Математическая биология и биоинформатика*. 2015. Т. 10. № 2. С. 427–435. doi: [10.17537/2015.10.427](https://doi.org/10.17537/2015.10.427).
18. Allen F.H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Cryst.* 2002. V. B58. P. 380–388.
19. Levitt M., Hirshberg M., Sharon R., Daggett V. Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. *Comput. Phys. Commun.* 1995. V. 91. P. 215–231.
20. Bois C. Structure du p-crésol à basse température. *Acta Cryst.* 1970. V. B26. P. 2086.

Материал поступил в редакцию 08.12.2015, опубликован 24.12.2015.