

# Молекулярная механика. Силовые поля.

## Структурная Биоинформатика

Головин А.В.<sup>1</sup>

<sup>1</sup>МГУ им М.В. Ломоносова, Факультет Биоинженерии и Биоинформатики

Москва, 2013

# Содержание

Базисы

Силовые поля

Ковалентные взаимодействия

Нековалентные взаимодействия

Варианты ММ



# Волновая функция

## Wikipedia \* :

Волновая функция — комплекснозначная функция, используемая для описания чистого квантового состояния системы. Обычно функция имеет комплексные значения, а для одной частицы это функция пространства и времени. Изменение волновой функции сравнимо с поведением волны.

**Физический смысл волновой функции** заключается в том, что согласно копенгагенской интерпретации квантовой механики плотность вероятности нахождения частицы в данной точке пространства в данный момент времени считается равной квадрату абсолютного значения волновой функции этого состояния в координатном представлении.



# Уравнение Шредингера

Итак обзовем оператором  $H$  (Гамильтониан):

$$H = \frac{-\hbar^2}{2m} \nabla^2 + V$$

тогда :

$$H\Psi = E\Psi$$

Для решения этого уравнения надо найти значения  $E$  и волновой функции. Это уравнение относится к типу дифференциальных уравнений с собственными значениями, где оператор действующий на функцию возвращает произведение скалярной величины на функцию.



# Операторы

**Ожидаемое значение** (можно рассматривать как среднее значение) какого либо свойства: энергии, положения, линейного момента, можно определить с помощью оператора.

**Пример:** гамильтониан это оператор для энергии можно сказать, что зная волновую функцию:

$$E = \frac{\int \dot{\Psi} H \Psi \partial r}{\int \dot{\Psi} \Psi \partial r}$$

Интегрировать надо по всем осям от  $-\infty$  до  $+\infty$ .

Надо учитывать, что волновая может быть сложным числом и поэтому комплексная составляющая указывается явно.



# Метод самосогласованного поля, SCF

Межэлектронное отталкивание вычисляется как влияние общего (среднего) поля на данный электрон, и это зависит только от положения данного электрона.

Это приближение позволяет повторить разделение переменных в сферических координатах.

$$H_i = \frac{-\hbar^2}{2m} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r_i} + \sum_{j \neq i}^N \left\langle \left( \frac{e^2}{4\pi\epsilon_0 r_{ij}} \right) \right\rangle_j$$

Эти уравнения называют одноэлектронными.

Суть решения состоит в итеративном изменении параметров в функциях, до тех пор пока изменение энергии не станет незначительным.



## Перейдём к молекулам:

Решать напрямую уравнения ХФ по отношению к молекулам, тяжело. Одной из успешных стратегий является введение базисных функций, т.е. волновая функция это комбинация одноэлектронных базисных функций и некоторых коэффициентов.

$$\psi_i = \sum_{\nu=1}^K c_{\nu i} \psi_{\nu}; \quad \frac{\partial E}{\partial c_{\nu i}} = 0$$



# Базис

- **Базисный набор это:** набор математических функций используемых для описания электронных орбиталей атомов в молекуле.
- **Ограниченный базис:** базисные функции в которых **не происходит** изменение параметров в функции в ходе расчёта молекулярных орбиталей.
- **Неограниченный базис:** базисные функции в которых **происходит** изменение параметров в функции в ходе расчёта молекулярных орбиталей.

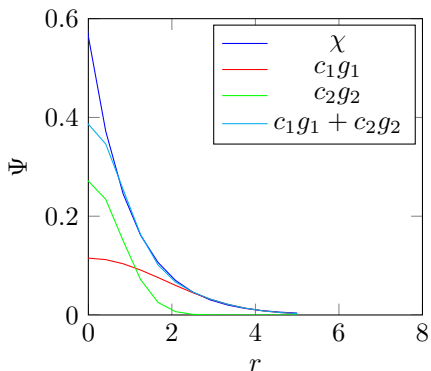
Часто такими математическими функциями является гауссиан:

$$\psi = de^{-\alpha r^2}$$





## Пример STO-2G для H:



$$\psi = \sum_1^2 c e^{-\alpha r^2}$$

$$\psi = c_1 g_1 + c_2 g_2$$

	$\alpha$	$c$
1	0.151623	0.678914
2	0.851819	0.430129

$\epsilon = -0.4665819$  a.u. = 12.697 eV.

В реальности 13.606 eV. И ошибка 87.7 кДж/моль



# Гауссианы:

$$1s = Ne^{\alpha r^2}; \quad 2p_x = Ne^{\alpha r^2} x; \quad 2p_y = Ne^{\alpha r^2} y; \quad 2p_z = Ne^{\alpha r^2} z;$$

$$3d_{xx} = Ne^{\alpha r^2} x^2; \quad 3d_{xy} = Ne^{\alpha r^2} xy; \quad 3d_{xz} = Ne^{\alpha r^2} xz;$$

$$3d_{yy} = Ne^{\alpha r^2} y^2; \quad 3d_{yz} = Ne^{\alpha r^2} yz; \quad 3d_{zz} = Ne^{\alpha r^2} z^2;$$

$$4f_{xxx} = Ne^{\alpha r^2} x^3; \quad 4f_{xxy} = Ne^{\alpha r^2} x^2 y; \quad 4f_{xxz} = Ne^{\alpha r^2} x^2 z;$$

И так далее.



# Описания базисных наборов для программы GAUSSIAN:

Общий вид обозначений от Поупл и коллег: **M-ijk..G**

- **M** – количество ограниченных гауссианов на один не валентный электрон
- Наличие двух и более букв после “-” означает, что валентные электроны описываются 2 и более функциями, каждая из которых состоит из линейной комбинации **i,j,k** гауссианов
- \* -Означает, что для тяжёлый атомов используются не только гауссианы характерные для конкретной орбитали, но и гауссианы следующей орбитали.

Например для углерода в **3-21\*G**: у валентных электронов с 3 гауссианами прибавляется 6 гауссианов для d-орбиталей.



# Описания базисных наборов для программы GAUSSIAN:

- \*\* то же самое, что и \* ,но добавляются 3 гауссиана для p-орбиталей к гауссианам H и He.
- + Означает добавление дополнительных гауссианов тех же орбиталей, но с маленьким значением  $\alpha$ . Этот шаг нужен для точного счёта систем где значительная электронная плотность удалена от ядра: электронные пары, анионы.



## Итак:

Запись Базиса	Количество гауссианов на орбиталь, для C	Применение
STO-3G	3	Большие системы
6-31G	6-не валентные 3+1-валентные	Системы без поляризации
6-31*G	То же самое + 6 функций типа l+1	Системы с анизотропией заряда
6-31**G	см. выше + 3 p-функции для H и He	Водородная связь
6-31+G(2df)	Поляризационные: 2*6 d-функций + 7 f- функций (см *) Диффузных : 4	Нужно там, где важно точно рассчитать высокую плотность электронов
6-311++ G(3df,3pd)	Диффузные на все атомы и поляризационные на C: 3*6 d типа + 7 f типа H: 3*3 p типа + 1 d типа	Если всё, что было до этого было не достаточно точным.



## Как не надо:

- **3-21++G(2df,2pd)**

Только две функции на валентную орбиталь и 3 поляризационных набора функции с кучей гауссианов и диффузные наборы.

- **6-311+G(2df)**

3 набора на валентную орбиталь, 3 поляризационных + диффузный набор для тяжёлых атомом и отсутствие поляризации для водорода?



# Уравнение Шредингера

$$\left( -\frac{\hbar^2}{m} \left[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] + V \right) \Psi(r, t) = i\hbar \frac{\partial \Psi(r, t)}{\partial t}$$

Или:

$$H\Psi = E\Psi; \quad H = \frac{-\hbar^2}{m} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r}$$

Современные базисы предполагают примерно 60 функций на атом. Итого: 900 функций на аминокислоту.

- Можно аппроксимировать электронную плотность уравнениями классической физики.



# Молекулярная механика (ММ)

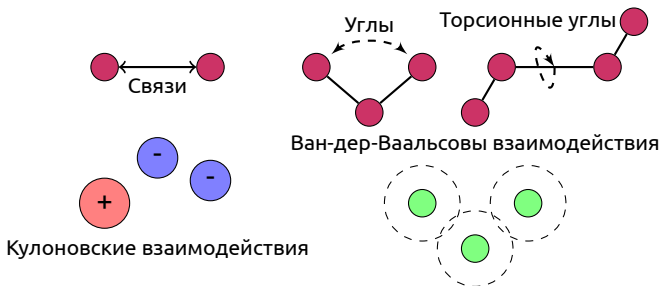
- В ММ электронная структура атома замещается на достаточно простые уравнения с параметрами.
- Наборы параметров называются силовыми полями.
- Используется допущение Борна-Оппенгеймера (электроны быстро адаптируются к движению ядер)
- Расчёт энергии происходит на основе положения ядер.
- Упрощения позволяют работать с большими системами
- В некоторых случаях ММ подходы могут давать результаты, сравнимые по точности с методами QM.





# Простое уравнение силового поля (СП)

$$\begin{aligned}
 U = & \sum_{bonds} \frac{k_i}{2} (l_i - l_0)^2 + \sum_{angles} \frac{k_i}{2} (\phi_i - \phi_0)^2 + \sum_{torsions} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \\
 & + \sum_{i=1}^N \sum_{j=i+1}^N \left( 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)
 \end{aligned}$$

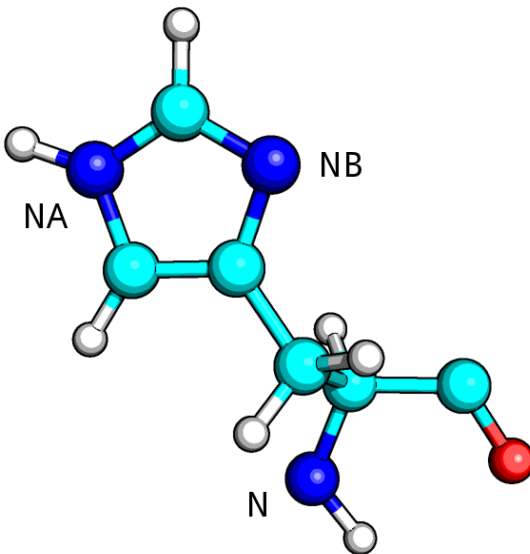


# Основные особенности силовых полей

- Большинство параметров неотделимо от поля.
- Параметризация ММ сильно зависит от целей исследования.
- Большинство силовых полей параметризованы для воспроизведения структуры.
- Силовые поля — это результат оптимизации параметров.
- Силовые поля — это эмпирически найденные данные.



# Типы атомов в СП



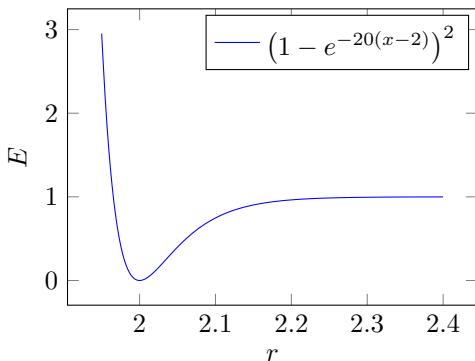
# Потенциал для описания связи

## Потенциал Морзе

$$U(l) = D_e \{1 - e^{[-a(l-l_0)]}\}^2$$

## Более распространённый

$$U(l) = \frac{k_i}{2} (l_i - l_0)^2$$



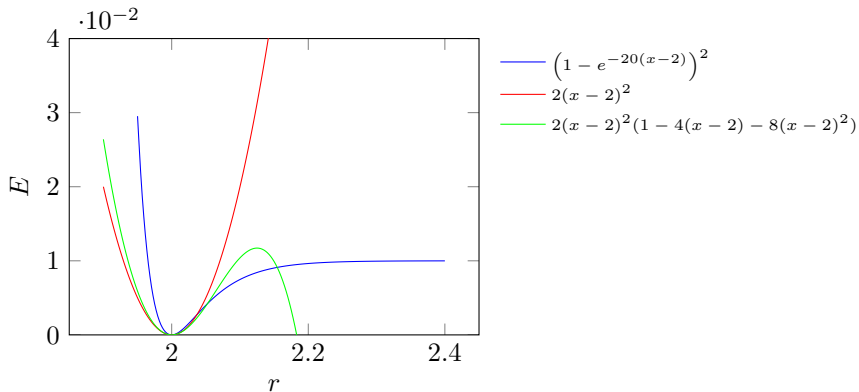
# Параметры при описании связи

$$E = \frac{k}{2}(r - r_0)^2$$

связь	$r_0, \text{Å}$	$k, \text{kcal mol}^{-1} \text{Å}^{-2}$
$C_{sp3}-C_{sp3}$	1.523	317
$C_{sp2}-C_{sp2}$	1.337	690
$C_{sp2}-O_{sp2}$	1.208	777
$C_{sp3}-N_{sp3}$	1.438	367



## Кубический и прочие варианты



$$U = \frac{k_i}{2}(l_i - l_0)^2(1 - k'(l_i - l_0) - k''(l_i - l_0)^2 - k'''(l_i - l_0)^3 - k''''(l_i - l_0)^4 \dots)$$



# Потенциал валентного угла

$$U(\phi) = \frac{k_i}{2}(\phi_i - \phi_0)^2$$

или

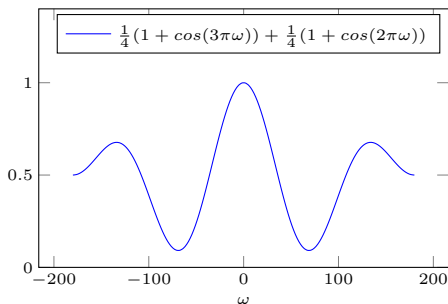
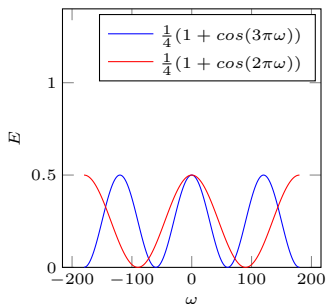
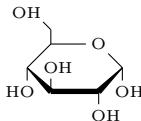
$$U(\phi) = \frac{k_i}{2}(\phi_i - \phi_0)^2(1 - k'(\phi_i - \phi_0) - k''(\phi_i - \phi_0)^2 - \\ - k'''(\phi_i - \phi_0)^3 - k''''(\phi_i - \phi_0)^4 \dots)$$



# Потенциал торсионного угла

$$U(\omega) = \sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma))$$

Рассмотрим O-C-C-O (сахар в ДНК)





# Потенциал торсионного угла

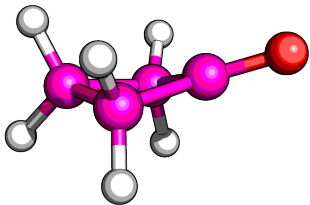
$$U(\omega) = \frac{V_1}{2}(1 + \cos\omega) + \frac{V_2}{2}(1 + \cos 2\omega) + \frac{V_3}{2}(1 + \cos 3\omega) \dots$$

- Для поля MM2 используют три члена.
- Поле OPLS использует ряды с 4-ю слагаемыми.

$$U(\omega) = \frac{1}{2} [F_1(1 + \cos\omega) + F_2(1 - \cos 2\omega) + F_3(1 + \cos 3\omega) + F_4(1 - \cos 4\omega)]$$



# "Неправильные" торсионные углы



$$U(\omega) = V_1(1 - \cos\omega)$$

$$U(\omega) = V_1(\omega - \omega_0)^2$$

Для циклобутанона кислород должен находиться в одной плоскости с  $C_1, C_2, C_3$ .

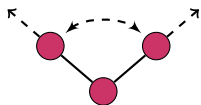
Используют потенциал, где перечисление не 1-2-3-4, а 1-4-2-3



# Кросс-составляющие в силовых полях

Кросс-составляющие отражают зависимость состояния одной связи или угла от состояния соседней связи.

Существуют: stretch-strech, stretch-bend, stretch-torsion



$$U(l_1, l_2) = \frac{K_{l_1 l_2}}{2} (l_1 - l_{1,0})^2 (l_2 - l_{2,0})^2$$

$$U(l_1, l_2, \phi) = \frac{K_{l_1 l_2 \phi}}{2} [(l_1 - l_{1,0})^2 + (l_2 - l_{2,0})^2] (\phi - \phi_0)$$



# Нековалентные взаимодействия

- Нековалентные взаимодействия являются определяющими в формировании структуры биополимеров.
- Так как эти взаимодействия реализуются через пространство, то часто они описываются как функции, обратно пропорциональные расстоянию между двумя атомами.



# Электростатические взаимодействия

Допустим, что поверхность единичного потенциала можно представить зарядами в центрах атомов.

Тогда электростатические взаимодействия будут описываться по закону Кулона:

$$U(q_1, q_2) = \frac{q_1 q_2}{4\pi\epsilon_0\epsilon_r r_{ij}};$$

$$U = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}$$



# Электростатические взаимодействия

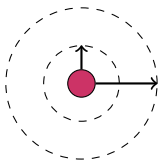
Очевидно, что количество вычислений растет значительно быстрее количества частичных зарядов.

Существуют следующие упрощения:

- Разрастание центрального мультиполя (ММ малых молекул)
- Двойное обрезание
- Потенциал реакционного поля
- Суммирование Эвальда



# Двойное обрезание



Для сферы А мы считаем все частичные заряды, а для сферы А-В мы будем считать взаимодействие групп зарядов с нашим атомом.

$$U_1 = \sum_{i=1}^{N_A} \frac{q_1 q_i}{4\pi\epsilon_0\epsilon_r r_{1i}} + \sum_{j=1}^{N_{group}} \frac{q_1 q_j}{4\pi\epsilon_0\epsilon_r r_{1j}}$$



# Потенциал реакционного поля

Основная идея: мы считаем, что за некоторым расстоянием плотность заряда одинаковая, и, следовательно, известна некая диэлектрическая проницаемость среды.

$$U_{ij} = \frac{q_1 q_i}{4\pi\epsilon_0\epsilon_r r_{1i}} \left[ 1 + \frac{\epsilon_{rf} - \epsilon_r}{2\epsilon_{rf} + \epsilon_r} \frac{r_{ij}^3}{r_c^3} \right] - \frac{q_1 q_i}{4\pi\epsilon_0\epsilon_r r_c} \frac{3\epsilon_{rf}}{2\epsilon_{rf} + \epsilon_r}$$





# Суммирование Эвальда

Основная идея: нам нужно учитывать не только заряды в ближайшем окружении, но и, как в кристалле, заряды, находящиеся в соседних ячейках.

$$U_{ij} = \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \sum_{z=1}^{N_z} \sum_{i=1}^N \sum_{j=1}^N N \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}$$

Это сходится, но очень медленно.



# Суммирование Эвальда

Эвальд предложил перевести этот ряд в сумму 2-ух быстро сходящихся рядов и константы.

$$U = U_{dir} + U_{rec} + U_0$$

$$U_{dir} = f/2 \sum_{i,j}^N \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \sum_{z=1}^{N_z} q_i q_j \frac{\text{erfc}(\beta r_{ij,n})}{r_{ij,n}}$$

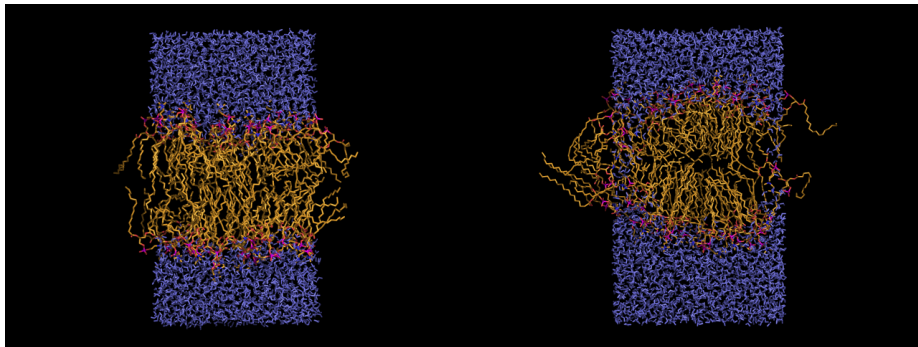
$$U_{rec} = \frac{f}{2} \pi V \sum_{i,j}^N q_i q_j \sum_{m_x} \sum_{m_y} \sum_{m_z} \frac{\exp(-\pi m/\beta)^2 + 2\pi i m (r_i \cdot r_j)^2}{m}$$

$$U_0 = \frac{f\beta}{\sqrt{\pi}} \sum_i^N q_i^2$$

Где бета - это параметр, определяющий соотношение прямого и обратного взаимодействий



# Суммирование Эвальда vs двойное обрезание



Self-assembly with PME

Self-assembly with Cut-off



# Ван-дер-Ваальсовы взаимодействия

- В основе природы Ван-дер-Ваальсовых взаимодействий лежат электронные эффекты: дисперсионные и обменные.
- В принципе, рассчитать такие эффекты можно в QM, но это далеко не тривиальная задача.
- В ММ нам надо считать такие взаимодействия быстро, на сегодняшний день наиболее часто используют потенциал Леонарда-Джонса:

$$U_{VdW} = \sum_{i=1}^N \sum_{j=i+1}^N 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$



# Ван-дер-Ваальсовы взаимодействия

Наряду с потенциалом Леонарда-Джонса используют потенциал Букингама:

$$V_{bh}(r_{ij}) = A_{ij} \exp(-B_{ij} r_{ij}) - \frac{C_{ij}}{r_{ij}^6}$$



# Взаимодействия между разными типами атомов

Константы для разных типов атомов будут разные. Для их определения существуют правила смешивания:

$$\sigma_{AB} = \frac{1}{2} (\sigma_{AA} + \sigma_{BB})$$

$$\epsilon_{AB} = \sqrt{\epsilon_{AA}\epsilon_{BB}}$$

Это не единственный вариант правила смешивания, но такой подход наиболее распространён для моделирования биологических систем



# Различия для 1-4 взаимодействий

- Так как 1-4 взаимодействия могут быть уже учтены в описании торсионного угла, то может быть, что силовых полях такие нековалентные взаимодействия не учитываются.
- В полях семейства AMBER, 1-4 VdW взаимодействия всё-таки учитываются, но их потенциал делится на 2.



# Водородные связи

- В силовых полях водородная связь часто описывается как комбинация Ван-дер-Ваальсовых и Кулоновских взаимодействий
- Существуют силовые поля, где водородная связь задаётся своим потенциалом на основе потенциала Леонарда-Джонса 10-12:

$$U_{HB} = \frac{A^{10}}{r} - \frac{C^{12}}{r}$$

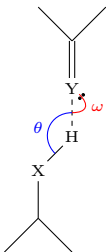




# Водородные связи

Для точного описания водородной связи вносят поправки, учитывающие геометрию водородной связи:

$$U_{HB} = \left( \frac{C}{d^6} - \frac{D}{d^4} \right) \cos^m \theta$$



$$U_{HB} = \left( \frac{A}{r_{H...Ac}^{10}} - \frac{C}{r_{H...Ac}^{12}} \right) \cos^2 \theta_{Don-H...Acc} \cos^4 \omega_{LP-Acc...H}$$



# Эффективный парный потенциал

Для системы из 1000 частиц существует 499500 парных взаимодействий и 166167000 тройных взаимодействий.

## Выход есть:

- Использование парного потенциала с 'правильной' параметризацией.
- Пример: использовать большие частичные заряды для фазы, чем для одной молекулы.
- Это работает для воды. 1.85 D vs 2.6 D



# Модели воды

- Вода - достаточно сложный объект.
- Важно, что модель воспроизводит как свойства одной молекулы, так и свойства фазы.

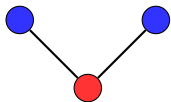
## **Существуют три основных класса моделей:**

- Простые модели
- Поляризуемые модели
- Ab initio модели

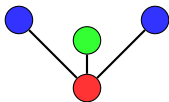


# Модели воды

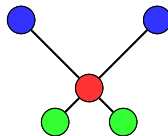
## Простые модели



SPC, TIP3P



TIP4P



ST2, TIP5P

В большинстве случаев применяют так называемые rigid body варианты моделей, хотя и существуют модели, где связи представлены потенциалами.

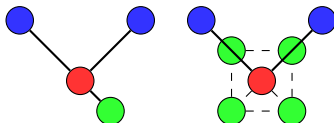


# Модели воды

## Поляризуемые модели

Есть два подхода:

- Смещать центр заряда кислорода относительно центра атома
- Добавить точки вокруг кислорода, в которых может меняться заряд

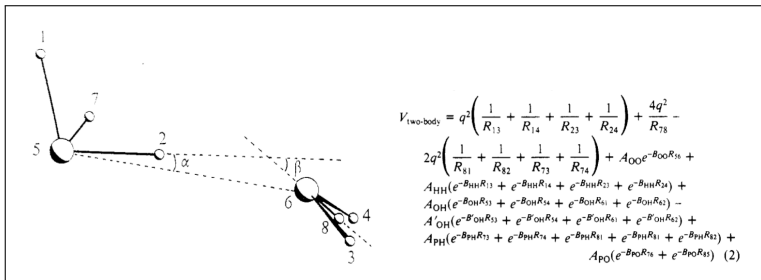


# Модели воды

## Ab initio модели

Базируются на квантовомеханических вычислениях как одной, так и нескольких молекул воды.

NCC модель:



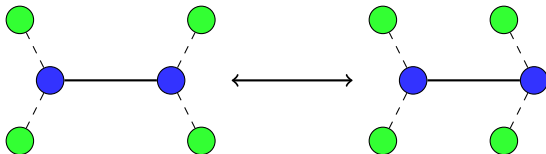
## Conclusions

- ▶ No water model available is able to reproduce all the water properties with good accuracy. All empirical models are fitted to something and the ab initio models suck....
- ▶ It is, never the less, possible to describe the force field of water using simple empirical models, and predict nature in many fields.
- ▶ It is worth considering which models to use when performing any simulation including water. Computational cost vs accuracy.

# Силловые поля с объединёнными атомами

Основная идея: не учитывать атомы водорода, не принимающие участие в образовании водородной связи. К массе атома без водорода добавляется 1.

Есть проблема:





# Молекулярная механика твердого тела

$SiO_2$  - типичный объект подобных исследований. Часто бывает необходимо наблюдать дефекты в образовании кристаллической структуры.

Ковалентную составляющую заменяют на модифицированные нековалентные потенциалы:

$$U = \sum_{i=1}^N \sum_{j=i+1}^N \left( 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)$$

GlassFF:

$$U = \sum_{i=1}^N \sum_{j=i+1}^N \left( D_0 \left[ e^{r(1-r_{ij}/r_0)} - 2e^{\frac{r}{2}(1-r_{ij}/r_0)} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)$$

