

Лекция 2. Моделирование. PyMol

Курс: Молекулярное моделирование биополимеров

Головин А.В.¹

¹МГУ им М.В. Ломоносова, Факультет Биоинженерии и Биоинформатики

Москва, 2017

Содержание

Лекции

Практикум

Общие рассуждения

Концепты

Введение

Визуализация с PyMol

Selections

Анимация

Моделирование и редактирование в PyMol

Скриптование в PyMol

Содержание курса, лекции

- Введение в квантовую механику, HF, DFT, MP.
- Молекулярная механика
- Молекулярная механика и минимизация энергии
- Переходные состояния
- Молекулярная динамика
- Гибридное моделирование QM/MM, Adaptive resolution
- Модификации молекулярной динамики: Метадинамика и прочее
- Введение в методы Монте-Карло

Содержание курса, лекции

- Анализ конформаций, MSM
- Моделирование структуры белков
- Поиск новых био-активных молекул и химоинформатика
- Свойства лигандов, построение лигандов.

Практические навыки

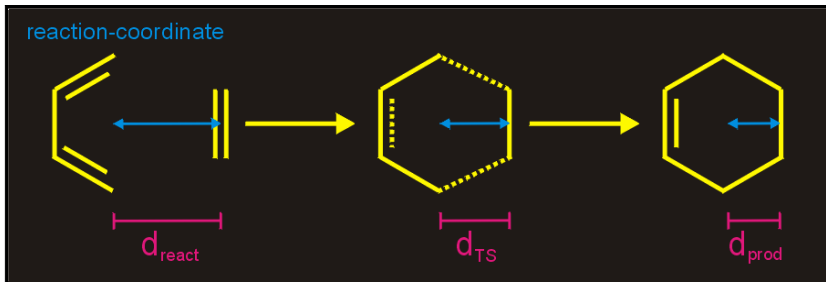
Software: Pymol

- Анимация
- Редактирование
- Моделирование
- Внесение мутаций
- Генерация структур
- Скриптование PyMol + Python

Практические навыки

Software: Полу эмпирические расчеты, Морас

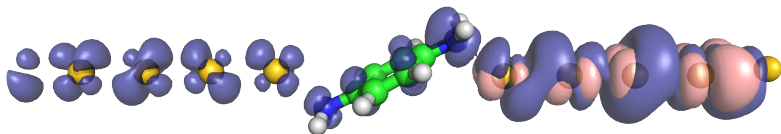
Работа с программой МОРАС для описания электронной структуры молекул полу эмпирическими методами



Практические навыки

Software: Gamess, ORCA

- Работать с программой GAMESS для описания электронной структуры *ab initio*
- Рассчитывать электронные свойства системы
- Предсказывать спектры



Практические навыки

Software: Gromacs

- Проводить моделирование молекулярной динамики в пакете GROMACS
- Параметризовать новые соединения
- Моделировать само сборку бислоя
- Моделировать конформационные переходы в ДНК



Дополнительные навыки

Software: Gromacs+Orca

- Моделировать взаимодействия ионов с биополимерами методами ММ/QM
- Моделировать химические реакции в белках

Практические навыки

Software: Modeller + Rosseta

- Строить структуру белка или РНК по гомологии
- Строить структуру комплексов белков с различными молекулами

Практические навыки

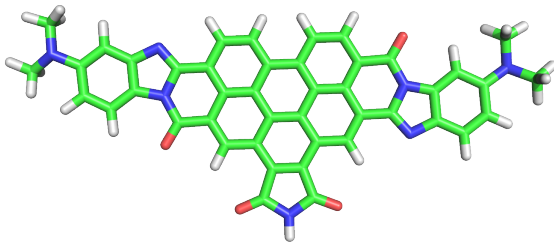
Software: Autodock vina +

- Находить полости в структуре биополимера
- Проводить докинг низкомолекулярных лигандов в найденные полости
- Проводить компьютерный скрининг баз данных низкомолекулярных веществ на связывание с белком

Практические навыки

Software: OpenBabel +

- Использовать химоинформатические базы данных
- Использовать SMILES, SMARTs
- Строить структуру лиганда на основе структуры белка (??)



```
O=C1c2c3cc4c5c6c3c3c(cc7c8c3c(ccc8c3nc8ccc(cc8n3c7=O)
N(C)C)c6ccc5c3nc5ccc(cc5n3c4=O)N(C)C)c2C(=O)N1
```

Перерыв

Немного рассуждений..

- Первые шаги к пониманию того, что вещество состоит из маленьких элементов сделал Лукреций, давно.
- Первые эксперименты по установлению структуры были проведены только в начале 20 века

Первые шаги

- Появились специальные молекулярные наборы из шариков и палочек
- Правильное использование химической информации позволило строить первые модели, очень похожие на результаты РСА.

Компьютер

- Формально для решения задач моделирования компьютер не является обязательным элементом
- Быстрый компьютер значительно увеличивает точность и следовательно достоверность моделирования.
- Количество вычислений отражает степень исследования конформационного пространства

Компьютер и программы

Those programs **always provide a result**, the evaluation of which is at liberty of the user. The programs **tend stubbornly** to calculate **every absurd application** and present a result-not only a number, but also a graph and represent a further instrument of seduction for the uncritical use of algorithms.

Для чего нужны модели?

1. Упрощение

Упрощение сложного объекта до анализа только той части которая предполагаемо является объектом интереса.

Для чего нужны модели?

2. Иллюстрация для раздумий

Модель является иллюстрацией для дедуктивного анализа очень сложных или многочисленных явлений.

Для чего нужны модели?

3. Механистическое представление больших систем

Часто модели отражают реальность не полностью, но моделируемой точности бывает достаточно для понимания рассматриваемой системы.

Для чего нужны модели?

4. Математическое моделирование

Моделирование кинетических схем превращений при каскадных энзиматических реакциях, позволяет найти оптимум действия фермента.

Финальный этап: Дизайн

Моделирование структуры, определение свойств это шаги к самому важному этапу :
дизайну или проектированию нового вещества с заданными свойствами.

Полезная информация

Система координат

Cartesian coordinates

C	0.98770	-0.03260	-0.09450
C	2.32350	-0.03260	-0.09450
H	0.42790	0.73340	0.43290
H	0.42790	-0.79870	-0.62190
H	2.88330	0.73340	0.43290
H	2.88330	-0.79870	-0.62190

Z-matrix

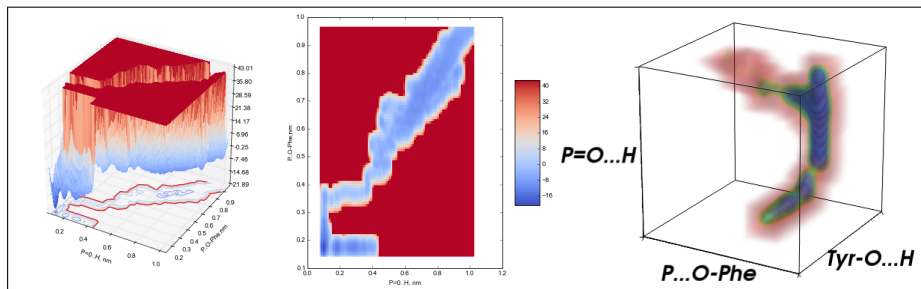
```

C 1 r1
H 1 r2 2 a1
H 1 r2 2 a1 3 d1
H 2 r2 1 a1 3 d1
H 2 r2 1 a1 3 d1
r2= 1.3358
r3= 1.0855
a1= 120.00
d1= 180.00

```

Полезная информация

Поверхность потенциальной энергии это отображение значения потенциальной энергии системы в зависимости от расположения атомов

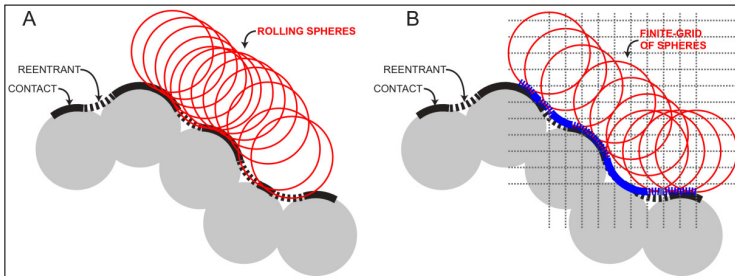


Компьютерная графика

- Когда-то компьютерная графика была малодоступным способом отображения молекул, но не сейчас.
- Компьютерная графика это только отображение, визуализация является только шагом первичного анализа при моделировании.

Поверхности молекул

- Поверхности единичного заряда
- Поверхности на основе Ван-дер-Ваальсового радиуса



Программное обеспечение

Специализация:

- Программы для визуализации хорошо отображает структуры.
- Софт для счёта хорошо считает. Часто расчёт занимает существенное время и следовательно выполняется в фоновом режиме и желательно на суперкомпьютере.

Аппаратное обеспечение

Основная тенденция это перенос расчёта на вычислительные комплексы

Особенности использования:

- Для удаленной работы графический интерфейс неудобен.
- Консольный режим предоставляет необходимую функциональность для запуска и первичного анализа.
- Детальный анализ проводится на локальной рабочей машине.

Визуализация с PyMol



Для чего нужен PyMol

- Визуализация pdb и прочих файлов с координатами атомов
- Изготовление высококачественных изображений
- Начальное редактирование структур

Системные требования

Компьютер: чем мощнее процессор и чем больше памяти, тем лучше

3D монитор не обязателен, но поддерживается

Операционная система: любая, под Linux проще установить, и он лучше работает с памятью.

Как установить?

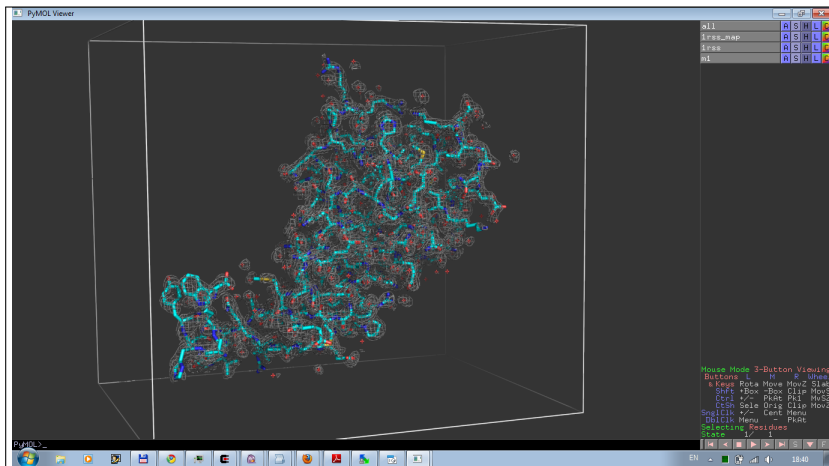
- Компиляция из исходников: <http://pymol.svn.sourceforge.net/>
- Установка бинарных пакетов в Ubuntu Linux: `sudo apt-get install pymol`
- Установка бинарных пакетов в Windows:
 - Ресурс для установки с python:
<http://www.lfd.uci.edu/~gohlke/pythonlibs/#pymol>
 - Компиляция под Windows:
<http://arcib.dowling.edu/~darakevn/installerpaper.pdf>

PyMol - это GPL программа?

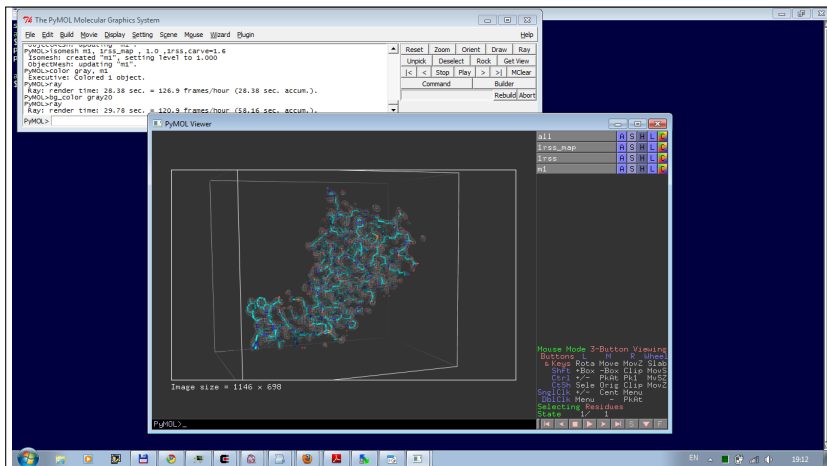
Да, PyMol это GPL-программа;

- исходный код доступен на sourceforge.net
- Бинарные пакеты для windows стоят денег и продаются:
<http://pymol.org/academic.html>
- Бинарные пакеты для Linux собираются майтенерами

PyMol



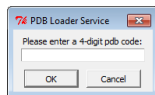
Основной вид



Как загрузить структуру?

- Из интернет:
 - в меню выбрать соответствующий plugin
 - или в командной строке: `fetch 1xxx`

- Локальный файл:
 - File->Open



Использование мыши

- Левый клик + движение = вращение молекулы
- Средний клик + движение = перемещение молекулы
- Правый клик + движение вверх/вниз = приближение/удаление молекулы
- Колесо = изменение уровня обрезания молекулы
- **Все манипуляции относятся к камере, а не координатам структуры**

Меню объекта/выборки

The screenshot shows the PyMOL Viewer interface. A 3D molecular model is displayed in the center. A context menu is overlaid on the model, listing the following objects: all, 1rss_map, 1rss, and m1. Each object name is followed by a grid of five colored buttons labeled A, S, H, L, and C. A red arrow points from the menu to a specific point on the model. Below the screenshot, the text "A,S,H,L,C" is written in red.

all	A	S	H	L	C
1rss_map	A	S	H	L	C
1rss	A	S	H	L	C
m1	A	S	H	L	C

A,S,H,L,C

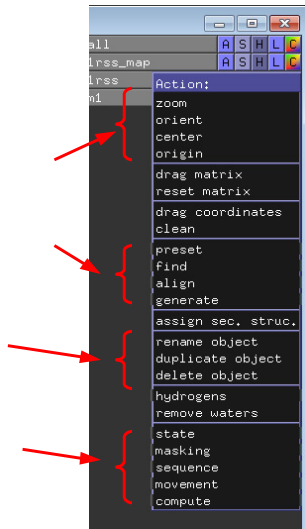
A=Action

Манипуляции с ориентацией

Предустановки изображения и т.д.

Манипуляция с объектом

Прочее



S=Show, H=Hide

The image displays a PyMol interface with eight different visualization styles for a protein structure, arranged in a 2x4 grid. The top row shows 'lines' (a pink wireframe), 'sticks' (a yellow stick model), 'spheres' (a cyan and red space-filling model), and 'surface' (a red and blue surface representation). The bottom row shows 'mesh' (a blue and red wireframe mesh), 'dots' (a grey and red dot model), 'ribbon' (a pink ribbon), and 'cartoon' (a yellow and white cartoon representation). To the right of the grid is a command menu with a list of styles and their corresponding keyboard shortcuts. At the bottom right, there is a 'Mouse Mode' section with various navigation and selection controls.

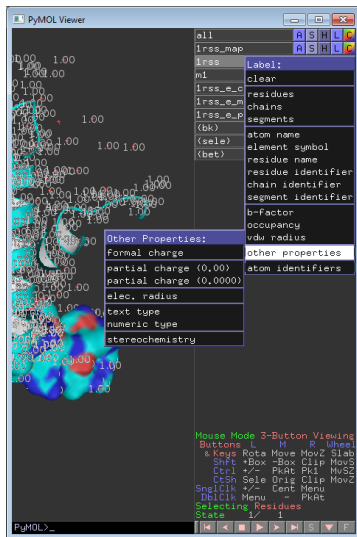
all	A S H L F
irss_map	A S H L F
irss	Show:
m1	as
irss_e_chg	lines
irss_e_map	sticks
irss_e_pot	ribbon
(bk)	cartoon
(sele)	label
(bet)	cell
nonbonded	
dots	
spheres	
nb_spheres	
mesh	
surface	
organic	
main chain	
side chain	
disulfides	

```

Mouse Mode 3-Button Viewing
Buttons L M R Wheel
a Keys Rota Move MovZ Slab
Shft +Box -Box Clip MovS
Ctrl +/- PkAt Pk1 MvSZ
CtrlSh Sele Orig Clip MovZ
SingleClick +/- Cent Menu
DoubleClick Menu - PkAt
Selecting Residues
State 1/ 1

```


L=Label



C=Color

The screenshot shows the PyMOL Viewer interface. On the left, a 3D molecular model is displayed with a color scale ranging from 0.00 (blue) to 1.00 (red). The model is a complex, multi-chain structure. On the right, the command console shows the following commands:

```

ell      A S H L C
lrsr_map A S H L C
lrsr     Atoms  Color:
ml       CHNOS... by element
lrsr_e_1 CHNOS... by chain
lrsr_e_2 CHNOS... by ss
lrsr_e_3 CHNOS... spectrum
(bk)     CHNOS... auto
(<sele) CHNOS... reds
(<bet)   CHNOS... greens
         CHNOS... blues
         CHNOS... yellows
set 2    magentas
set 3    cyans
set 4    oranges
set 5    tints
set 6/H  grays
  
```

At the bottom of the console, there is a section for mouse controls:

```

Mouse Mode 3-Button Viewing
Buttons L M R Wheel
a Keys Rotate Move MovZ Slab
ShFt +Box -Box Clip MovS
Ctrl +/- PKAt Pk1 MvSz
Dcsh Sele Driz Clip MovZ
SnglClk +/- Cent Menu
DbIClk Menu - PKAt
Selecting Residues
State 1/ 1
  
```

Выборки

- Можно задать с помощью кликов мыши, удерживая SHIFT
- Удобнее писать выражения в командной строке

Например: *Select backbone, name ca+c+n*

Операторы множеств

- Логические операторы AND, OR, NOT
Операция OR может быть записана как ”.

Упражнение: Документ PDB содержит описание структуры, состоящей из белка, фрагмента ДНК и молекул воды. Что получится, если задать следующие команды ?

select protein or dna

select protein and dna

select not water

- Оператор WITHIN(...)
select all within 3.5 of resi 20
select s1, (byres n. ca) within 3.5 of resn LIG

Help selections

Длинное

name <atom names>
 resn <residue names>
 resi <residue identifiers>
 chain <chain ID>
 id <original-index>
 hydrogen
 all
 visible
 hetatm
 byres <selection>
 byobj <selection>
 around <distance>
 expand <distance>
 in <selection>
 like <selection>

Короткое

n. <atom names>
 r. <residue names>
 i. <residue identifiers>
 c. <chain identifiers>
 h.
 *
 v.
 br. <selection>
 bo. <selection>
 a. <distance>
 e. <distance>
 l. <selection>

<selection> within <distance> of <selection>
 <selection> w. <distance> of <selection>

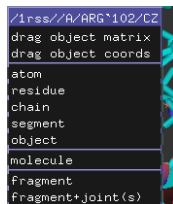
Примеры выборов

sel=select

- sel s1, n. ca and c. A : все атомы CA в цепи A
- sel s2, n. ca and (c. A or c. B) : атомы CA цепей A и B
- sel s3, resn GLU and resi 100 : остаток 100 если он GLU
- sel s4, resi 100-120+130 : атомы остатков 100-120 и 130
- sel s5, byres(name CG) : атомы остатков где есть CG

Иерархическое определение выборки

Легко увидеть иерархию правым кликом по атому



sel s1, a/102/cz : атом cz в остатке 102

sel s2, 100-120/N and c. A : атомы N в остатках 100-120 цепи a

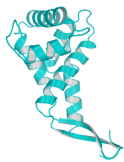
sel s3, a/100+120/ : все атомы остатков 100 и 120 в цепи A

Трассировка лучей, команда ray

Подробнее: <http://www.pymolwiki.org/index.php/Ray>



No ray



ray_trace_mode,0



ray_trace_mode,1



ray_trace_mode,2



ray_trace_mode,3

Настройки изображения

<http://www.pymolwiki.org/index.php/Category:Settings>

- PyMol содержит порядка 600 настроек
- Не все документированы
- Большинство интуитивно понятны
- Настройки доступны через меню или в командной строке набрать:
set первые буквы имени опции и клавиша tab для достроения

Примеры

#initial setup

`viewport 600, 600` — размер графического окна

`set auto_zoom, off` — не приближать новые объекты

`set auto_show_lines, off` — не показывать линии автоматически

`set auto_show_selections, off` — не показывать выборку автоматически

#cartoon parameters

`set cartoon_fancy_helices, 1` — изменение вида спиралей

`set cartoon_highlight_color, grey60` — цвет внутренней стороны спиралей

`set cartoon_dumbbell_length, 1.0` — ширина ленты в спирали

`set cartoon_rect_length, 1.40000` — ширина ленты в бета

`set cartoon_loop_radius, 0.3` — толщина неструкт. участка

`set cartoon_smooth_loops=0` — без сглаживания

Анимация в PyMol

Если структура содержит более чем одну модель, то в PyMol можно анимировать движение молекулы переходом от одной модели к другой



Анимация, основы

GUI :

Вращение вокруг объекта на N секунд:

- Movie->Program->Camera->X-Roll->N Seconds
- Movie->Program->Camera->Y-Roll->N Seconds

Покачивание:

- Movie->Program->Camera->X-Rock->X-Degrees over N-Seconds

Пример

- Action->Preset->Technical (viewer gui)
- Scene->Store->F1
- zoom i. 90 # увеличение остатка 90
- Scene->Store->F2
- Movie->Program->Scene Loop->Y-Rock->4 Seconds Each
- File-> Save movie

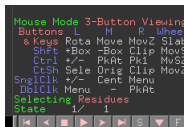
Результат



Анимация, терминология

- Объект и выборка : смотри выше
- states: конформация или набор координат
- scene: позиция камеры и отображение
- frames: это кадры в анимации, содержит state и scene

Movie panel:



Анимация, команды

`mset 1 -55` : задать анимацию от 1 до 55 state на 55 кадров (frames)

`mset 1 x90` : задать анимацию первого state от 1 до 90 кадров

`mset 1 x30 1 -15 15 x30 15 -1` : первые 30 кадров state 1, следующие 15 кадров это состояния 1-15, следующие 30 кадров состояние 15, следующие 15 кадров состояния от 15 до 1

Анимация, команды

`mview` : команда для создания ключевых точек

Пример :

- `mset 1 x100`
- `frag leu #` создаём LEU
- `orient #` ориентируем его
- `mview store #` запоминаем ключевую точку
- `frame 100 #` переходим в кадр 100
- `zoom ID 10 #` увеличиваем атом №10
- `mview store #` запоминаем ключевую точку
- `mview reinterpolate #` делаем интерполяцию

Результат mview



Дополнительные команды

- `mmatrix` : устанавливает вид для первого кадра
- `util.mrock` : покачивание сцены на определённый угол
- `util.mrock(start, finish, angle, phase, loop-flag)`
- `util.mroll` : вращение вокруг оси Y
- `util.mroll(start, finish, loop-flag)`
- `mdo` : (устарело) запуск какой-либо команды в заданном кадре

Сохранение анимации

Старый путь:

```
set ray_trace_frames,1  
png movie
```

Нужны программы avidemux, Virtual Dub, mencoder для того, чтобы собрать ролик с нужным сжатием (кодек)

Новый путь: File->Save movie ; есть недостаток, старый офис понимает только avi с определённым кодеком

Моделирование и редактирование в PyMol

- Можно перемещать объекты и сохранять их новые координаты
- Можно рассчитать вторичную структуру
- Можно менять координаты отдельных атомов
- Можно вносить мутации в белок (но не НК)
- Можно конвертировать L->D аминокислоты
- Можно добавлять протоны
- Можно выравнивать в пространстве молекулы
- Можно добавлять некоторые фрагменты из библиотеки и собственные

Перемещение объектов

Рекомендуемый порядок действий:

- `set retain_order #` надо сохранить порядок атомов
- `create newobj, sele #` создаём новый объект, страховка
- `translate [0,10,0], newobj #` перемещаем
- `rotate x,90,newobj #` вращаем
- `save newfile.pdb, newobj`

Операции по перемещению и вращению можно делать мышкой в режиме editing

Изменение координат отдельных атомов и объектов

`alter_state 1,(pdb1cse),x=x-10.0`
Или `translate [0,10,0], A/100/NZ`

Удаление связей, но не атомов

- Выберите первый атом, ctrl+middle click, выберите второй атом, ctrl+middle click
- И unbond или ctrl+D

Внимание! Координаты атомов не меняются, только исчезает изображение связи

Мутация аминокислот

- Запустите wizard->mutagenesis
- Выберите аминокислоту для мутации
- Справа выберите, на что мутировать
- Выберите ротамет с помощью управления movie
- Закончите процедуру с Apply

Добавление протонов

Работает с молекулами, т.е. объектами

```
create gln, A/101/
```

```
h_add gln
```

Или через меню action объекта.

Есть вероятность, что протоны будут добавлены неверно, если PyMol неправильно угадал валентность тяжёлых атомов.

Суперпозиция в пространстве

Задача достаточно нетривиальная, и есть разные пути:

Белки:

`align, super, fit`

Другое:

`pair_fit`

Желательно указывать родственные атомы в молекулах

`pair_fit (trna10 and resid 10:15 and name P), (ref4 and resid 10:15 and name P)`

Добавление органических фрагментов или а.к.

- С помощью `ctrl+middle click` выделите шариком атом, к которому будет присоединяться фрагмент
- В меню `Build` выберите нужный фрагмент
- С помощью `ctrl+left click` выберите торсионный угол
Или
- Создайте свою молекулу (`ChemSketch`)
- Сохраните как `pkl` в `<pymol_path>/data/chempy/fragments`
- `editor.attach_fragment('pk1','my_fragment_name',11,0)`
11 - это номер атома в фрагменте для связи

Sculpting, что ЭТО?

Это похоже на real-time оптимизатор геометрии, но это алгоритм, который старается сохранить значения длины связей, углов, торсионных углов при изменении координат.

Как запустить sculpting?

У вас достаточно мощный компьютер? Тогда:

- Переводим мышь в режим редактирования
- Выбираем "auto-sculpting" из меню Sculpting
- Выбираем Sculpting из меню Wizard
- Выбираем центральный атом для модификаций
Ctrl-middle-click
- Тянем атом в любую сторону ctrl-left-click-and-drag

Скриптование в PyMol

Возможны как скрипты из команд, так и скрипты на Python

Запуск скриптов из команд:

@ [myfile.pml](#)

Запуск скриптов на питоне:

run [myfile.py](#)

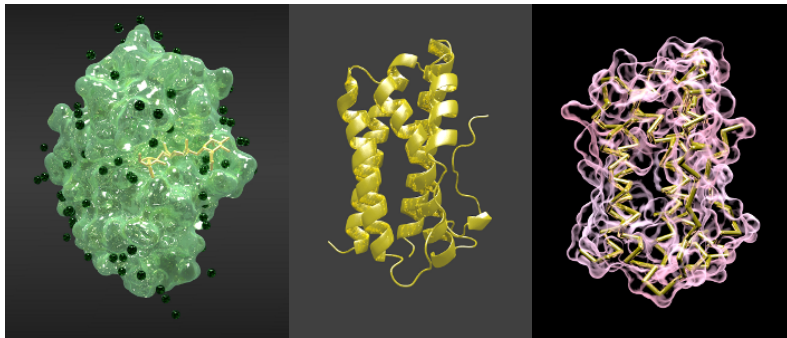
Пример

```
fetch 1c1l, async=0
as lines, n. C+O+N+CA
zoom i. 4+5
mset 1 x1440
mview store
python
for x in range(0,144):
    cmd.frame((10*x)+1)
    cmd.zoom("n. CA and i. "+str(x) + "+" + str(x+1))
    cmd.mview("store")
python end
frame 288
mview store
mview reinterpolate
```


Результат



Объекты из PyMol можно использовать в разных 3D программах





Анимация структуры в Blender

