

Пятигорский медико-фармацевтический институт – филиал ГБОУ ВПО
ВолГМУ Минздрава России

ЛЕКЦИЯ

«Создание трехмерной компьютерной модели
молекулы в программе Hyper Chem»

Автор: к.ф.н. Глушко Александр Алексеевич

Пятигорск 2012 г.

Создание трехмерной компьютерной модели молекулы

Исходной информацией для создания модели молекулы обычно является структурная формула. В итоге получается трехмерная молекулярная структура или набор конформаций молекулы. В создании такой модели есть ряд этапов:

1. Перенос структурной формулы молекулы в компьютерную программу
2. Построение грубой трехмерной модели из двухмерной.
3. Проверка правильности ввода структуры (стереоизомерия, заместители, связи)
4. Оптимизация геометрии
5. Конформационный анализ
6. Сохранение полученных молекулярных структур в файл

Ввод химической структуры.

В программе Hyper Chem ввод структуры возможен путем открытия файла, копированием через буфер, либо ручным редактированием. Программа поддерживает следующие типы химических файлов:

- *.hin (основной формат Hyper Chem)
- *.ent (формат базы данных BrookHaven Protein DataBase)
- *.skc (редактор химических формул Isis Draw)
- *.mol (MDL формат)
- *.zmt (формат программы MOPAC)
- *.ml2 (формат Tripos)
- *.chm (редактор химических формул Chem Draw)

Возможно также копирование структур через буфер, однако при данном способе иногда могут возникать затруднения (вставка не происходит).


Программа HyperChem имеет широкие возможности редактирования трехмерных химических структур. Для выбора атома используется первая кнопка на панели управления (рис. 1) , при двойном щелчке по ней происходит вызов панели выбора элементов (рис. 2):



Рис 1. Панель управления программы Hyper Chem.

На данной панели можно выбрать необходимый элемент. Затем путем нажатия левой кнопки мыши можно устанавливать атомы выбранного

элемента в рабочей области, а также связывать их ковалентной связью протягиванием связи между атомами при помощи движения мыши с нажатой левой кнопкой от атома к атому.

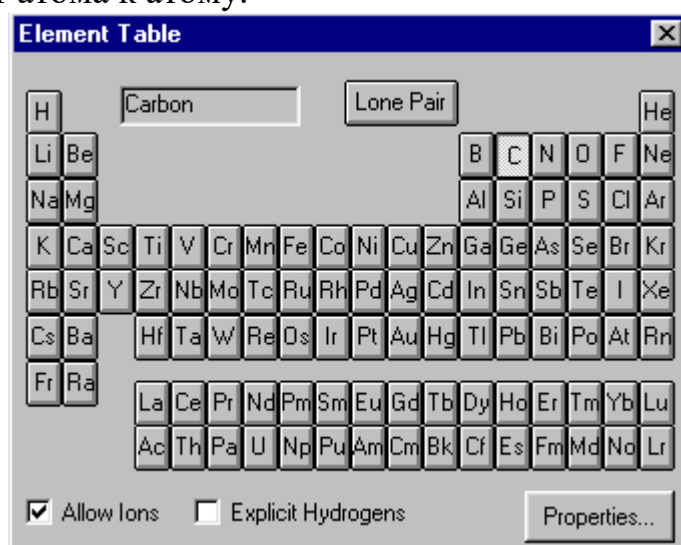



Рис 2. Панель выбора элементов (выбран углерод).

Важно учитывать, что редактирование происходит в плоскости экрана. При этом возможно создание плоской молекулы. Например, выберите углерод в панели выбора элемента и постройте шестиатомный цикл. Теперь при помощи клавиши  на панели управления перейдите в режим вращения. Теперь у вас есть возможность с помощью мыши (при нажатой левой кнопке) вращать структуру в пространстве. В HyperChem есть средство, позволяющее быстро построить приближенную модель молекулы исходя из плоской структуры без включенных в нее атомов водорода. Вызовите данную функцию через меню Build->Add H & Model Build. При этом к атомам углерода в соответствии с валентностями будут добавлены атомы водорода, а длины связей и валентные углы установлены в соответствии со строением молекулы метана (рис. 3).

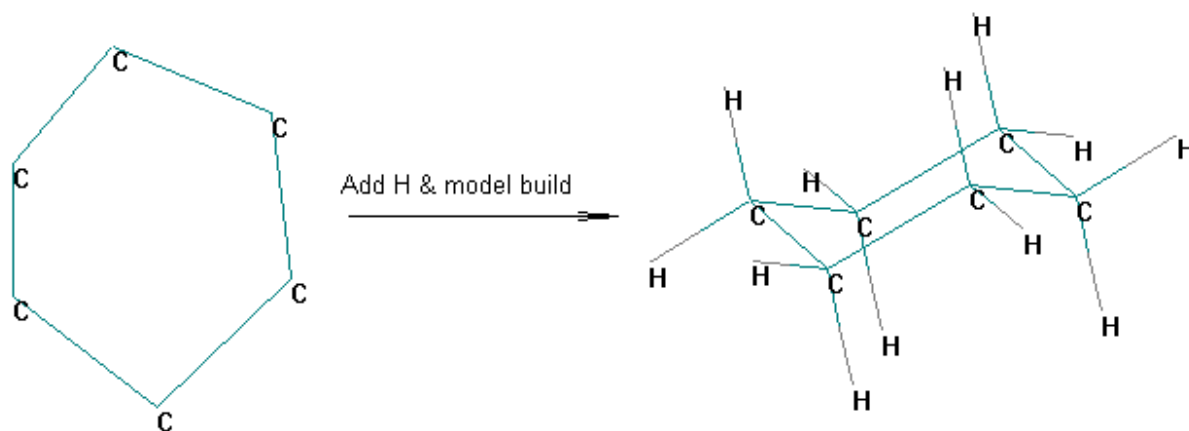


Рис 3. Построение модели молекулы циклогексана.

Получившаяся структура является весьма грубым приближением. Для получения более точной геометрии молекулы требуется проведение оптимизации геометрии. Использование процедуры Add H & Model Build не рекомендуется для крупных молекул. Рекомендуется аккуратно пользоваться данной процедурой в случае наличия хиральных атомов в молекуле, т.к. при этом часто нарушается стереоизомерия. После ее проведения обязательно необходима внимательная проверка всех хиральных атомов в молекуле.

Другой способ ввода данных для программы HyperChem заключается в использовании уже созданных при помощи программы Isis Draw плоских моделей молекул. При этом можно переносить структуры через буфер обмена, либо открывать их в виде файлов.

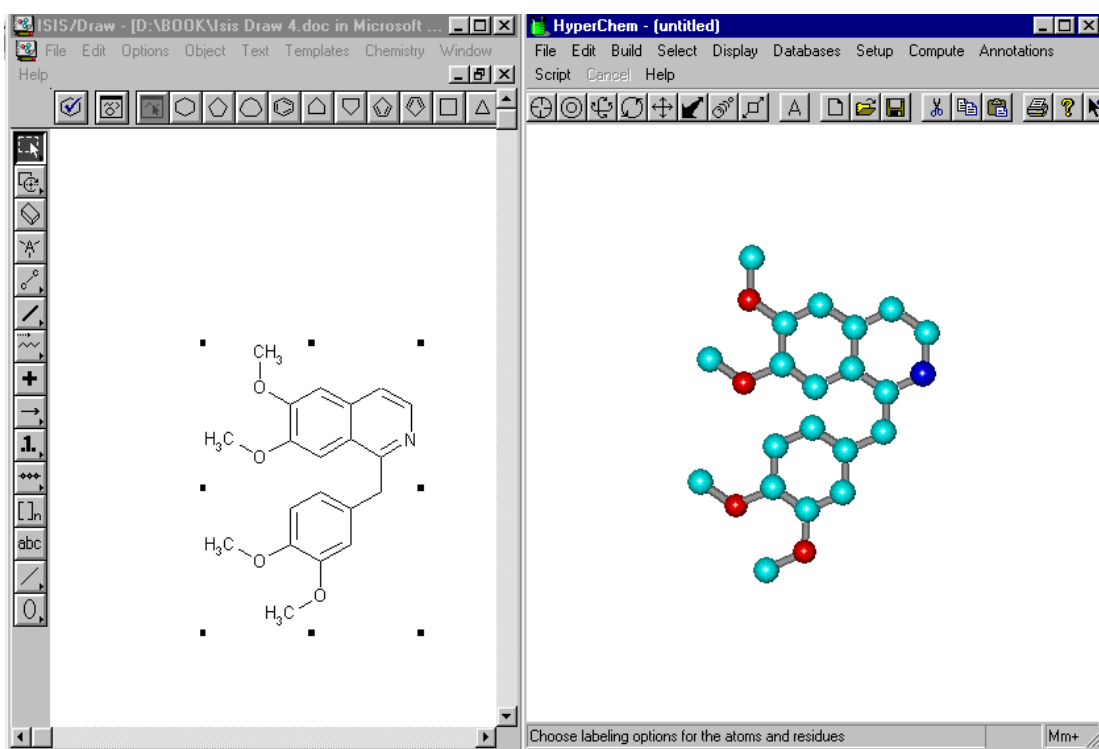


Рис. 4. Пренос плоской структуры из Isis Draw в Hyper Chem.

Теперь структуру можно перевести в трехмерный формат, для чего используется пункт меню Build->Add H & Model Build (рис. 5). Однако данную команду следует использовать с осторожностью, т.к. в случае молекул со стереоизомерией, возможно ее нарушение. Процедура построения модели совершенно не учитывает стабильность конформации и в результате могут получаться весьма напряженные конформации молекул.

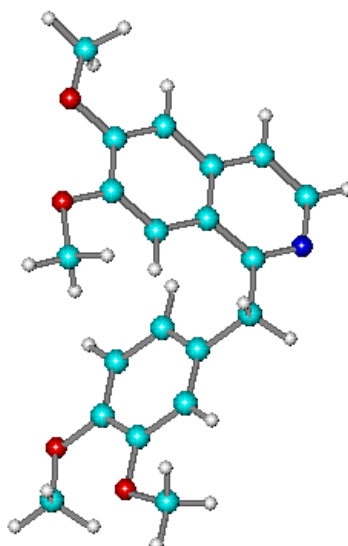


Рис. 5. Трехмерная структура папаверина (результат действия процедуры Add H & Model Build).

После того, как создание трехмерной модели завершено, структуру необходимо сохранить в формате *.hin.

Оптимизация геометрии

Все расчеты проводятся через меню Compute. В меню Compute выберите Geometry Optimization. При этом возникнет форма изображенная на рис. 6:

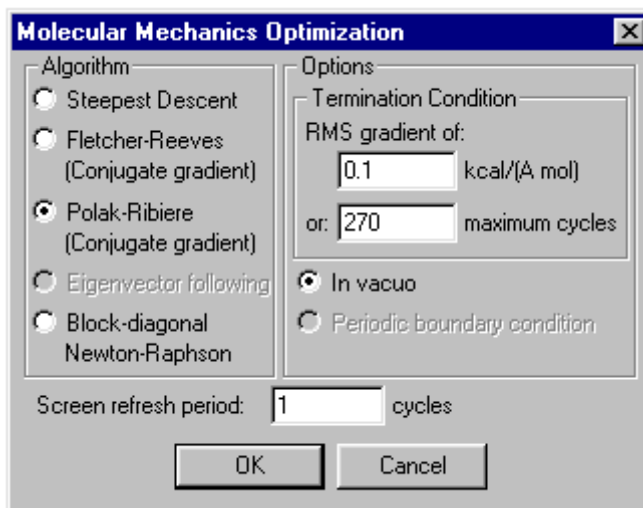


Рис. 6. Параметры оптимизации геометрии.

В рамке Algorithm возможен выбор алгоритма оптимизации. От выбора алгоритма зависит скорость и успешное завершение оптимизации геометрии.

Steepest Descent – метод наискорейшего спуска, позволяет оптимизировать достаточно быстро структуры находящиеся далеко от точки равновесия (минимума потенциальной энергии), однако не годится для точной оптимизации.

Алгоритмы **Fletcher-Reeves** и **Polak-Ribiere**, являются более интеллектуальными с точки зрения поиска минимума потенциальной энергии и позволяют значительно быстрее найти равновесную точку.

Block-diagonal Newton-Raphson – блок-диагональный метод Ньютона-Рафсона использует для оптимизации матрицу вторых производных, что позволяет достигать весьма высокой точности расчета, которая недостижима для вышеприведенных методов. Однако данный метод оптимизации более длительный, и целесообразно переходить к нему после применения методов **Fletcher-Reeves** и **Polak-Ribiere**, если необходима очень высокая точность.

В поле **RMS gradient** задается порог градиента. В случае падения градиента ниже заданного порога, оптимизация завершается.

Возможно задание максимального количества циклов оптимизации в поле **maximum cycles**, обычно этот параметр устанавливается автоматически исходя из анализа структуры.

Screen refresh period – период обновления изображения – количество циклов оптимизации, в течении которого обновляется изображение на экране. Параметр позволяет повысить скорость оптимизации в случае большого количества циклов, т.к. экономит машинное время, затрачиваемое на прорисовку структуры.


Ознакомившись с диалоговым окном оптимизации выберите метод **Steepest Descent**, если он не установлен по умолчанию. Установите градиент 0.1 ккал/моль, и нажмите «ОК».

Начнется процесс оптимизации, однако он будет идти весьма медленно и неэффективно. Внизу экрана вы можете наблюдать как меняется энергия системы и градиент. В меню нажмите «Cancel» и снова вызовите **Compute->Geometry optimization...** Выберите метод **Polak-Ribiere**. Сновапустите оптимизацию. Весьма быстро будет достигнут пороговый градиент 0.1 и процесс завершится. Значительно дольше происходит оптимизация с использованием метода **Ньютона-Рафсона**, однако в ряде случаев метод позволяет достигать величин градиента, недостижимых с использованием метода **Polak-Ribiere**.

Конформационный анализ.

Результат оптимизации зависит от стартовой геометрии молекулы и приводит к локальному минимуму потенциальной энергии, однако в случае, если молекула имеет большое количество конформаций, процесс поиска глобального минимума потенциальной энергии (наиболее стабильной конформации) значительно усложняется.

В случае простых молекул возможен перебор всех конформаций. Рассмотрим простой пример. Создадим две модели молекулы бутана (рис. 7) – анти- и гош-конформеры. Установите силовое поле mm+. Сначала введите одну из конформаций (например анти-), затем проведите оптимизацию геометрии. Произведите расчет энергии молекулы в одной точке (**Compute->Single Point**) и запишите ее значение. Затем, перейдите в режим выделения,

нажав на кнопку панели управления . выделите четыре связи, которыми образован двугранный угол (C-C-C-C) и измените его значение на 0 (Edit->Set Bond Torsion) – при этом получится стартовая геометрия для гош-конформации, затем (ВАЖНО), снимите выделение нажав правую кнопку мыши на темном фоне графического окна. Проведите оптимизацию геометрии.

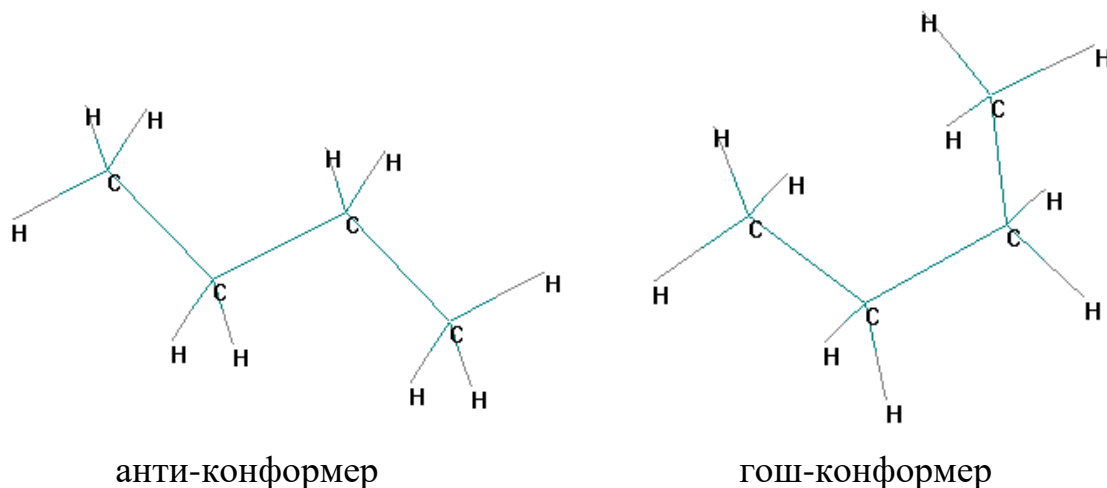


Рис. 7. Две конформации молекулы бутана.

В случае использования силового поля mm+ энергии составляют:

3.0377 для гош-конформации;

2.1745 для анти-конформации.

Следовательно можно сделать вывод, что анти-конформер стабильнее чем гош-конформер.

Существует весьма наглядный способ визуализации зависимости потенциальной энергии от двугранного угла – **построение потенциальной кривой**. Выделите двугранный угол C-C-C-C и Запустите построение потенциальной кривой (Compute->Potential...), при этом возникнет меню выбора параметров построения потенциальной кривой (Рис. 8).

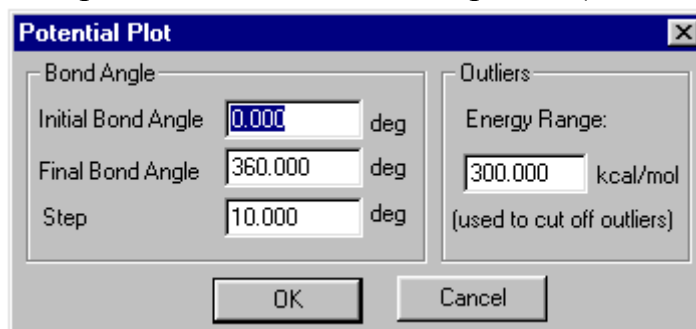


Рис.8. Параметры построения потенциальной кривой.

Установите параметры Initial Bond Angle (начальное значение угла) и Final Bond Angle (конечное значение угла) 0 и 360 градусов соответственно, а Step (шаг) 10 градусов. Запустите процедуру нажав на «ОК». При этом по окончании вращения будет построена потенциальная кривая (рис. 9), на которой виден минимум потенциальной энергии для вращения вокруг

данного двугранного угла. Сравните полученный угол минимальной энергии с ранее полученным результатом. Какой конформации бутана он отвечает, гош- или анти- ?

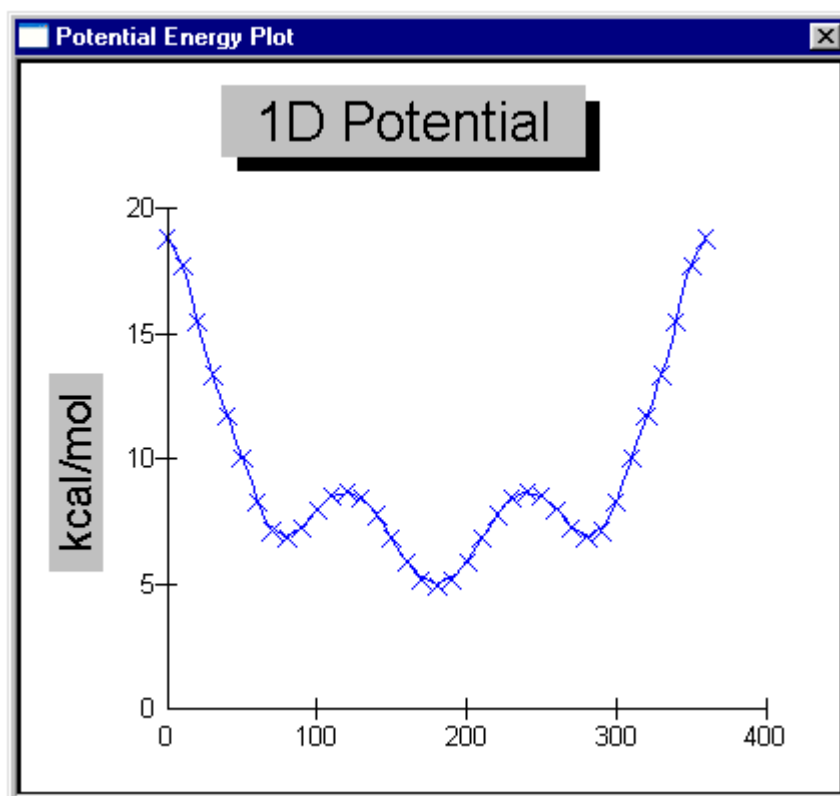


Рис. 9. Потенциальная кривая молекулы бутана.

Однако в ряде случаев использование потенциальной кривой для оценки оптимального угла не рекомендуется. Например, в случае, когда при повороте возникают значительные стерические напряжения и тем более когда возникает перекрывание фрагментов химической структуры. В таком случае используется автоматический конформационный анализ. Для этого в Nuser Chem 6.01 предусмотрена специальная процедура конформационного анализа.

Конформационный анализ.

Рассмотрим конформационный анализ на примере глутаминовой кислоты (рис. 10).

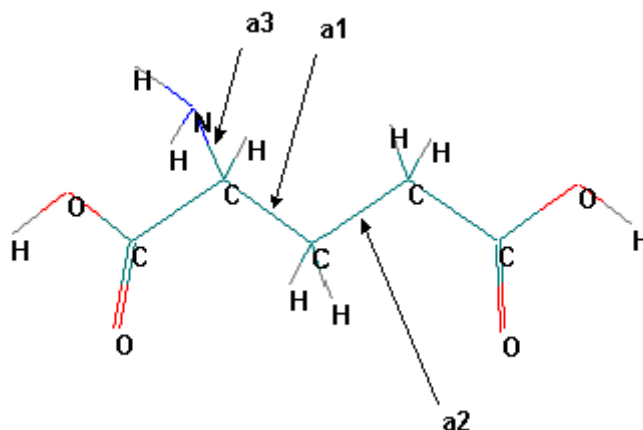


Рис. 10. L-глутаминовая кислота.

Прежде всего необходимо указать, какие валентные углы будут варьироваться в процессе конформационного анализа. К сожалению это реализовано в Nuser Chem 6.01 не очень удобно. Прежде, чем приступить к анализу необходимо выделить и присвоить каждому из углов имя. Выделите три связи, соответствующие валентному углу a1 (для вращения вокруг связи C-C), путем выделения двух прилежащих связей и связи указанной на рис. 8 стрелкой a1. Затем выберите пункт меню Select->Name Selection... Укажите имя угла a1. Аналогично присвойте имена двум оставшимся углам (сначала выделяя а затем присваивая имя).

Теперь запустите конформационный анализ Compute->Conformational Search... (рис. 11).

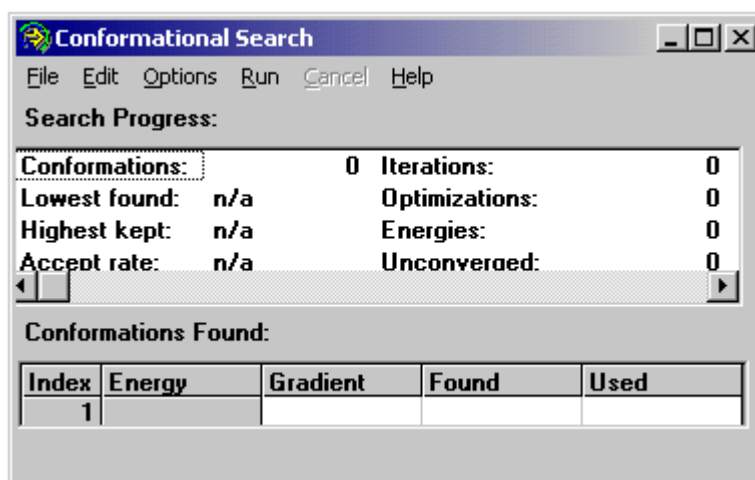


Рис 11. Окно конформационного анализа.

В меню конформационного анализа выберите Options->Choose torsions... (рис. 12). В левой колонке (Named Torsion Angles) должен быть список выбранных углов (a1,a2,a3). Выделите их все и нажмите кнопку Add. Углы будут добавлены в колонку Torsions to Vary (углы, которые будут варьироваться) (рис. 12).

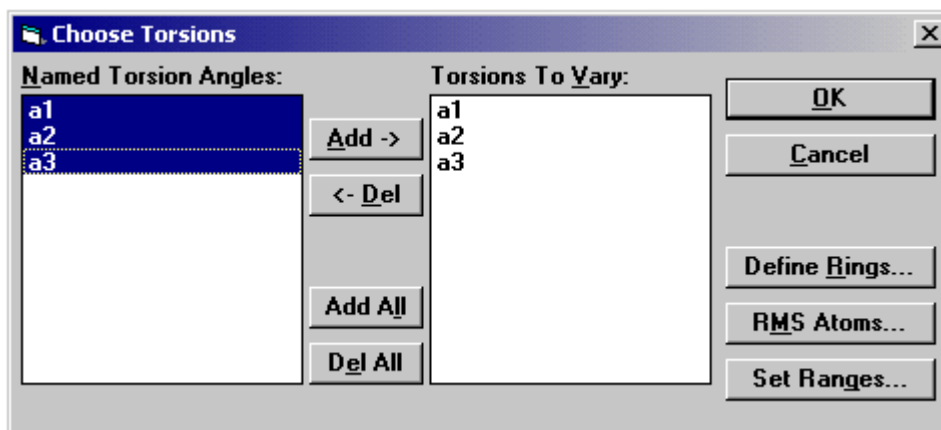


Рис. 12. Выбор торсионных углов для конформационного анализа.

Варьирование углов используется для получения различных стартовых точек для оптимизации. Углы могут варьироваться с некоторым шагом и в определенном диапазоне, которые задаются в Options->Setup Ranges... (рис. 13).

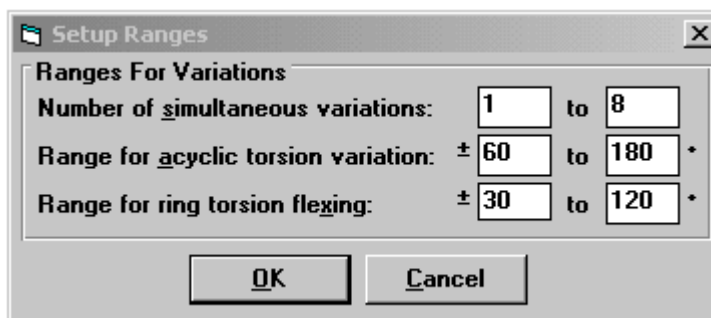


Рис. 13. Выбор диапазона и шага варьирования валентных углов.

Для выбора метода конформационного поиска выберите пункт меню Options->Search Method... (рис. 14).

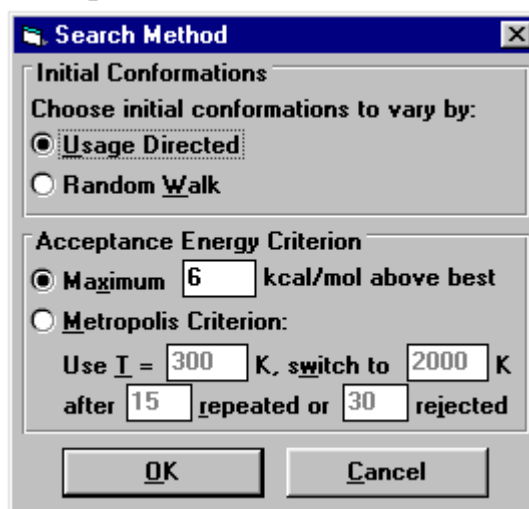


Рис 14. Выбор метода конформационного поиска и настройка параметров. В рамке Initial Conformations – выбор начальных конформаций: **Usage Direction** – по мере использования.

Random Walk – случайный выбор (повышает вероятность полного охвата всего конформационного пространства).

В рамке **Acceptance Energy Criterion** критерий выбора конформации в соответствии с ее энергией:

Maximum – максимальное отклонение энергии от энергии конформации с низшей энергией, в случае превышения конформация не используется.

Metropolis criterion – вероятностный выбор конформации в соответствии с распределением Больцмана (задается температура).

Перед запуском поиска в меню File активизируйте опцию **Autosave**. При этом вся информация в процессе поиска будет записываться в файл. Запустите конформационного поиска Run->Start. Создайте при этом файл конформационного анализа, в котором будет сохраняться вся информация о ходе анализа, в том числе и структуры различных конформеров, которые можно в дальнейшем вызывать.

В процессе анализа таблица со свойствами конформаций будет расширяться по мере нахождения новых конформеров (рис. 15). При этом ее строки отсортировываются по энергии. Таблица имеет следующие колонки:

Energy – энергия конформера.

Gradient – градиент в конце оптимизации.

Found – сколько раз данная конформация была найдена.

Used – сколько раз данная конформация была использована для генерации новых конформаций.

Далее следуют колонки со значениями валентных углов, подлежащих варьированию.

Conformational Search - SEARCH.HCS

File Edit Options Run Cancel Help

Search Progress:

Conformations: 33 Iterations:
 Lowest found: -10.6646 Optimizations:
 Highest kept: -8.174095 Energies:
 Accept rate: 0.6923077 Unconverged:

Conformations Found:

Index	Energy	Gradient	Found	Used
1	-10.6646	0.0935579	1	3
2	-10.22056	7.967126E-0	2	2
3	-10.19221	9.874085E-0	3	5
4	-10.14444	9.068906E-0	3	4
5	-9.94322	9.110635E-0	2	5
6	-9.895847	8.045439E-0	2	2
7	-9.888053	8.828578E-0	2	14
8	-9.825215	9.585679E-0	1	5
9	-9.79701	8.794494E-0	6	4
10	-9.765873	7.574522E-0	2	2
11	-9.730405	6.669938E-0	2	5
12	-9.711513	9.542049E-0	1	1
13	-9.699095	0.0931299	3	2
14	-9.659108	9.845752E-0	1	2

Рис. 15. Результаты конформаионного анализа.

Основные параметры приведенные в поле Search Progress.

Conformations – количество конформаций.

Lowest found – конформация с наименьшей энергией

Highest kept – конформация с наибольшей энергией

Остановить процедуру поиска можно путем нажатия Cancel. Однако можно задать количество конформаций в меню Options->Limits. После остановки анализа его можно возобновить Run->Resume. Теперь после завершения можно посмотреть структуры полученных конформаций. Для этого нужно выделить соответствующую строку таблицы выбрать пункт меню Edit->Put Molecule, либо нажать сочетание клавиш Ctrl+U.

Искусство конформационного поиска.

Для малых молекул конформационный поиск возможно проводить и вручную, т. к. количество не велико, а наиболее стабильная конформация бывает обычно очевидна. Тем не менее, для проведения такого поиска требуются определенные навыки, и порой выбор бывает чисто интуитивным. Безусловно, здесь присутствует элемент субъективности, но такой подход позволяет значительно ускорить обработку больших массивов молекул, хотя это порой и требует определенного терпения. Однако можно предложить ряд рекомендаций, которые могут оказаться полезными.

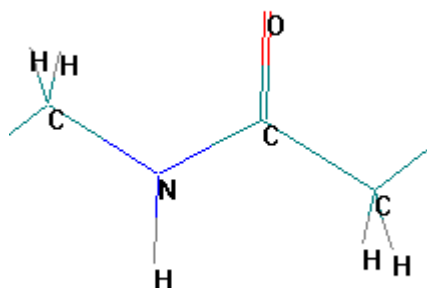


Рис. 18. Пептидная связь.

4. В циклических молекулах обычно меньше энергия конформаций с экваториальным расположением заместителей. Однако при этом важно соблюдать стереоизомерию молекулы. Например, чтобы изменить положение заместителя с аксиального на экваториальное:

- 1) Удалите атом водорода у его основания и поместите его на противоположную сторону (рис. 19).
- 2) Запустите оптимизацию геометрии.

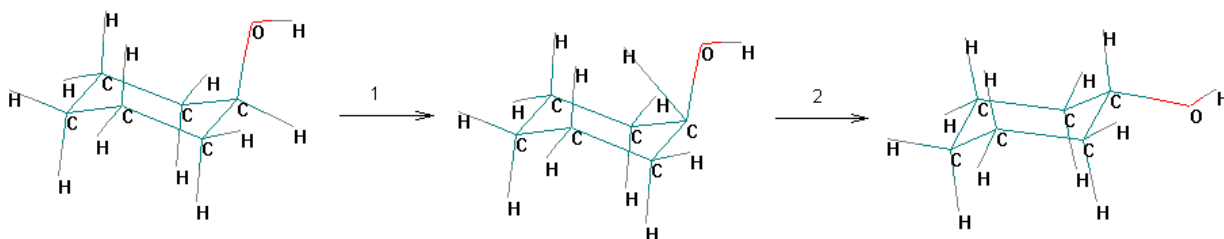


Рис. 19. Перевод заместителя из аксиального в экваториальное положение.

5. Для циклических молекул конформация кресло обычно стабильнее. В программе. Для перевода циклогексана из конформации «ванна» в конформацию «кресло» можно воспользоваться следующим приемом (рис. 20).

- 1) Добавьте в молекулу атом тяжелого элемента (у него достаточно большая длина связи), как показано на рис. 20;
- 2) Проведите оптимизацию геометрии (полного завершения дожидаться необязательно);
- 3) Сотрите атом тяжелого элемента и добавьте атомы водорода Build->Add Hydrogens. Запустите оптимизацию геометрии.

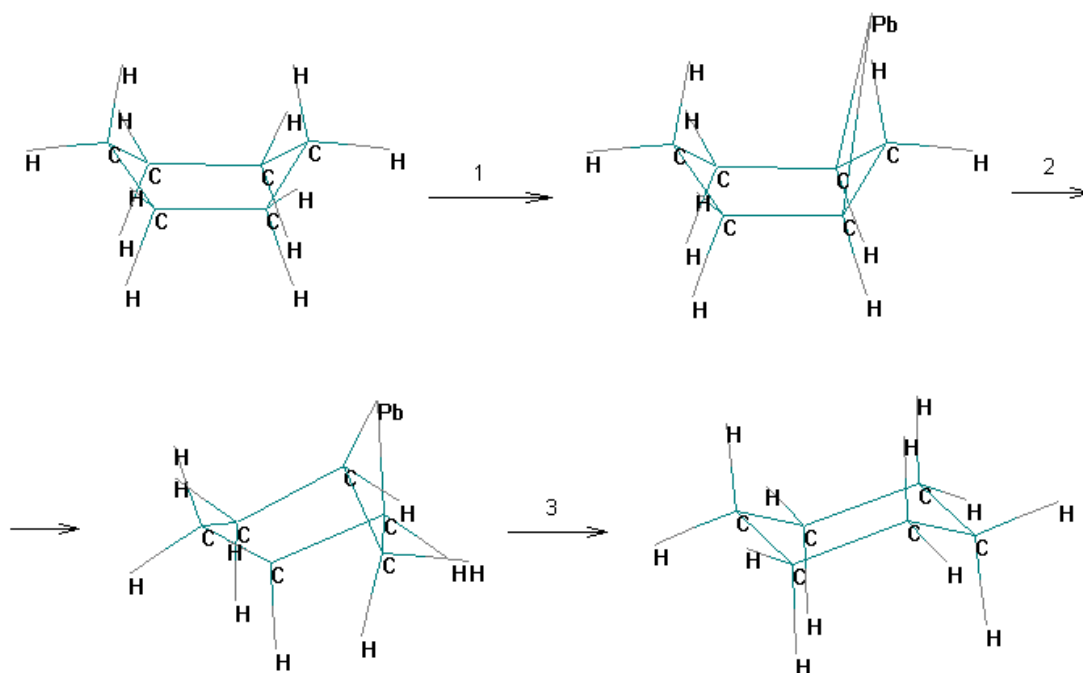


Рис. 20. Изменение конформации циклогексана.

Изменить конформацию можно и простым переносом атома углерода, однако данный метод может оказаться весьма удобным в случае наличия сложных заместителей. В последнем случае, даже при переносе атома углерода, молекула вероятнее всего снова вернется в прежнюю конформацию.