

## 構造生物学実習 (NMR による構造計算と構造評価)

### CNS / CNX “Crystallography & NMR Systems” Axel T Brunger’s Laboratory, YALE UNIVERSITY

#### 1. 分子動力学と NMR 構造計算

##### 1.1 溶液 NMR から得られる情報

###### 化学構造の情報

- ・ 原子団・官能基・化学シフト
- ・ 化学結合の情報・結合定数(J)
- ・ 水素結合の情報・HD 交換/結合定数

###### 立体構造の情報

- ・ プロトン間の距離・NOE
- ・ 二面角・結合定数(J)
- ・ ペプチド主鎖の二面角・化学シフト
- ・ 分子の運動性・緩和現象
- ・ 分子の磁場への配向・残余双極子
- ・ 常磁性金属イオンからの距離

##### 1.2 分子動力学(MD)

コンピュータの中に蛋白質分子を仮想的に構築して、それを物理法則にしたがって動かしてみるための計算(詳細は情報科学特論などの講義を参照のこと)。とはいえ、仮想的に分子を構築してやるにしても、実際の化学結合の性質(たとえば単結合は回転可能だが二重結合は回転しない、原子と原子は VdW 半径より近くには近づかない、正と負の電荷は引き合、など)の現実の性質をなるべく正確に反映していなければ意味がない。そうした「ルール」を分子の持つエネルギーと各原子(団)の距離や角度の関係式で既述したものを「力場」という。力場はエネルギーを記述する関数(近似式)の線形結合である。種々の近似式の組み合わせとなっているため、計算の利用目的によって適宜使い分ける必要がある。

NMR/X-ray の立体構造決定に用いられるプログラムでは、計算効率を優先させる代わりに、精密さを犠牲にした力場を用いることが多い。なぜな

ら計算の目的が、コンピュータ内で分子シミュレーションを行うことではなく、測定によって得られた実験値を矛盾なく満たす分子モデル(=座標)を得ることだからである。また、特別な場合を除いて、構造計算は真空中で行い、水分子は含まれないで行う。

具体的には bond 長をほとんど固定にしながら、peptide 主鎖の可動な単結合のまわりの二面角のみを回転させながら、NMR や X 線のデータを満たす立体構造を求めてゆく。

ここで更に注意を要するのは、これまでにいろいろな「力場」が論文により発表され、またプログラムに実装されているが、どの力場がどの種の計算機実験に向いているのか、また目的とする計算に、その力場の精度が適しているのか、吟味しなければならないことである。残念ながら、立体構造決定の専門家は、計算機科学の専門家ではない。したがって力場の不備によって、得られた実験データが正しかったとしても、誤った(または不正確な)立体構造を提出してしまうことがある。それを避けるためにも、とくに計算科学の専門家と共同研究をしている場合を除いては、業界標準ともいべき構造計算プログラムを使用して立体構造計算を行う場合が多い。

##### 1.3 distance geometry(DG)と simulated annealing (SA)

DG はもともとは NOE によって得られる二つのプロトン間の距離の情報に、ポリペプチドの化学構造に基づく結合距離の情報を加味・集積して、それを満たすプロトンの 3 次元座標を行列計算をもとに求めるアルゴリズムのことであった。しかし、意味のある蛋白質の構造を計算するためにア

ミノ酸残基の結合角に無理のない解を求める必要があり、いずれにせよ MD 計算の力場を用いることになる。そのため、現在一般的に行われている DG は、MD の計算を行いながら、NOE や J から得られた NMR の情報を偽エネルギーとして盛り込んで、NMR の情報を全て満たしたときに分子の total のエネルギーが低くなるように分子を「動かし」、最終的な構造を得る、というプロセスをとる。この方式のことを **restrained molecular dynamics** (制限付き分子動力学計算) と呼ぶ。また、計算に盛り込む NMR の情報のことは距離制限情報または角度制限情報と呼び、**restraint / constraint** と呼ぶ。

標準的な MD のプログラムであれば、それに適当な形式で距離・角度制限情報を付加することで、原理的にはどの MD プログラムであっても NMR の構造解析に用いることが可能である。

一方、SA は「焼きなまし法」「徐冷法」と呼ばれていて、前述の制限的分子動力学計算におけるテクニックの一つである。コンピュータの力場の中で、「温度」とは各原子のもつ化学エネルギーと各原子のもつランダムな運動エネルギーの和として記述されている。実際にプラスチック製の分子モデルで、分子模型を組み立てることを想像して欲しい。正しい順番で組み上げていかないと、側鎖同士がぶつかり合ってしまう、モデルを最終的によい形にもっていくことができない。もし途中でどこかが引っかかってしまった場合には、なんとかすり抜けさせなければならない。ということは、モデルの大まかな形が決定するまでは、結合角や結合長が自由に伸び縮みしたほうがモデル組み立てがやりやすい、という道理である。それを計算機の中で実現するために、計算初期では 2000-4000K に相当する温度で **simulation** を行い、計算終期で徐々に室温に相当する温度まで下げてくるというテクニックが、NMR・X 線の構造決定の際に効率が良いことが、多くの例によって示されている。

## 1.4 CNS

CNS は Yale 大学 AT Brunger の研究室で開発された分子動力学計算エンジン X-PLOR の発展形として、特に **X 線構造解析** と **NMR 構造決定** のパッケージとして提供されていて、両者の構造決定において頻繁に利用される。

CNS は CHARMm という汎用生体分子用分子動

力学計算+力場を用いて、分子動力学計算を行うプログラムである。X 線や NMR の構造精密化のために「**simulated annealing**」法 (分子動力学を高温で行い、徐々に温度を下げるという手法) に特化した X-PLOR というソフトに発展。その後さらに進化を遂げて現在の形となった。

前述のように温度の上げ下げや実験データ由来のエネルギーの比率の増減を自在に制御できるような、専用の言語を備えている。しかし、日常的に汎用するような計算処理のためには、CNS 言語を覚える必要はなく、WWW-browser を通じて **input file** を簡便に作成できるようになっているのが、最大の特徴である。

## 1.5 CNS を用いて立体構造計算をするときの流れ

今回の実習では以下の繰り返しを非常に頻繁に行うので、やり方を覚えておいてほしい。なお、コマンドによってリダイレクト記号「</>」を使う場合と使わない場合があるので要注意。

1. `cns_edit hoge.inp`  
↓
2. `[SAVE UPDATED FILES]`  
↓
3. `cns < hoge.inp > hoge.out`  
↓
4. `less hoge.out`

## 2. 計算を始める前に

### 2.0 各自 Linux で端末にログインする。

```
[nmr_@pc101] cd [return]
```

これでホームディレクトリに移動した。

### 2.1 cns で計算する directory を準備する。

```
[nmr_@pc101] mkdir CNS [return]
```

```
[nmr_@pc101] cd CNS [return]
```

### 2.2 実習で使うファイルを自分のディレクトリにコピーする

```
[nmr_@pc101] cp /home/share/data/cns/*.inp _ [return] ##period を忘れずに
```

```
[nmr_@pc101] cp /home/share/data/cns/er 2* _ [return] ##period を忘れずに
```

```
[nmr_@pc101] cp
/home/share/data/cns/molmol.mac _ [return]
(## period を忘れずに)
```

## 2.3 cns が動くかどうかチェック

```
[nmr_@pc101] cns [return]
```

CNSsolve> という prompt が返ってきて  
入力待ちの状態になる。

help [return]と打ち込んでみる

説明が表示される。

**stop**と入力して cns を終了。

## 2.4 cns 基本コマンド

**cns\_info** : cns 関連コマンドの基本的な使い方を表示

**cns\_web** :cns のオンラインマニュアルと WWW インタフェースを表示するために Netscape 画面を呼び出す。

**cns\_solve**:cns 計算の本体

**cns** : 同上

**cns\_edit** : input file を WWW browser 上で編集するためのインタフェース。

**cns\_header** : header 表示

## 3. データの準備

\*\*は絶対必要なファイル

\*は無くともよいファイル

### 3.1 cns で計算するための主な入力ファイル。

配列ファイル\*\*

→ amino 酸 3 文字表記の配列ファイル

【例 1・er2.seq】

距離制限情報ファイル\*\*

【例 2・er2\_noe.tbl】

角度制限情報ファイル\*

【例 3・er2\_aco.tbl】

水素結合情報のファイル\*

【例 4・test\_hb.tbl】

そのほか

J 結合情報のファイル\*

【詳細はオンラインマニュアル参照】

残余双極子相互作用のファイル\*

<sup>13</sup>C の化学シフトのファイル\*

### 3.2 自分で独自に解析を始めたときは、上記のフ

ァイルを【例】のフォーマットに従って editor  
で作成する。

適当な script を用いて作成してもよい。例えば Perl 言語や Awk, C 言語などで小さいプログラムを自作する。(例えば配列ファイルを WWW 上からダウンロードしてきた FASTA 形式のファイルから作る場合など。NMR や X 線の構造計算では、このようなファイルの形式を変更する作業が割と多い。editor で手作業で行うのもよいが、打ち間違いなどがあると悲劇である)。

NMRDraw / Pipp などのスペクトル解析ソフトウェアで、帰属・解析して作成した NOESY peak table を「NOE 強度→プロトン間距離」に換算して、CNS 形式のファイルに変更する C のプログラムや Perl script がいくつかの NMR lab. より無料で入手可能なことが多い。

(例: トロント大学 [http://nmr.uhnres.utoronto.ca/ikura/public\\_files/pseudo\\_corr/pseudo\\_corr.tar](http://nmr.uhnres.utoronto.ca/ikura/public_files/pseudo_corr/pseudo_corr.tar))

このようなフリーで配布されているスクリプト類は、マニュアルなどが完備していない場合が多いので、スクリプトを解読しながら使わなければならない。経験者に教えてもらうのが最も手っ取り早い。もちろん自作してもかまわない。

今回の実習ではあらかじめ作成しておいたファイルを使うことにする。

## 4. 入力ファイル(input file)の準備

### 4.1 配列ファイルより分子構造ファイル(mtf)を作成する

必要な情報	配列ファイル	【er2.seq】
出力(作成)する mtf file の名前		【er2.mtf】

■まずこのファイルを作成する

■ファイルの名前を決める

**SS-bond** の有無 [true]  
【5-20, 12-37, 17-28】

**generate parameters hydrogen flag**  
[true]

使用するパラメータファイル・ライブラリ

【CNS\_TOPPAR:protein-allhdg.top】

【CNS\_TOPPAR:protein.link】

【CNS\_TOPPAR:protein-allhdg.param】

■力場などを指定できる

その他の付帯情報も必要

糖鎖の有無                   なし\*  
水の情報の有無               なし\*  
Hの情報の有無               なし\*  
その他オプション多数 (ほとんど省略可能)

(1)[nmr\_@pc101]

**cns\_edit generate\_seq.inp [return]**

Netscape が立ち上がって、generate\_seq.inp を編集する画面が立ち上がる。

↓

各ボックスに必要事項を入力する  
(前記の変更箇所のみ)

↓

[SAVE UPDATED FILES]ボタンを押す  
[保存する場所・ファイル名・上書きするか  
どうか聞いてくるので[yes]で答える]

↓

[generate\_seq.inp という名前で/home/s02  
/xxx/CNS/に保存する]

(2)[nmr\_@pc101]

**cns < generate\_seq.inp > generate\_seq.out  
[return]**

先に述べたようにもともと cns は interactive な program である。標準入力 (キーボード) から CNS 言語によるコマンドを受け付けて、標準出力にログを返す。

更に実行中に作成された結果は、別のファイルとして保存される。

非常に長いコマンドを決まりきった操作のために何度も入力するのは煩雑であるので、大抵の作業には対応した input file が用意されており、それを「<」リダイレクト記号でファイルから読ませてやる。標準出力に返ってくるログも膨大なので、「>」リダイレクト記号でファイルに保存する。

(3)[nmr\_@pc101] **ls -l [return]**

確認する

正しく er2.mtf が作成されていれば O.K.

(4)[nmr\_@pc101] **more generate\_seq.out**

**[return]** エラーがないか確認する

(この作業はとくに正常に終了しなかった場合に重

要。もしきちんとできていなかった場合は、このファイルの中身を読んで、どこがまずかったか確定せよ。)

#### 4.2 mtf ファイルより初期構造ファイル (er2\_extended.pdb)を作成する

(立体構造計算に必要な、初期座標を作成する。random に伸びた構造を発生させる。)

必要な情報 = 構造ファイル                   **[er2.mtf]**

(1)[nmr\_@pc101] **cns\_edit generate\_extended.inp [return]**

Netscape が立ち上がって、generate\_extend ed.inp を編集する画面が立ち上がる。

各ボックスに必要事項を入力する(下記の変更箇所のみ)

■mtf file を指定

■作成された結果を保存するファイル名  
er2\_extended.pdb

↓

[SAVE UPDATED FILES]ボタンを押す  
[保存する場所・ファイル名・上書きするか  
どうか聞いてくるので[yes]で答える]

(2)[nmr\_@pc101] **cns < generate\_ext  
ended.inp > generate\_extended.out [return]**  
数分かかる。

(3)[nmr\_@pc101] **ls -l [return]**                   確認する

(4)[nmr\_@pc101] **more generate\_exten  
ded.out [return]** エラーがないか確認する  
正しく er2\_extended.pdb が作成されていれば O.K.

#### 4.3 構造計算を行うときのダイアグラムを決定し input file を作成する

##### 構造計算のダイアグラムとは?

CNS の分子動力学計算には NMR の構造計算に適したいくつかのモードがあり、それぞれに対応した input file が用意されている。

■dg\_sa.inp 主鎖のディスタンスジオメトリとシミュレーテッドアニーリング

↑・・・今回の実習ではこれを用いる。

■anneal.inp シミュレーテッドアニーリング。  
直交座標系と二面角(torsion angle)のダイ

ナミクスが選べる。

二面角空間でのダイナミクスは計算速度が速い。

■ **anneal\_cv.inp**            上記の **cross validated version**

■ **ensemble.inp**            距離制限情報を **ensemble** として扱う SA 法。

ホモダイマーの解析などに用いる。

これらのダイアグラムの中では、計算を行う温度やステップ数、NOE や二面角制限のエネルギーの重みを、それぞれ最適な数値に変化させなければ、よい収束構造が得られない。

そのために **input file** を編集して自分の目的に合わせる。

### (1) [nmr\_@pc101] **cns\_edit dg\_sa.inp** [return]

前回と同様に慎重に必要なパラメータを全て入力する

parameter file

CNS\_TOPPAR:protein-allhdg.param

structure file            **[er2.mtf]**

initial pdb file name

**[er2\_extended.pdb]**

distance geometry            ◇true

DGSA                            ◇true

number of regularized file

0

random seed (乱数)

2 0 0 6 0 6 0 8 #適当

trial or accept

◇trial

number of trial / accepted structures ?

30

print accepted structures

◇false

print trial structures            ◇true

generate average structure        ◇false

...

(何行か省略)

...

計算パラメータ(何行か省略)

...

このあと、SA 計算に必要なパラメータを入力するが、デフォルトからの変更はほとんどないので、今回は説明しない。通常、必ず確認する

べきパラメーターは次の 3 点

■ 高温で何ステップ計算を行うか

700~3000

(今回出席番号が偶数の人 : 1000 奇数の人 : 1500)

■ 徐冷のときに何ステップ計算を行うか

700~3000

(今回出席番号が偶数の人 : 1500 奇数の人 : 1000)

■ 最後の最適化の計算ステップ

200~1000 (今回 200)

...

(何行か省略)

...

NMR 実験データの入力パート

NOE distance restraints file (1)

er2\_noe.tbl

(複数のファイルが入力可能なようになっている。今回は 1 つしか用いない)

(数行下)

NOE averaging mode (1)            center

(数行下)

hydrogen bonding                            [skip]

(空欄のまま)

(数行下)

j-coupling                                    [skip]

(数行下)

chemical shift                                [skip]

(数行下)

diffusion restraints                        [skip]

(数行下)

chemical shift                                [skip]

(数行下...かなり下のほうに)

...

other restraints / dihedral angle

er2\_aco.tbl

...

base name for input coordinate

[skip]

(ここに名前を指定すると random ではない構造からでも計算できる。(今回は空白))

...

(一番下のほうに)

base name for output coordinate

pdb/er2

↑できあがりファイルの頭文字

この例ではディレクトリ `pdb` の下に `er2_1.pdb` などの名前でファイルができる。

↓

**[SAVE UPDATED FILES]** ボタンを押す  
[保存する場所・ファイル名・上書きするかどうか聞いてくるので `[yes]` で答える]

**[nmr\_@pc101] ls -l dg\_sa.inp [return]**  
できているかどうか確認してほしい。  
`file` のサイズを周りの人と比べて、極端に少ない場合には `dg_sa.inp` ファイルが正確にできていない可能性が高い。上のステップを繰り返して、もう一度作り直すこと。

## 5. 構造計算を行う

(1) デスクトップよりもう一枚ターミナル・ウィンドウを開く (計算監視用)

そちらのウィンドウでもと忘れずに `"cd"` とする

**[nmr\_@pc101] cd CNS [return]**

(2) 計算結果を入れるディレクトリを作成

**[nmr\_@pc101] mkdir pdb [return]** ←できあがりファイル `er2_**.pdb` はこの下にできる

■■これを忘れてよく `cns` がとまる■■

(3) 計算開始

**[nmr\_@pc101] cns < dg\_sa.inp > dg\_sa.out [return]**

60分くらいかかるので昼食休憩。  
でもそのまゝに

↓

(4) 計算経過を監視する もう一枚のウィンドウの方から

**[nmr\_@pc101] tail -f dg\_sa.out [return]**

\* `tail` はファイルの終わり数行を画面に出すコマンド。

\* `tail -f` で、現在生成している最中の、最新の結果の終わり数行が画面に常にでていくようになる。

\* `[CTRL][C]` で抜ける。

しばし待つ。

## 6. 結果の解釈と評価

通常 NMR の構造計算はランダムな構造から

30~200 個計算し、NMR より与えられた実験情報をもっともよく満たす構造上位 10~20 個を選んで結果とする。

結果を選ぶ基準は

【1】 NOE / 二面角などの NMR 情報をよく満たすこと。

【2】 ファンデルワールス半径でのコンタクトに原子の衝突がないこと。

【3】 原子の各結合長・結合角が統計値と大きくずれないこと。

【4】 1~3 を満たしているかどうかは、それぞれ CNS が出力した PDB ファイルの header 部分に書かれている、

`E(noe)/E(cdih)/E(vdw)/E(bond)/E(angle)` などの計算結果のエネルギー値に反映されるので、それらの総和 `E(total) / E(overall)` がよい指標になる。

今回の実習では 30 個計算した中から 5~10 個の構造を手作業で選ぶ。

(1) **[nmr\_@pc101] cd pdb [return]**

(2) **[nmr\_@pc101] ls -l [return]**

`er2_embedded_**.pdb`

計算途中に出力される初期構造ファイル

`er2_**.pdb`

計算結果の構造ファイル

`er2a_**.pdb`

↑ より `Etotal` の低いものを `cns` が選んだものの。

今回の実習では `er2_**.pdb` を手作業でチェックして `accept` する構造を選ぶ。

(3) **[nmr\_@pc101] more er2\_1.pdb**

とりあえず、ファイルを見てみる 【例 5】

およそ 30 行目くらいに violations (NMR 情報や正しい化学構造に対する違反) の記述がある。

→→→ 【少ないほどよい】

また 50 行目くらいに energy 項 `E(overall)` とか `E(noe)` とかがある。

→→→ 【小さいほどよい】

(4) **[nmr\_@pc101] more ../dg\_sa.out [return]**

どの NOE が violation しているかななどの情報は、`dg_sa.out` に入っているが、長大なファイルなので手作業で確認するのはあまり現実的でない。通常は系統的に violation している NOE を探すために各研

研究室で自作の Perl script を用意して使用している例が多い。

今回の実習では、violation している NOE の情報の編集は行わない。

(5) 自作 Perl script を用いて pdb 構造を取捨選択する。

```
[nmr_@pc101] analysis.pl er2 [return]
```

または

```
[nmr_@pc101] analysis.pl er2 | sort -nb +5.0 [return]
```

(自作の perl script の例 「analysis.pl」 付録【例 6】)

このようにすれば、overall のエネルギーを指標に最適の構造を簡単に選ぶことができる。

overall のエネルギーを指標に上位 5~10 個を選ぶ。Energy の値にギャップがあるところで切る。

どの pdb を採用するか、ノートなどに書きとめておくこと。

## 7. 結果の表示・・(MOLMOL)

分子表示プログラム MOLMOL を用いて結果を表示する。NMR で計算を行う場合は、「同じ化学構造をもった蛋白質の座標」複数を重ね合わせするという機能を多用する。そのため、Rasmol や SwissPDB Viewer や InsightII といったソフトよりも MOLMOL が便利である。

MOLMOL のもう一つ便利な機能は、決まりきった操作を頻繁に行う場合に、そのコマンドの集まりをマクロファイル(一種のスクリプト)にして保存しておくことである。今回は

MOLMOL の使い方は下記のホームページを参考にしてください。

<http://homepage.mac.com/yabyab/howtomolmol/molmol0.html>

ここでは、まず MOLMOL のマクロを編集して、重ねあわせを行う。(大文字・小文字に注意)

(0) まず cns の計算結果のあるディレクトリに molmol.mac をコピーしてくる。

```
[nmr_@pc101]
```

```
cp /home/share/data/cns/molmol.mac
```

```
molmol.mac [return]
```

(1) エディターで各自、「molmol.mac」を編集。

```
-----molmol.mac-----
```

```
InitAll yes
```

```
ReadPdb er2_1.pdb
```

#順不同でよいので上で選んだ構造のファイル名をならべる。

```
ReadPdb er2_17.pdb
```

```
ReadPdb er2_21.pdb
```

```
ReadPdb er2_11.pdb
```

```
ReadPdb er2_9.pdb
```

```
.....
```

```
SelectAtom ''
```

#↑ (シングルコーテーションを 2 回続ける、空白ははさまない)

```
SelectAtom ' :3-35 & bb '
```

```
-----ここまで
```

(2) [nmr\_@pc101] molmol [return]

MOLMOL を起動する。

(3) XMacUser molmol.mac [return]

MOLMOL のコマンド入力行のところにカーソルを持っていく。

実際には大文字で XMU molmol.mac と入力するだけで上記のように入力されるはずである。

画面上にいくつかのエラー表示が現れる。これは cns の pdb format が純正とは若干違うために出ている警告なので、気にしなくてよい。

[OK]を数回クリックすると消える。

(4) Fit to\_first [return]

これで、3-35 残基目までの backbone を指標に、1 番目の座標への重ね書きができた。

(5) 見やすくする。

右上ボタンの「bb」をクリック backbone の意味である。

その 5 つ下の「Show Sel.」をクリック

これで主鎖だけの絵ができた。

(6) sidechain も表示してみる。

「Show all」をクリック

「heavy」をクリック

「Show sel.」をクリック

## (7) ribbon 図も作ってみる。

「ribbon」をクリック

text window に結果が表示されるので[close]とする。  
ribbon が表示される。

ribbon は絵は美しいがこの後の解析には向かない。

ribbon を消して line 図に戻る。

コマンド入力行より

→ 「RemovePrim [return]」 で 消える  
右のボタン「line」をクリック

## (8) Ramachandran Plot 作成

コマンド入力行より

→ 「SelectAngle "」 [return] ←シングルコーテ  
ーション2回

メニューより「Fig -> Ramachandran」をクリック

↓

「OK」をクリック

↓

メニューより「File -> Plot -> PostScript」をクリック

file name 入力 [rama.ps] [return]

メニューより「Fig -> Off」をクリックで終了

## (9) mean structure と r.m.s.d table の作成 (advanced)

メニューより「Edit -> Molecule -> Mean Structure」  
をクリック

「OK」をクリック

これで mean という名前で平均構造ができた。

メニューより「Calc -> RMSD」をクリック

「fit\_pairs」を check して「OK」

「Save」をクリック

file name 入力「CalcRmsd.txt」として「OK」をク  
リック

## (10) report の提出。

以下のとおりに操作せよ。

■mkdir result

■cp er2\_○○.pdb result/. [return]

( ↑ この作業を繰り返して、自分で選んだエネ  
ルギーの低いベストな構造を結果ディレクトリに5  
個コピーする)

■cp rama.ps result/. [return] ←同様に  
ramachandran を結果ディレクトリにコピー

■cp rmsd.txt result/. [return]

■tar czvf 出席番号.tgz result/. [return]

出席番号は二桁で。例 04.tar とか 28.tar とか

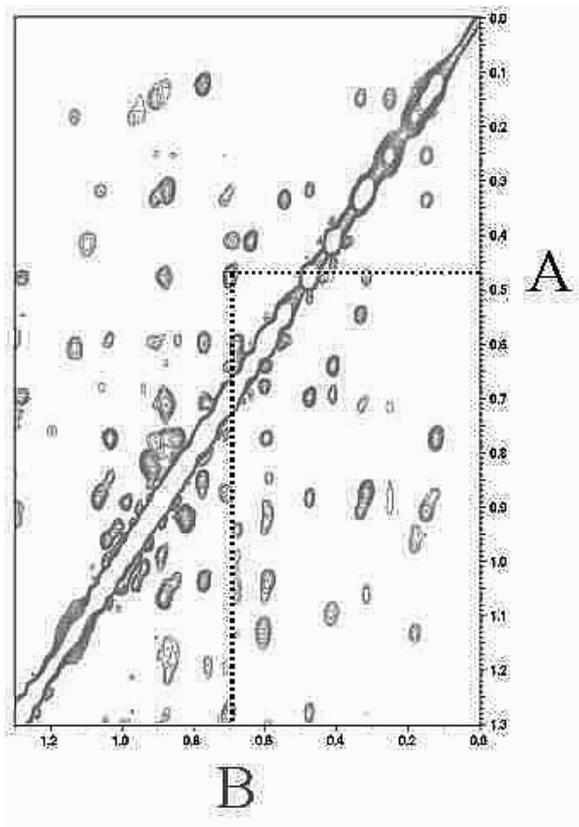
■cp 出席番号.tgz /home/share/report/cns/

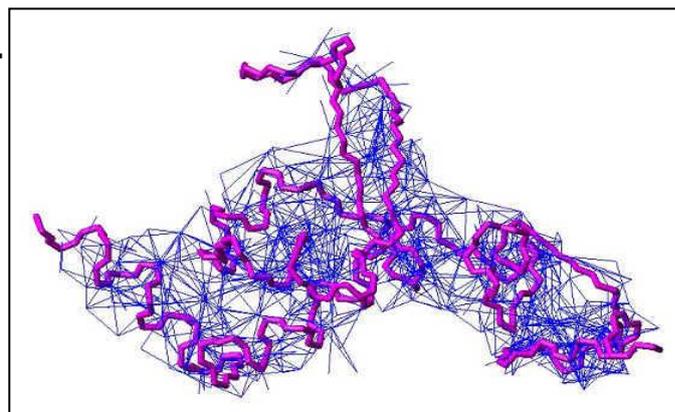
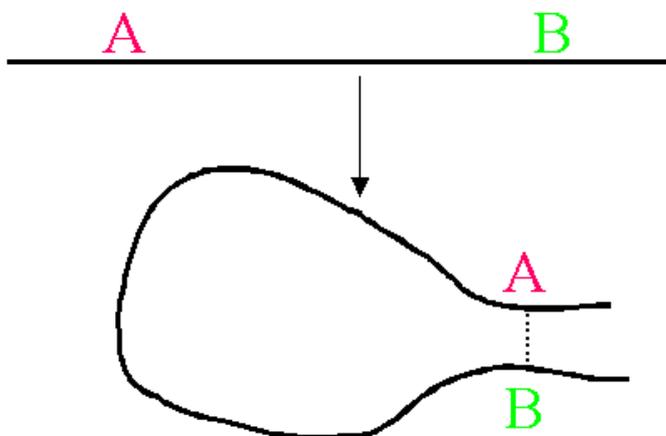
以上でおしまい。ご苦労様でした (^^) /~~

## 8. Advanced : 構造の評価・・

### (procheck\_nmr / AQUA)

更に詳細に構造を評価したり、構造計算の結果  
を元に帰属したNOEの表から誤りを探したり、  
あるいはPDBデータベースに登録するための  
「最終精密化版構造」を作製するためには、よ  
り詳細な解析を行わねばならない。その後、更  
に入力パラメータや実験データのうちの不適  
切な値を修正し、何回か構造計算を重ね、最終  
構造にいたる。そうした「蛋白質の健康」をチ  
ェックするためのパッケージが procheck\_nmr  
と AQUA である。実習では用いないが、そうし  
たツールを活用するということ覚えておい  
て欲しい。





## 9. ADVANCED2 : NOE の自動解析と完全自動の立体構造計算

通常、溶液 NMR 法を用いて蛋白質の構造決定を行う際に、【構造制限情報】として主として寄与するのは、【NOE (核オーバーハウザー効果)】によるプロトン間距離制限情報である。蛋白質を構成するアミノ酸の主鎖または側鎖の任意の 2 個のプロトンがおよそ 5Å より近ければ、交差緩和により NOE 相関ピークが観測される。NOE はプロトン間距離の 6 乗に反比例する。NMR による立体構造計算の原理は、なるべく多くのプロトン対の間に観測される NOE から、数多くのプロトン間距離情報を取得して、その制限情報を満たす立体構造の解のアンサンブルを数学的に得ることである。経験的に、個々の制限情報の精密さよりも、タンパク質のアミノ酸残基一残基あたりの情報の量 (たとえば一残基あたりの NOE の数) と、制限情報の集合全体の self consistency が、正確かつ精密な構造決定に寄与することが知られている。

今回の実習では、スペクトル表示・主鎖帰属の概要は行ったが、側鎖の帰属と NOE の帰属、および距離情報の作成は体験してもらわなかった。これまで NOE の帰属作業は、従来の手作業による NMR 解析において、側鎖の帰属以上に時間と手間のかかる作業であった。しかも蛋白質では、分子量の増大によって分子相関時間が長くなり、信号の線幅も増加する。そのため高磁場を利用して、 $^{13}\text{C}/^{15}\text{N}$  などの多次元 NMR を用いても、プロトンの化学シフトの重なりは回避できない。測定時間の制限からは、4 次元 NMR を用いたとしても、3 次元、4 次元目に十分な

デジタル分解能を得ることは難しい。そのため、任意の NOE の交差ピークを、化学シフトのみから唯一対のプロトン間の NOE であると帰属することは通常困難である。このように、化学シフトの情報からのみでは一意的には帰属を確定できない NOE のことを【ambiguous な NOE】と呼び、一対のプロトン対に帰属が確定できる【unambiguous (unique) な NOE】と呼んで区別する。分子量の小さい蛋白質では、unambiguous な NOE 由来の制限情報のみを用いて立体構造決定を行っても、正しい fold が決定できる例が多かった。そのため、まず手動で帰属した unambiguous な NOE から、構造計算を行って低分解能の中間構造のアンサンブルを得て、そこから得られるプロトン間の距離を解析して、「中間構造に矛盾しないように」ambiguous な NOE の帰属と解釈を進めていき、段階的に【制限情報】の数を増やしていきながら構造計算を繰り返しつつ精密化を進める、というのが従来方法であった。

しかし、このような方法では、初期に得られる中間構造が正しいかどうかを判断する基準に乏しく、誰がやっても同じように正しい高品質な立体構造が得られるという確証がない。そこで、構造決定の MD 計算の初期段階から ambiguous NOE が含んでいる構造制限情報を盛り込み、その帰属を動的に行う方法論が開発され、後の自動構造決定への道を開いた。現在、多くの構造プロテオミクス研究プロジェクトでは、この自動 NOE 帰属の方法を採用している。

### 9.1 ARIA

自動帰属方法のうち、最初に成功を収めたのが、Nilges らにより開発された【ARIA】である。ARIA の特徴は、構造計算エンジンとして今回実習に用い

た CNS を用いて、そこに **ambiguous distance restraints (ADR)** を取り扱えるような標的関数を導入したことである。ADR は、たとえば任意の一つの NOE ピークを与える可能性のある全てのプロトン対に対して、それらのプロトン対間の距離のアンサンブル平均の和に対してポテンシャルを発生する関数として設計されている。ADR が与える距離制限情報は、エラーを多く含んでいる情報であるにもかかわらず、少なくとも一つは正しい距離情報を含んでいると考えられる。そのため、より多くの NOE peak をお互いに矛盾なく解釈することのできる帰属の組み合わせが、最終的に「正しい」構造へとアンサンブルを収束させることを促す、と考えられている。標準的な ARIA の計算プロトコルでは、8 回の iteration を経て最終構造を導出する。最初の iteration では、ADR の多くは、大きなバイオレーションを発生する。しかし最初の iteration で得られた構造アンサンブルを各 NOE ピークごとに解析することで、(i) 系統的にバイオレーションを与えるノイズピークの除去と (ii) 各 NOE のそれぞれの帰属可能性のあるプロトン対について、構造アンサンブルから逆算することにより、実際の NOE ピークへの寄与の計算を行う。その後、寄与の高いプロトン対の組から順にその NOE ピークの帰属として採用していき、あらかじめ設定された cut off 値 (0.8-1.0) を満たせば、それよりも寄与の低いプロトン対を帰属可能性のリストから除外する。この方法で、iteration ごとにそれぞれの NOE ピークの多義性を解消していき、それとともにエネルギーの最小化と構造の収束が達成される。

なお、自分の Linux マシンに ARIA を導入するときには、CNS とともに両者のソースコードからコンパイルする必要がある。(必然的に、論文で発表された、便利で最新の機能を使おうと思うものは、日常的にこのような作業に直面する。)

## 9.2 CYANA (CANDID + DYANA)

ETH (現・理研 GSC) の Güntert は、自身が開発した NMR 用構造決定プログラム DYANA に、ambiguous distance restraints を取り扱えるようにし、また NOE 自動帰属機能を実装したプログラム CANDID を発表した。(CYANA は両者を組み合わせ統合した際のプログラム名である)。

DYANA は分子の共有結合に関して CNS よりも単純化された力場を用いて、二面角座標系で MD を行うことを特徴としており、直交座標系の MD よりも計算効率がすぐれている。CYANA では ARIA が導入した ADR に、更にいくつかの ambiguous NOE を的確に自動帰属するための方法論が追加された。以下に ARIA との差異を列挙する。(i) ambiguous NOE に対

して与えられる、初期の帰属可能性の組み合わせの、複合的な帰属による順位付けと、順位の低い帰属解の除外。(ii) NOESY ピークに対して化学シフトの一致度の高い帰属解の順位を上げる。(iii) 3D/4D NOESY における、対称的な相関ピークが存在している帰属解の順位を上げる。(iv) ペプチドの化学構造を考慮して、3bond または 4bond の距離にあるプロトン間での帰属解の順位を上げる。(v) 「ネットワークアンカリング」により、全ての peak の NOE 帰属解の全体を見渡して、互いに NOESY ピークを与えあう数個のプロトンの組のサブセットをネットワークとして発見し、そこに属する帰属解の順位をあげる。これは、中間構造を規定しないでも self consistency の高い NOE 距離情報の組み合わせを同定して、その情報に重みをつけて構造計算に反映される過程にほかならない。(vi) 中間的な構造アンサンブルを用いた、NOE 帰属可能性の評価、ランク付けとノイズの除去。(vii) Constraint combination を行い、任意に選んだ複数の長距離 NOE (5 残基以上離れたアミノ酸残基間の NOE) 由来の距離情報を、論理和で結合した新たな仮定の距離制限情報を生成し、それを用いて構造計算を行う。これは、長距離 NOE における誤った帰属が蛋白質の誤った global fold への収束を引き起こすことをある程度予防することのできる手法である。CYANA の標準的なプロトコルでは 7 回の iteration (CYANA cycle) を経て最終構造を収束させる。

## 9.3 NOE 自動帰属を行う際の問題点

以上の方法を用いれば、NOE の手作業による帰属無しで、構造計算開始から、数時間から数日でよく収束した global fold を決定することができる。いずれの方法でも、最初の iteration において誤った構造に収束しないようにするために、側鎖、とくにメチル基と芳香族プロトンの帰属を精密になるべく完全に行っておくことが重要である。更に、NOE データから peak file を作成するときの peak pick 時の、ノイズとシグナルを見分ける閾値の設定が、作業の効率を大きく左右する。実際には、automatic peak picking だけでは不十分で、手作業で各 peak の妥当性と位置を吟味すべきである。更に、ambiguous NOE を帰属するための化学シフトの誤差の許容値を、実際のスペクトルの質に合わせて適宜設定する必要がある。また計算を始める前に、スペクトルや測定条件の違いによるわずかな化学シフトのずれなどを丁寧に補正しておくことがよい結果につながる。同時に、二次構造化学シフトや主鎖の二面角から決定された二次構造情報や、残余双極子相互作用、水素結合など、NOE 以外の構造制限情報を、iteration 初期で活用することで、誤った構造に収束する危険性を低減することができる。

=付録=====

ファイルフォーマットなどの例

■■ 例 1

```
er2.seq
ASP PRO MET THR CYS GLU GLN ALA MET ALA SER
CYS GLU HIS THR MET CYS GLY TYR CYS GLN GLY
PRO LEU TYR MET THR CYS ILE GLY ILE THR THR
ASP PRO GLU CYS GLY LEU PRO
```

■■ 例 2

```
er2_noe.tbl
assign (resid 1 and name HA) (resid 2 and name HD#) 3.550 1.7 0.0 ! 1 ASP- HA 2
PRO QD 3.55
assign (resid 1 and name HA) (resid 19 and name HE#) 7.630 5.8 0.0 ! 1 ASP- HA 19
TYR QE 7.63
assign (resid 1 and name HB#) (resid 2 and name HD2) 5.500 3.7 0.0 ! 1 ASP- HB2 2
PRO HD2 5.50
assign (resid 1 and name HB#) (resid 2 and name HD3) 5.500 3.7 0.0 ! 1 ASP- HB2 2
PRO HD3 5.50
assign (resid 1 and name HB#) (resid 2 and name HD2) 5.500 3.7 0.0 ! 1 ASP- HB3 2
PRO HD2 5.50
assign (resid 1 and name HB#) (resid 2 and name HD3) 5.500 3.7 0.0 ! 1 ASP- HB3 2
PRO HD3 5.50
assign (resid 1 and name HB#) (resid 2 and name HD#) 4.960 3.2 0.0 ! 1 ASP- QB 2
PRO QD 4.96
assign (resid 2 and name HA) (resid 3 and name HN) 2.620 0.8 0.0 ! 2 PRO HA 3
MET HN 2.62
assign (resid 2 and name HA) (resid 19 and name HD#) 7.640 5.8 0.0 ! 2 PRO HA 19
TYR QD 7.64
assign (resid 2 and name HA) (resid 19 and name HE#) 7.630 5.8 0.0 ! 2 PRO HA 19
TYR QE 7.63
```

■■ 例 3

```
er2_aco.tbl
assign (resid 1 and name n) (resid 1 and name ca)
(resid 1 and name c) (resid 2 and name n) 1.0 135 80 2

assign (resid 2 and name n) (resid 2 and name ca)
(resid 2 and name c) (resid 3 and name n) 1.0 125 160 2

assign (resid 2 and name c) (resid 3 and name n)
(resid 3 and name ca) (resid 3 and name c) 1.0 180 310 2

assign (resid 2 and name c) (resid 3 and name n)
(resid 3 and name ca) (resid 3 and name c) 1.0 -55 300 2
```

■■ 例 4

```
( example of hb file ... format is same as noe)
assign (resid 10 and name O) (resid 33 and name N) 3.000 1.2 0.0 ! 10 HIS O 33
LYS N 3.00
assign (resid 12 and name O) (resid 31 and name N) 3.000 1.2 0.0 ! 12 ALA O 31
LEU N 3.00
assign (resid 16 and name O) (resid 29 and name N) 3.000 1.2 0.0 ! bulge
assign (resid 29 and name O) (resid 14 and name N) 3.000 1.2 0.0 ! bulge
assign (resid 31 and name O) (resid 12 and name N) 3.000 1.2 0.0 ! 31 LEU O 12
ALA N 3.00
```

```

assign (resid 32 and name O) (resid 40 and name N) 3.000 1.2 0.0 ! 32 VAL O 40
LYS N      3.00
assign (resid 30 and name O) (resid 42 and name N) 3.000 1.2 0.0 ! 30 PHE O 42
VAL N      3.00
assign (resid 28 and name O) (resid 44 and name N) 3.000 1.2 0.0 ! 28 TYR O 44
ARG N      3.00
assign (resid 44 and name O) (resid 28 and name N) 3.000 1.2 0.0 ! 44 ARG O 28
TYR N      3.00
assign (resid 42 and name O) (resid 30 and name N) 3.000 1.2 0.0 ! 42 VAL O 30
PHE N      3.00
assign (resid 40 and name O) (resid 32 and name N) 3.000 1.2 0.0 ! 40 LYS O 32
VAL N      3.00
assign (resid 38 and name O) (resid 34 and name N) 3.000 1.2 0.0 ! 38 SER O 34
TRP N      3.00

```

```

■■ 例5
head -68 er2_1.pdb | tail -42
REMARK 104 dihedral restraints with scale factors of 5; 200; 400
REMARK 0 planarity restraints with a scale factor of NA
REMARK NCS restraints not used.
REMARK =====
REMARK          bond, angles, improp, vdw(<1.6), dihedral
REMARK violations :      0      0      0      0      18
REMARK RMSD       : 0.0023  0.421  0.193      26.900
REMARK =====
REMARK          noe, cdih, coup, oneb, carb-a, carb-b,
REMARK violations :      0      0      0      0      0  -----
REMARK RMSD       : 0.025  0.021  0.000  0.000  0.000  0.000
REMARK 0.2/2 viol.:      1      0      0
REMARK =====
REMARK          dani, sani
REMARK violations :      0      0
REMARK RMSD       : 0.000  0.000
REMARK .2/.1 viol.:      0      0
REMARK =====
REMARK Protons          violations, rmsd
REMARK all      :          0      0.000
REMARK class 1:          0      0.000
REMARK class 2:          0      0.000
REMARK class 3:          0      0.000
REMARK class 4:          0      0.000
REMARK =====
REMARK overall = 71.6455
REMARK bon      = 3.03865
REMARK ang      = 27.5915
REMARK imp      = 1.5788
REMARK vdw      = 21.1843
REMARK harm     = 0
REMARK noe      = 18.2464
REMARK coup     = 0
REMARK oneb    = 0
REMARK carb    = 0
REMARK prot    = 0
REMARK dani    = 0
REMARK sani    = 0
REMARK cdih    = 5.775784E-03
REMARK ncs     = 0
REMARK =====
REMARK DATE:03-Jul-2001 13:50:13      created by user: hiroakih

```

■■ 例6 analysis.pl

```
#!/usr/bin/perl
die("usage : analysis.pl <base name>¥n") if $#ARGV < 0;

for($i=1;$i<101; $i++){

    $filename = $ARGV[0] . "_" . $i . ".pdb" ;
    if(!-f $filename) {exit;};
    open(FILE, "$filename");
    while(<FILE>){
        if (/ATOM/) {&print_E; last};
        split;
        $E{$_[1]}=$_[3];
        1;
    }
    close(FILE);
};

sub print_E {
    printf("%-12s (overall, noe, cdih, vdw) %8.2f %8.2f %8.2f %8.2f¥n", $filename,
    $E{"overall"},
    $E{"noe"}, $E{"cdih"}, $E{"vdw"});
}
}
```

CNS が出力した多くの pdb ファイルのヘッダから、計算結果に関連するパラメータを読み取って、それのみを表にして表示する perl script の例

## 構造生物学実習 (NMR による構造計算と構造評価)

### MOLMOL

#### “MOLEcule analysis and MOLEcule display” Kurt Wütrich’s Laboratory, ETH Switzerland & Bruker Spectrospin AG.

#### はじめに

MOLMOL は生体高分子の三次元構造の表示・解析・処理を行うプログラム(分子構造ビューワー)である。横浜市立大学大学院・生体超分子科学専攻の大学院実習では RasMol, SwissPDBViewer, MolScript などの分子構造ビューワーの利用法を学んできたが, NMR 実習では MOLMOL を使用する。それは, MOLMOL を開発した研究室が NMR による蛋白質立体構造決定に対して先駆的研究を行ってきた ETH の Wuthrich 研であり, とくに NMR 立体構造の表示については特別に使いやすい機能が整備されているからである。

#### MOLMOL の機能は

1. PDB ファイルを読み込み表示可能.
2. CYANA などの構造計算の COR 形式のファイルを読み込み表示が可能.
3. NMR 構造計算で得られた「同一の分子の複数の立体構造」(＝構造アンサンブル)を取り扱って表示可能. 特に重ねがきをおこなったり, 比較する領域を変化させながら構造の差異 (RMSD) を計算したりすることが可能.
4. CYANA などの構造計算の制限距離情報を読み込んで分子上に表示させることができる.
5. Ramachandran プロットを作成できる.
6. 分子表面を計算して表示できる.
7. 静電ポテンシャルを計算して表示できる.
8. 例えば二本の helix の相対角度を計算することが可能.

9. 多くの操作を MACRO として記述して保存が可能. たとえば複数の分子に同じような配色で絵を作成したいときなどに, クリック操作を繰り返さなくて良いので便利.

10. Win/Mac/Linux などで動作する

欠点としては

11. 電子密度マップを書かせることができない.
12. 結晶中の分子の対称操作などは行うことができない.
13. 半透明な分子や影の処理などにはバグがあったりする.

#### MOLMOL の起動

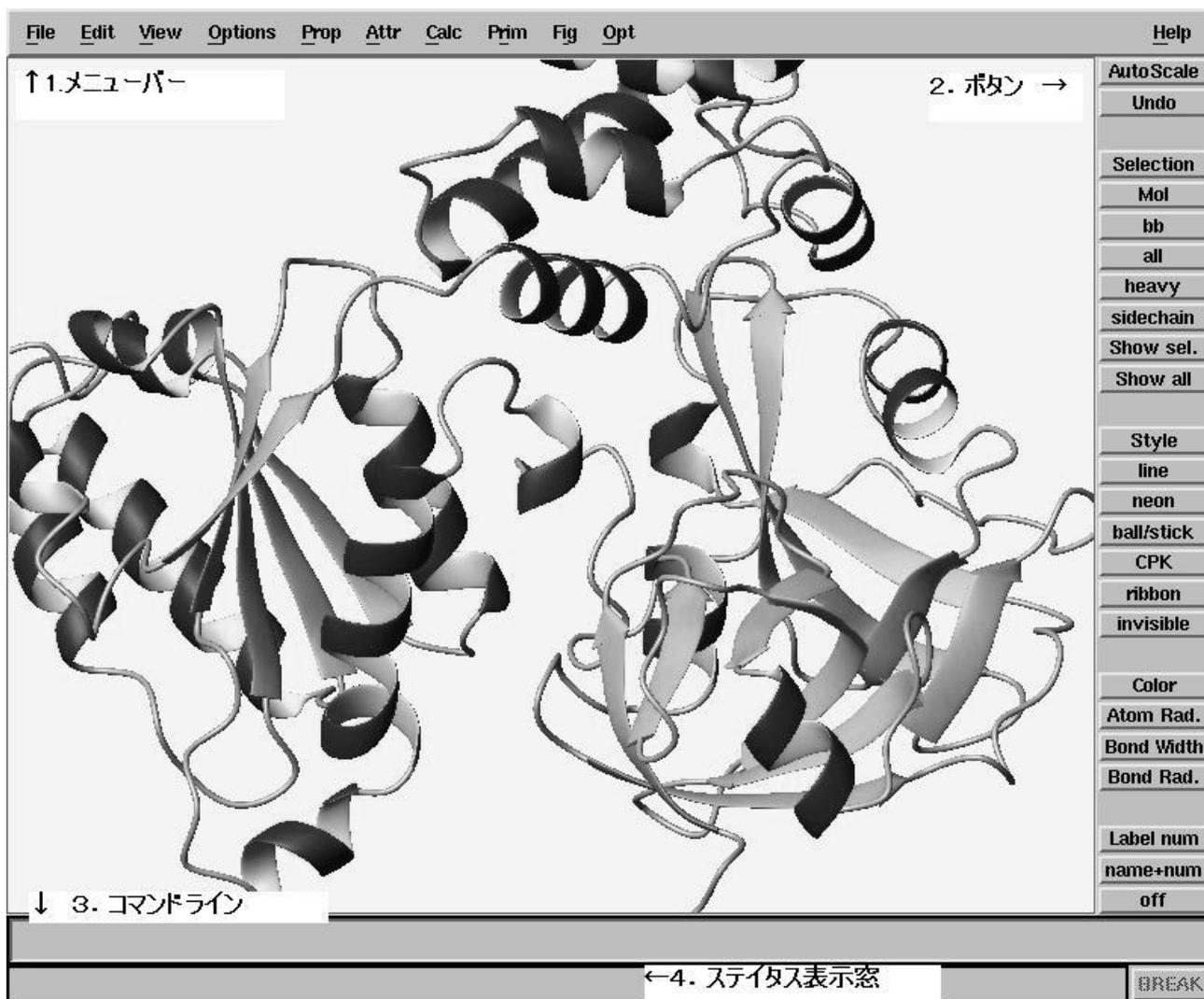
Linux 版の場合 (今回の実習の場合) コマンドラインから

```
[hiroaki@pc101] molmol er2_11.pdb [return]
```

er2\_11.pdb を表示する.

プログラムの状態によって主鎖だけを表示したり, 側鎖もコミで表示したりすることがあるが「とりあえず主鎖だけ表示」するためのボタン操作の流れを説明すると

1. 全て消す  
[all]→[invisible]
2. 主鎖を選んで線画表示  
[bb] →[line]



## MOLMOL のウィンドウの説明

**Main Window:** 分子を表示するエリアの上部にメニューバー, 右にボタンがなっている. ボタンはよく使う内容が5グループにわけて並んでいて, 上からズーム系, 分子原子選択系, 表示 (スタイル) 系, 表示 (色・大きさ) 系, 表示 (ラベル) となっている. 画面下側にはコマンドライン入力行とステータス表示行(窓)がある. 分子をマウスでクリックすると何をクリックしたのかが表示されるのはここである.

分子表示エリア内部の分子をマウスでクリックして, 移動・回転・ズームを行うことができる.

マウス左→回転

マウス中央→平行移動

マウス右→メニュー

マウス左+中→ズーム  
イン・アウト

## その他のウィンドウ

メニューバー→Options

→ (最下段) User

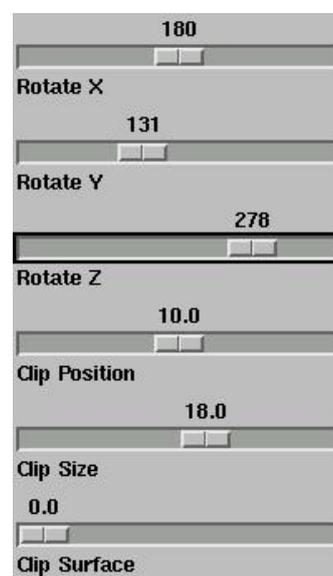
interfaces のチェックボ

ックスをクリックする

ことで更に2つのウィ

ンドウを開かせること

ができる.

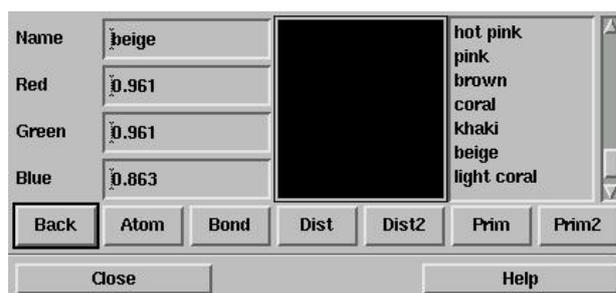
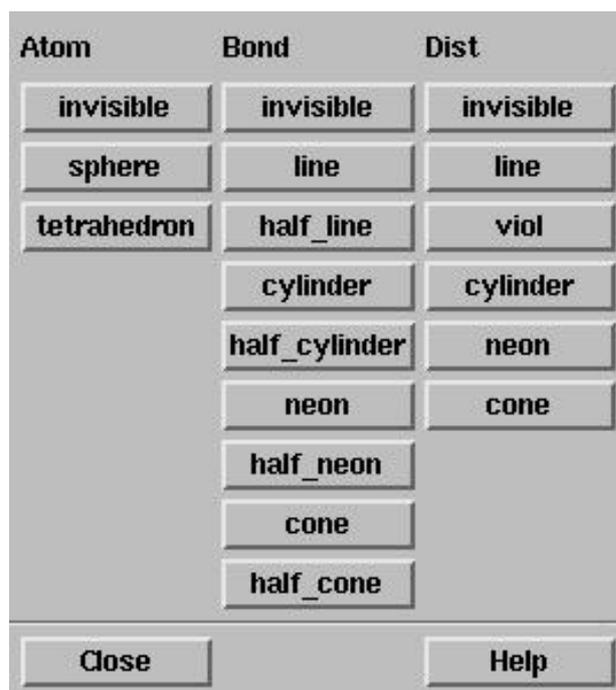
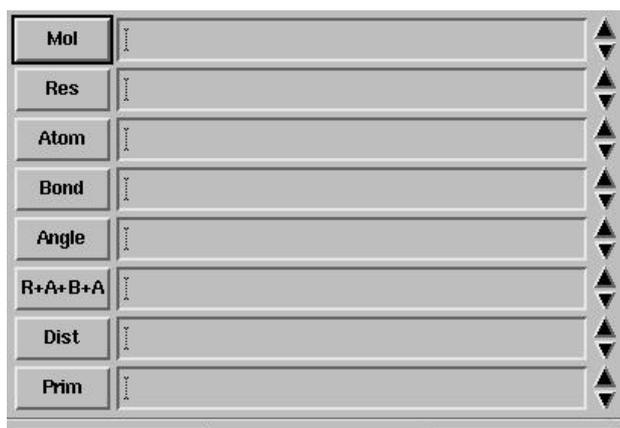


**Valuator box** : 分子の回転を行う.

**Log Window**: これまでのコマンドのログが記録される. 自作のマクロを創るときも個々を参考にすれば良い.



また, 画面右側の右側のボタンのうち, **Selection** (原子の選択を行うダイアログ), **Mol** (表示する分子を選択決定するダイアログ), **Style** (選択中の原子やボンドの表示属性を決定するダイアログ), **Color** (色を決定するダイアログ) のボタンで, 対応するダイアログがそれぞれひらく.



## MOLMOL の使い方のミソ

MOLMOL の使い方は基本的には上記のウィンドウやボタンで操作を行ったり, メニューバーからコマンドを選択することで行う. またカーソルをコマンドラインにもって行って, キーボードでコマンドを入力することでも操作可能である (コマンド補完機能が備わっているので, 最初の頭文字だけ覚えていれば操作可能である).

MOLMOL の操作で初心者がとっつきにくいのは

1. primitives の概念 と
2. 原子の選択のしかた・文法 であろう.

**Primitives** とは MOLMOL が作画して取り扱っている描画オブジェクトのことで, ヘリックスの板状矢印やコイル, また MOLMOL で計算させた分子表面がそれに相当する. たとえばリボン図を描かせたと

して、**SelectAtom** で ' ' (シングルコーテーション2回, MOLMOLでは空文字を入れると「全て」を意味する)として[Style][invisible]とクリックしてもリボン図は消えない. これはリボン図のリボン表面は**Atom / Bond**とはべつの**Primitive**という属性でそれぞれ記述されているからである.

**SelectPrim ' ' [retutn]**

**[invisible]**

とやると消すことができる.

原子の選択のしかたにも癖があるので要注意である. 詳細は MOLMOL のヘルプやオンラインマニュアル, または日本語の使い方のページの紹介に譲るとして, ここでは簡単なルールを書いておく.

(1) コマンドラインから入力するとき→コマンドの引数はかならずシングルコーテーションで挟む

例 : **SelectAtom 'heavy'**

(2) コマンドラインから入力するときの値は数字の場合にはそのまま, 文字列の場合にはダブルコーテーションで囲む

例 : **SelectRes 'num=20..40'** (残基 20~40 を選ぶ)

**SelectRes 'name="LYS+''** (残基リジンを選ぶ)

なお, **GLU** と **GLU-**, **LYS** と **LYS+**を別々に取り扱うなど, かなり融通が利かないつくりなので注意.

(3) **SelectionDialog** から同じ情報を入力する場合にはシングルコーテーションは不要 (これもまた混乱を招く原因)

(4) ただし **SelectAtom** 文を上手につかうとかなり高度な選択ができる.

例 : **SelectAtom 'res.name="ARG+" & name="CA''**

以上のように MOLMOL では内部でとりあつかっている分子・残基・原子・結合・結合角にそれぞれ属性を値をふりわけており, それを適切に選択し, そこに適切な値をいれていくことで色や大きさを変えることができるようになっていく. 以下に MOLMOL のマニュアルから抜粋したいくつかの属性値や値の

表を載せておく.

item	value	type	explanation
mol	num	integer	molecule number
mol	number	integer	molecule number
res	num	integer	residue number
res	number	integer	residue number
prim	num	integer	primitive number
prim	number	integer	primitive number
mol	name	string	molecule name
res	name	string	residue name
atom	name	string	atom name
angle	name	string	angle name
atom	shift	float	chemical shift of atom
atom	bfactor	float	B factor of atom
atom	vdw	float	van der Waals radius of atom
atom	charge	float	partial charge of atom
atom	heavycharge	float	charges on heavy atoms
atom	avgcharge	float	averaged charges on heavy atoms
atom	simplecharge	float	simple charges from setup file
atom	d	float	distance from reference atom(s)
bond	len	float	bond length
angle	val	float	angle
dist	val	float	distance
dist	limit	float	limit of constraint
dist	viol	float	violation of constraint
dist	upl	bool	true if distance is upper limit
dist	lol	bool	true if distance is lower limit
dist	hbond	bool	true if distance is H-bond
atom	attr	integer	graphics attribute index
bond	attr	integer	graphics attribute index
dist	attr	integer	graphics attribute index

prim	attr	integer	graphics attribute index
------	------	---------	--------------------------

item	name	description
any	all	Always set for all items. Do not modify!
any	selected	Set for selected items. Most commands operate on items that have this property set.
any	displayed	Only items with this property set are displayed. Can be useful for making items (especially molecules) temporarily visible/invisible, use the <a href="#">Molecule Dialog</a> for a convenient way of setting it.
mol	movable	Only molecules with this property set are affected by <a href="#">interactive manipulations</a> and commands that transform molecules. Useful for moving molecules relative to each other, use the <a href="#">Molecule Dialog</a> for a convenient way of setting it.
atom	visible	Set for all atoms that have a display style that makes them visible. Convenient to restrict a selection to only visible items.
bond	visible	Set for all bonds that have a display style that makes them visible.
dist	visible	Set for all distances that have a display style that makes them visible.
prim	visible	Set for all primitives that have a display style that makes them visible.

item	name	description
atom	bb	Backbone atoms of a protein or nucleic acid.
bond	bb	Backbone bonds of a protein or nucleic acid.
atom	sc	Side chain atoms of a protein or nucleic acid.
bonds	sc	Side chain bonds of a protein or nucleic acid.
atom	heavy	Heavy atoms.
bond	heavy	Bonds between two heavy atoms.
atom	heavysc	Heavy side chain atoms of a protein or nucleic acid.
bond	heavysc	Bonds between two heavy side chain atoms of a protein or nucleic acid.
atom	phos	Atoms in phosphate group of a nucleic acid.
bond	phos	Bonds in phosphate group of a nucleic acid.
atom	sugar	Atoms in sugar group of a nucleic acid.
bond	sugar	Bonds in sugar group of a nucleic acid.
atom	base	Atoms in base of a nucleic acid.
bond	base	Bonds in base of a nucleic acid.
atom	proton	Hydrogen atoms.
atom	ca	Atoms with name CA.

atom	pseudo	Pseudo atoms (name starting with Q).
atom	lonepair	Lone pair (name starting with LP).

文法としては

Item.name = name

Item.name = value

のように使い、集合に対しては& (または小文字で and 論理 seki), | (または小文字で or 論理和), ! (または小文字で not, 否定)などを作用させることができる。