

YASARA チュートリアル タンパクーリガンド・ドッキング計算

(株) アフィニティサイエンス

概要：YASARA Structure (Version 18.4.24) によるタンパクーリガンドのドッキング操作をチュートリアル形式で説明する。本チュートリアルでは、アルツハイマー病の発症に関与すると考えられるアミロイド β タンパクの産生および凝集性を決定する酵素 β -site APP cleaving enzyme 1 (BACE1) の複合体構造 (PDB ID:5I3V) を用いて、リガンドの再ドッキングを実施する。

1. ドッキング準備

1.1 作業ディレクトリの作成と YASARA (GUI) の起動

ドッキングを実施するにあたり入出力ファイルを格納するディレクトリ (フォルダ) を新規に作成する (Windows ではエクスプローラ、Linux では mkdir コマンド等を使用)。作成場所やファイル名は任意で問題ないが、YASARA 上で表示が乱れるため、日本語を含まない方が好ましい。

例

(Windows) C:\Users\username\Desktop\YASARA_DATA\tuto_docking

(Linux) /home/username/YASARA_DATA/tuto_docking

その後、YASARA (GUI) を起動する。

「Options」→「Working directory」を選択し、先ほど作成したディレクトリを指定し作業ディレクトリを設定する。

1.2 PDB ダウンロードと構造表示

「File」→「Load」→「PDB file from Internet」を選択し、「5I3V」を入力、PDF ファイルをダウンロードすると、画面上にタンパク 3D 構造が表示される。

※インターネット未接続の場合、PDB ファイルを別途入手し、「File」→「Load」→「PDB file」等から構造を読み込む。

「View」→「Style scene」またはファンクションキーで分子の表示方法を変えることができる。本チュートリアルでは、「Ribbon」(F6 キー) に設定している。

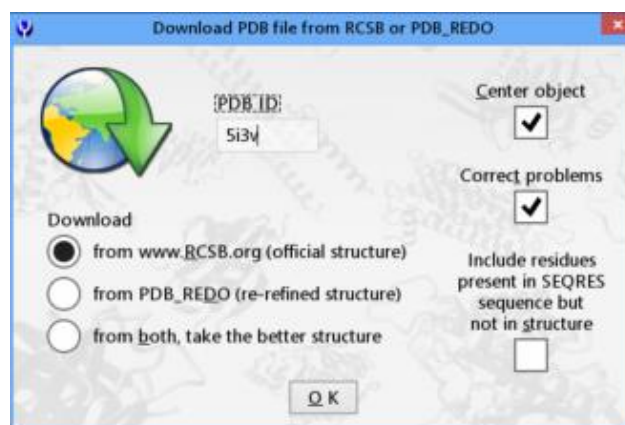


Figure 1.1 PDB ファイルのダウンロード

1.3 ドッキング前処理①（クリーニング・不要分子削除）

解析のため、「Edit」→「Clean」→「All」を選択して、PDB 構造のクリーニング処理（欠損原子・水素付加、結合次数の調整など）を行う。

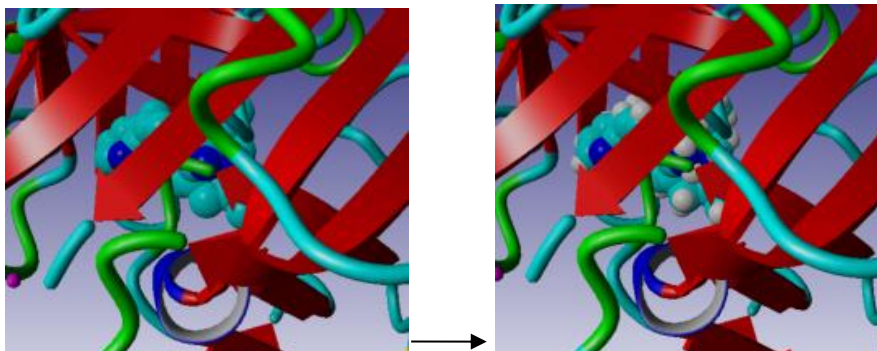


Figure 1.2 Clean 処理による水素分子（白球）の付加

続けて、レセプター（タンパク）とリガンド（低分子）以外の不要なものを取り除くため、メニュー「Edit」→「Delete」→「Molecule」を実行して、水分子（Waters）および不要分子を削除する（今回は、上から 1 番目と 3 番目の Molecule のみを残し、その他は削除する）。

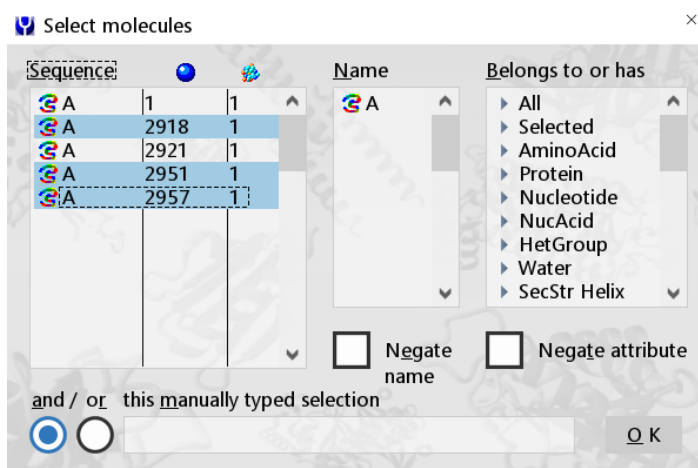


Figure 1.3 Delete の実行（Ctrl キー＋クリックで複数選択が可能）

1.4 ドッキング前処理②（分子名変更・オブジェクト分割・ドッキングセル設定）

画面右側 HUD の Molecule 名（“Mol”に続く部分）がデフォルトではすべて“A”になっているため、それぞれの Molecule を右クリックして「Name」を選択し、適当なものに変更する。本チュートリアルではレセプターに“R”、リガンドに“L”と設定した。名前はアルファベット一文字のみ選択可能（大文字・小文字別）。

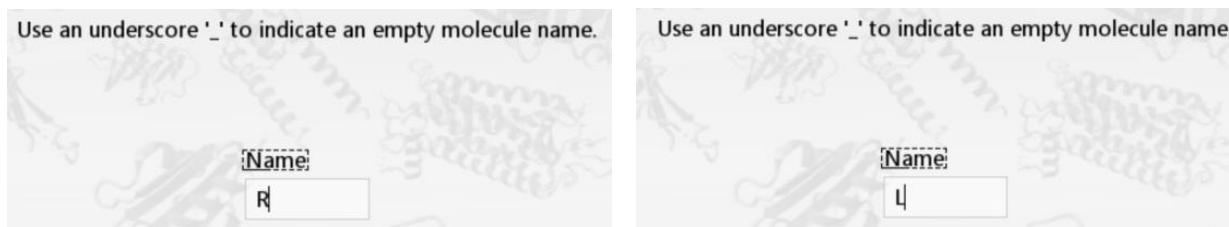


Figure 1.4 Molecule 名の変更

この段階では同じ Object の中にレセプターとリガンドが含まれている状態のため、別々の Object に分割する。画面右側 HUD の Object（5i3v）を右クリックして、「Split」を選択、Center の設定は好みのほうを選択し、「Split at all molecules」をクリックする。（本チュートリアルでは、分割後の Objects が同じ（ローカル）座標系を持つ「Do not Center ~」を選択する。もう一方の「Center the newly created objects ~」を選択すると、幾何中心がそれぞれのローカル座標の原点と一致するように新規オブジェクトがシフトされる。）

SCENE CONTENT				
Obj	Name	Vis	Act	Atom
1	5i3v	Yes	Yes	1
	Mol R	No	Yes	1
	Mol L	Yes	Yes	5767
2	-----	No	No	-----
3	-----	No	No	-----

SCENE CONTENT				
Obj	Name	Vis	Act	Atom
1	5i3vR	Yes	Yes	1
	Mol R	No	Yes	1
2	5i3vL	Yes	Yes	5767
	Mol L	Yes	Yes	5767
3	-----	No	No	-----

Figure 1.5 Split 処理による Object の分割

結合ポケット付近にドッキング範囲を指定する。画面右側 HUD のリガンド Object (5i3vL) を右クリックし、「Select」→「newly」を選択することでリガンドを選択状態にする。次に「Simulation」→「Define simulation cell」を選択、「around selected atoms」をクリックすることでリガンド周辺にドッキング範囲(SimCELL)が出現する。

※SimCELL の設定を省略する場合、タンパク全体を対象としたドッキングが実施される。

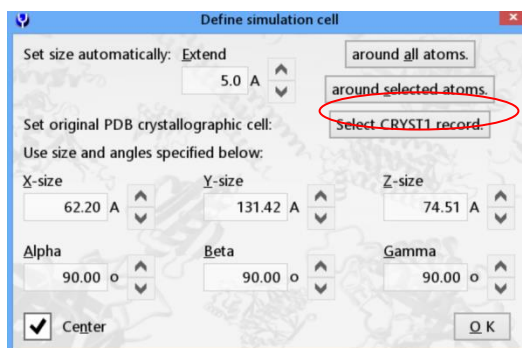


Figure 1.6 SimCELL の定義

2. ドッキング実行

YASARA でドッキングを行うには、GUI から直接ドッキング解析・条件設定を行う簡易的な方法 (①) と、マクロを用いて実行する方法 (②) の二つの手段がある。ドッキング結果の解析機能等の観点から、後者のマクロ利用が推奨されているが、2つの手法について紹介する。

2.1 [方法①]簡易手法

2.1.1 力場設定

ドッキングを実行するには力場の設定が必要になる。「Simulation」→「Force field」を選択し、力場を選択して、「OK, and if ~ parameters」をクリックする。(本チュートリアルでは AMBER03 を選択)

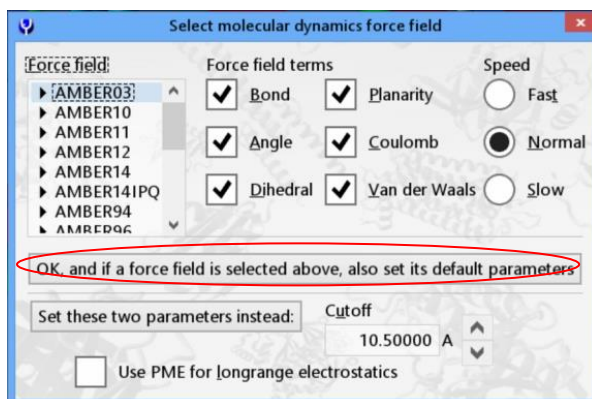


Figure 2.1 力場設定

2.1.2 ドッキング Experiment 実行

「Options」→「Choose experiment」→「Docking ligand to receptor」を選択し、「Proceed with experiment」をクリックする。まずレセプターを指定 (5i3vR)、次にリガンドを指定 (5i3vL)、そしてドッキングプログラム (VINA)、条件 (例 Docking runs = 10, Cluster RMSD = 5.00A)、出力ファイル名 (任意) を指定して「OK」をクリックするとドッキングが開始される。

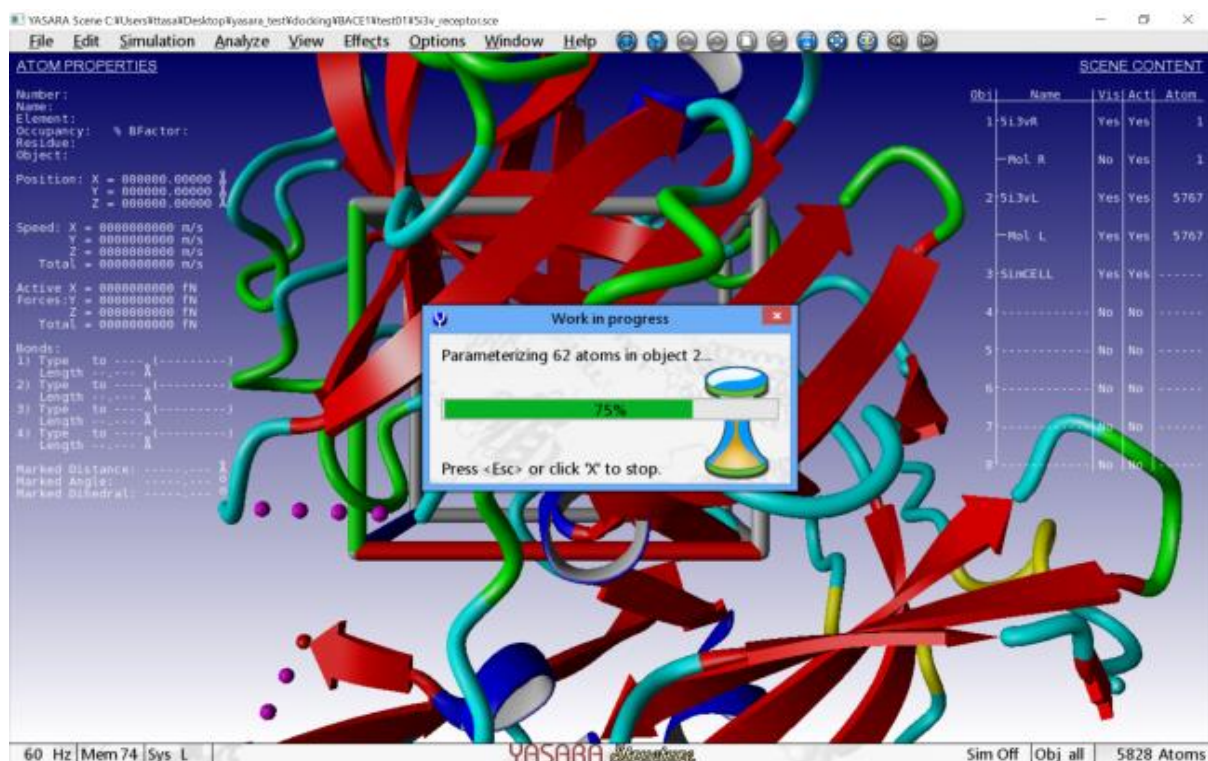


Figure 2.2 ドッキングシミュレーション実行中画面

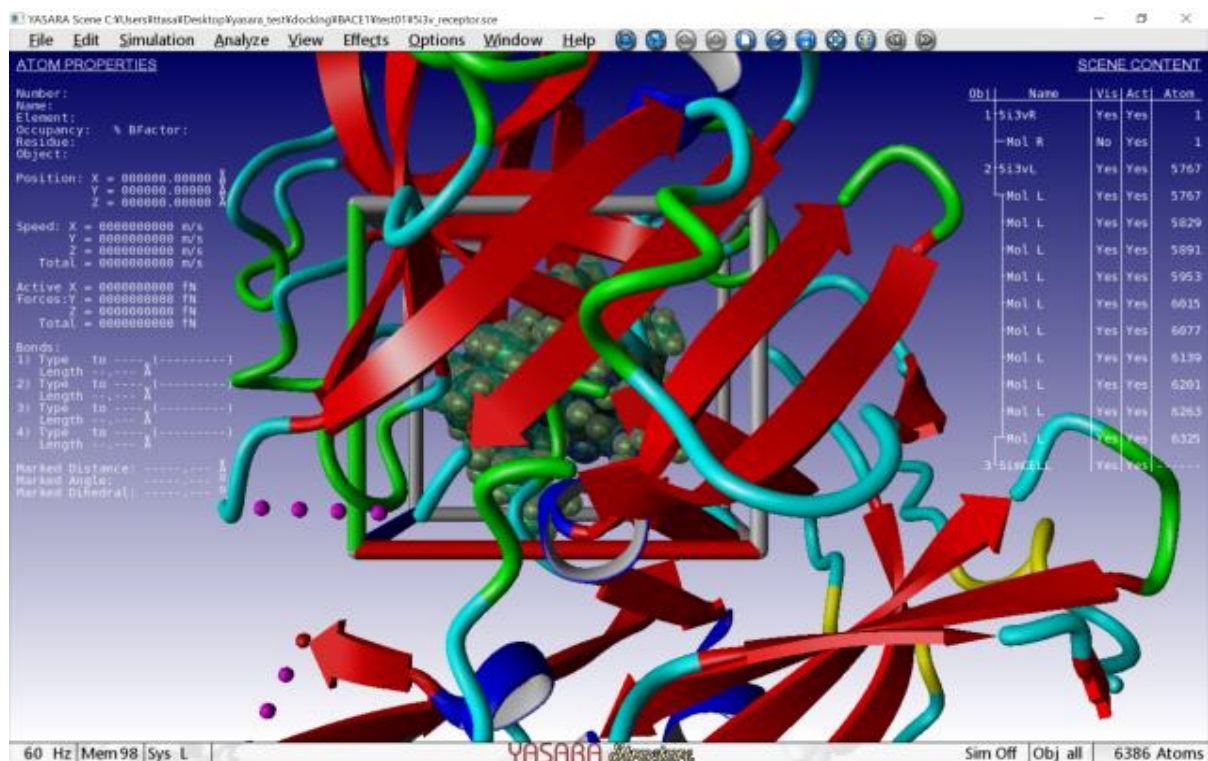


Figure 2.3 ドッキングシミュレーション計算終了画面 (結合エネルギーのスコアが高い順に並べられている)

2.2 [方法②]マクロファイルを利用する方法

ドッキング計算の実行に当たっては、パラメータやコマンドが記述されたマクロファイルを読み込んで自動的に実行する方法も存在し、現行版ではこちらを使用することが推奨されている。マクロを利用する場合、リガンドとレセプターそれぞれの構造ファイルを準備する必要がある。以下ではソフトウェア付属のドッキング計算用マクロファイルを利用する流れを紹介する。

2.2.1 リガンドとレセプターファイルの保存

リガンドのみのファイルを pdb、yob 形式のいずれかで作業ディレクトリ内に保存する。本チュートリアルでは、yob 形式で保存する。「File」→「Save as」から「YASARA Object」を選択し、リガンド Object (5i3vL) を指定して保存する。ファイル名の末尾は必ず「_ligand(.拡張子)」にする。

(例 5i3v_ligand.yob)

続いて、レセプターのみのファイルを作業ディレクトリ内に保存する。pdb、yob、sce のいずれの形式でも構わないが、結合ポケットを指定する場合は、SimCELL の情報も一緒に保存可能な sce 形式を選択する必要がある。本チュートリアルでは、sce 形式で保存する。画面右側 HUD でリガンド Object (5i3vL) を右クリックし、「Delete」を実行して削除、レセプターと SimCELL のみを残す。続いて、「File」→「Save as」から「YASARA Scene」を選択、ファイル名の接頭辞(prefix)部分はリガンドファイルとそろえるようにし、末尾は必ず「_receptor(.拡張子)」にする。

(例 5i3v_receptor.sce)

2.2.2 ターゲット (MacroTarget) 設定

「Options」→「Macro&Movie」→「Set target」で、ターゲットを設定する。

作業ディレクトリを選択し、レセプターとリガンドのファイル名の共通 (接頭辞) 部分をターゲット名として設定する。

(例 C:\Users\username\Desktop\YASARA_DATA\tuto_docking\5i3v)

GUI から設定する場合、Remove チェックボックスの「file extension」, 「from underscore」, 「nothing」を利用可能。

2.2.3 ドッキング実行

「Options」→「Macro&Movie」→「Play macro」でドッキングを実行する。「dock_run.mcr」または「dock_runensemble.mcr」を選択する。前者では単一のレセプターのみでドッキングを行い、後者では側鎖の回転異性体(side-chain rotamers)を考慮して数種のレセプターを自動で作成し、ドッキングを行う。本チュートリアルでは「dock_run.mcr」を使用する。

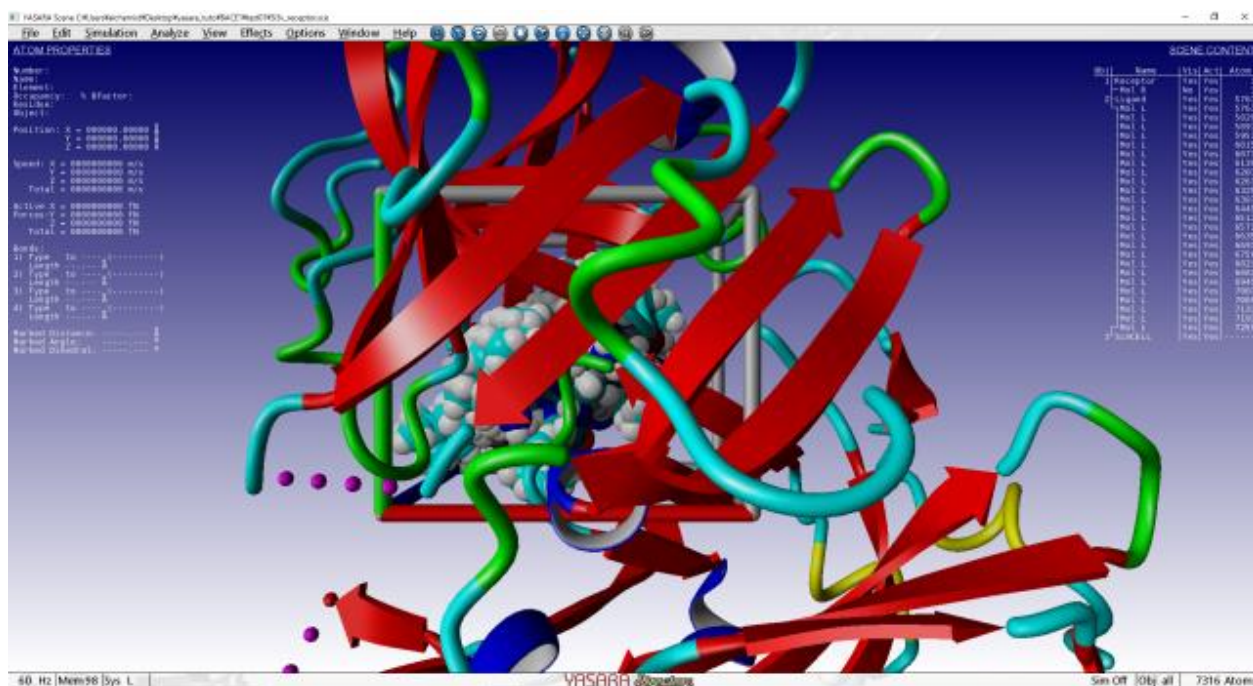


Figure 2.4 ドッキングシミュレーションマクロ (dock_run) の実行結果

2.3 コマンドラインからマクロを実行する方法

最初に[方法②]の手順と同様にレセプター・リガンドの両ファイルを作成する。次に Windows であればコマンドプロンプト (Win キー+R キーでファイル名を指定して実行ダイアログを開き、名前に cmd を入力し OK ボタンをクリックすることで起動可)、Linux であればターミナルプログラム (gnome-terminal など) を開き、次の手順に従いコマンドを入力、Enter キーを押すことで実行することができる。

- ① YASARA インストールディレクトリ内の YASARA 実行ファイルを指定。
スペースを空け、テキストモード用オプション (-txt) を指定。
- ② スペースを空け、次に YASARA インストールディレクトリ内の mcr ディレクトリに格納されているマクロを指定。
- ③ スペースを空け、最後に MacroTarget を指定。
(全体をダブルクォーテーション (") で囲み、ターゲットをシングルクォーテーション (') で囲む)。
※dock_run.mcr を適当なディレクトリにコピー・編集する事で MacroTarget や各種パラメータの設定も可能。

実行例

(Windows) C:\Users\username\Desktop\yasara\YASARA.exe -txt C:\Users\username\Desktop\yasara\mcr\dock_run.mcr
"MacroTarget='C:\Users\username\Desktop\YASARA_DATA\tuto_docking\5i3v'"

(Linux) /home/username/yasara/yasara -txt /home/username/yasara/mcr/dock_run.mcr
"MacroTarget='/home/username/YASARA_DATA/tuto_docking/5i3v'"

3. ドッキング結果の解析

3.1 ログファイル (マクロ使用時)

マクロを実行してドッキングを実行すると作業ディレクトリ内にログファイル(.log)が作成され、詳しいドッキング結果を確認することができる。レセプター・リガンド間の結合エネルギー(Binding energy)が強いものから順に並べられており(値が大きい方が結合力が強い)、その他に解離定数や結合に関与するアミノ酸残基が記載されている。
また、ドッキング結果のクラスタリングも自動で行われ、その結果が続けて出力されている。

5i3vlog - メモ帳

ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)

Global docking result analysis

25 VINA docking runs of the ligand object 2 to the receptor object 1 yielded the following results, sorted by binding energy [more positive energies indicate stronger binding, and negative energies mean no binding]

Run	Bind. energy [kcal/mol]	Dissoc. constant [pM]	Contacting receptor residues
001	000009.5090	00000000107105.2031	ASP R 32 GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 GLN R 73 GLY R 74 LYS
002	000009.5090	00000000107105.2031	ASP R 32 GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 GLN R 73 GLY R 74 LYS
003	000008.4440	00000000646347.3125	GLY R 34 SER R 35 SER R 36 VAL R 69 PRO R 70 TYR R 71 GLY R 74 LYS R 75 TRP
004	000008.4440	00000000646347.3125	GLY R 34 SER R 35 SER R 36 VAL R 69 PRO R 70 TYR R 71 GLY R 74 LYS R 75 TRP
005	000008.3530	00000000753650.1875	ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 TYR R 71 TRP R 76 LYS
006	000008.3530	00000000753650.1875	ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 TYR R 71 TRP R 76 LYS
007	000008.1100	00000001135772.0000	LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 GLN R 73 GLY R 74 LYS
008	000008.1100	00000001135772.0000	LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 GLN R 73 GLY R 74 LYS
009	000008.0650	00000001225396.5000	LEU R 30 ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 TYR R 71 GLN
010	000008.0650	00000001225396.5000	LEU R 30 ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 TYR R 71 GLN
011	000007.8660	00000001714531.8750	SER R 35 SER R 36 ASN R 37 VAL R 69 TYR R 71 LYS R 75 TRP R 76 LYS R 107 PHE
012	000007.7960	00000001929550.8750	LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 GLN R 73 GLY R 74 LYS
013	000007.6280	00000002562134.2500	GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 GLN R 73 GLY R 74 LYS R 75 TRP
014	000007.5420	00000002962378.5000	LEU R 30 ASP R 32 GLY R 34 SER R 35 SER R 36 VAL R 69 PRO R 70 TYR R 71 LYS
015	000007.5260	00000003043468.7500	LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 GLN R 73 GLY R 74 LYS
016	000007.3670	00000003980314.5000	LEU R 30 GLY R 34 VAL R 69 TYR R 71 GLN R 73 GLY R 74 LYS R 75 TRP R 76 ASP
017	000007.3330	00000004215409.0000	LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 TRP R 76 LYS
018	000007.3230	00000004287161.5000	GLY R 34 SER R 35 SER R 36 VAL R 69 PRO R 70 TYR R 71 LYS R 75 TRP R 76 PHE
019	000007.3180	00000004323494.0000	ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 PRO R 70 TYR R 71 TRP
020	000007.2840	00000004578958.5000	GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 PRO R 70 TYR R 71 THR R 72 GLN
021	000007.2570	00000004792349.5000	ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 THR R 72 TRP R 76 PHE R 108 ILE
022	000007.2170	00000005127065.5000	GLY R 13 LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 TRP R 76 LYS
023	000007.1990	00000005285219.5000	ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 TYR R 71 GLN R 73 GLY
024	000007.0750	00000006515627.5000	GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 GLN R 73 GLY R 74 TRP R 76 ASP
025	000007.0170	00000007185728.5000	GLY R 13 LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 TRP R 76 PHE

After clustering the 25 runs, the following 7 distinct complex conformations were found:
[They all differ by at least 5.0 Å heavy atom RMSD]

Clu	Bind. energy [kcal/mol]	Dissoc. constant [pM]	Contacting receptor residues
001	000009.5090	00000000107105.2031	ASP R 32 GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 GLN R 73 GLY R 74 LYS
002	000008.3530	00000000753650.1875	ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 TYR R 71 TRP R 76 LYS
003	000007.7960	00000001929550.8750	LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 TYR R 71 GLN R 73 GLY R 74 LYS
004	000007.6280	00000002562134.2500	GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 GLN R 73 GLY R 74 LYS R 75 TRP
005	000007.3330	00000004215409.0000	LEU R 30 ASP R 32 GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 TRP R 76 LYS
006	000007.3180	00000004323494.0000	ASP R 32 GLY R 34 SER R 35 SER R 36 ASN R 37 VAL R 69 PRO R 70 TYR R 71 TRP
007	000007.0750	00000006515627.5000	GLY R 34 SER R 35 VAL R 69 PRO R 70 TYR R 71 GLN R 73 GLY R 74 TRP R 76 ASP

While the table above lists the best binding energy in each cluster, it is sometimes helpful to also look at the energy spread [average and standard deviation], the dissociation constant has been recalculated from the average binding energy:

Clu	Members	Bind. energy spread [kcal/mol]	Dissoc. constant [pM]
001	007	000008.0459+-000000.7170	00000001265738.3992
002	003	000007.8120+-000000.4653	00000001878289.0600

Figure 3.1 ドッキング log ファイル

3.2 マクロ dock_play による解析

ドッキング結果を一覧する場合、「Option」→「Macro&Movie」→「Play macro」を選択してから、「dock_play」を指定することで、ドッキング結果を連続で確認できる。

「Proceed to cluster analysis」を選択すると、クラスター分類したものや、レセプター・リガンド間の相互作用（H-Bonds, Hydrophobic, Pi/Pi, Cation/Pi）について確認することができる。

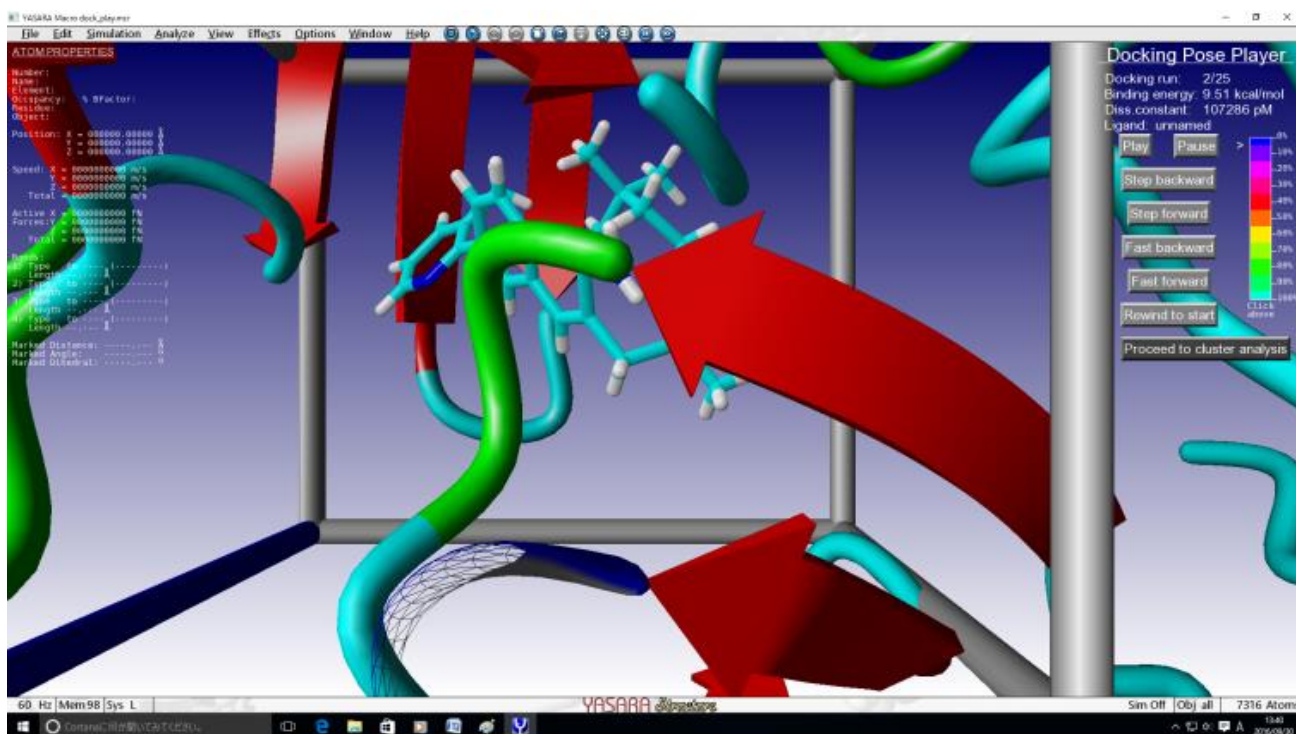


Figure 3.2 「dock_play」実行画面

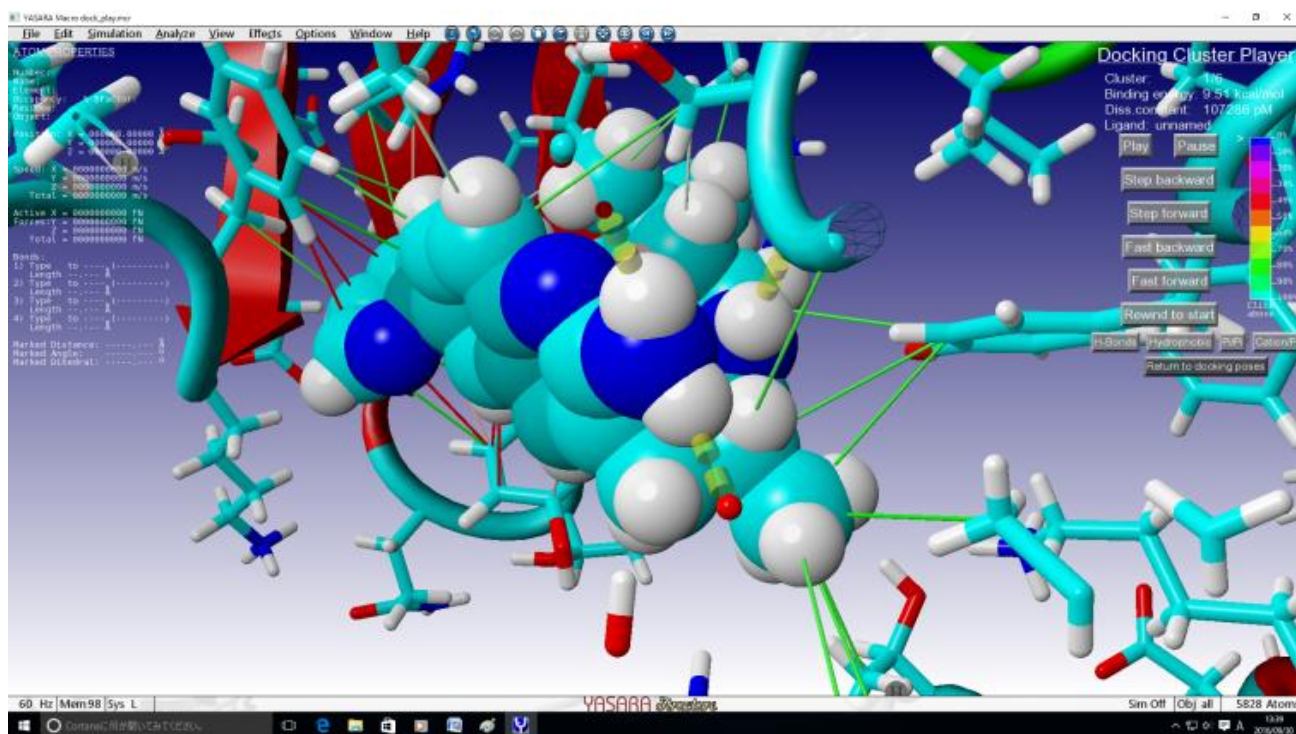


Figure 3.3 「dock_play」クラスター分類や相互作用の表示

3.3 RMSD 計算

ドッキング解析で得られた結合ポーズと結晶構造のリガンド座標を比較するため、RMSD 値を計算することができる。

コマンドラインからドッキング解析を実施した場合は、YASARA GUI を起動しドッキング解析結果である YASARA Scene ファイルをロードする。これには、「File」→「Load」→「YASARA Scene」として“MacroTarget.sce”ファイルを指定するか GUI 上に該当ファイルをドラッグ&ドロップする（本チュートリアルでは 5i3v.sce）。

続いて、「File」→「Load」→「PDB file from Internet」または「PDB file from local PDB」等から RMSD 計算対象の結晶構造を読み込む（PDB file from Internet を実行済みの場合、ローカルディスク上に PDB ファイルが自動保存されている PDB ID に 5i3v を入力し、OK ボタンをクリックする）。

次に「Analyze」→「Align」→「Objects with MUSTANG」を選択し、source オブジェクト（動かしたい方、ここでは結晶構造データ（5i3v））と target オブジェクト（動かしたくない方、ここではレセプター（Receptor））を指定し、アライメントを行う。

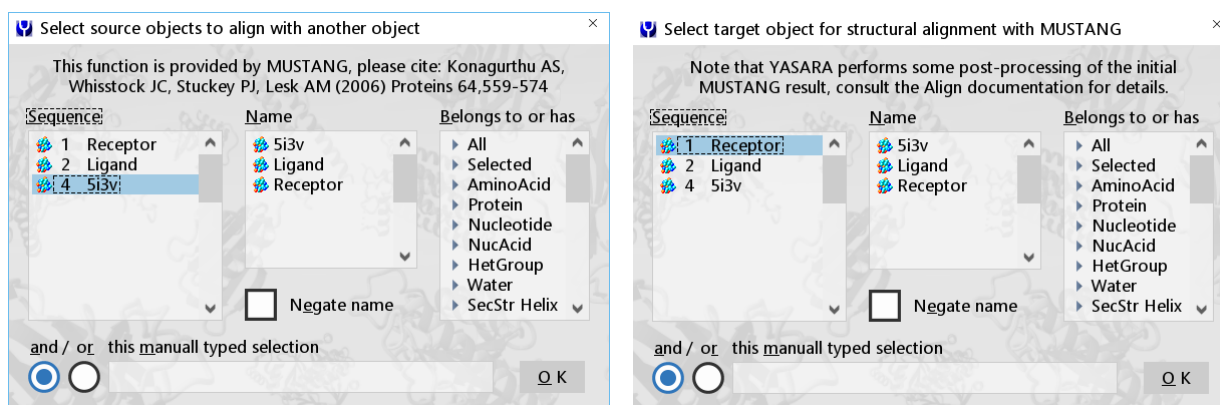


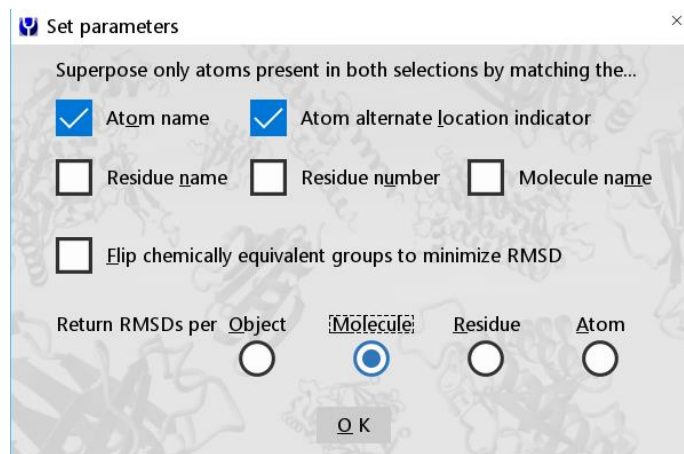
Figure 3.4 アライメント画面

その後、RMSD を計算する分子名を変更する（この例では、ドッキングリガンドを 1、結晶構造のリガンドを 0 と設定）。

Obj	Name	Vis	Act	Atom
2	Ligand	Yes	Yes	5767
	Mol 1	Yes	Yes	5767
	Mol L	Yes	Yes	5829
	Mol L	Yes	Yes	5891
	Mol L	Yes	Yes	5953
	Mol L	Yes	Yes	6015
	Mol L	Yes	Yes	6077
	Mol L	Yes	Yes	6139
	Mol L	Yes	Yes	6201
	Mol L	Yes	Yes	6263
	Mol L	Yes	Yes	6325
	Mol L	Yes	Yes	6387
	Mol L	Yes	Yes	6449
	Mol L	Yes	Yes	6511
	Mol L	Yes	Yes	6573
	Mol L	Yes	Yes	6635
	Mol L	Yes	Yes	6697
	Mol L	Yes	Yes	6759
	Mol L	Yes	Yes	6821
	Mol L	Yes	Yes	6883
	Mol L	Yes	Yes	6945
	Mol L	Yes	Yes	7007
	Mol L	Yes	Yes	7069
	Mol L	Yes	Yes	7131
	Mol L	Yes	Yes	7193
	Mol L	Yes	Yes	7255
3	SimCELL	Yes	Yes	-----
4	5i3v	Yes	Yes	7317
	Mol A	Yes	Yes	7317
	Mol 0	Yes	Yes	10234
	Mol A	Yes	Yes	10237
	Mol A	Yes	Yes	10267

Figure 3.5 RMSD 計算対象分子（左；赤色が結晶構造）と分子名の変更例（右）

続いて、「Analyze」→「RMSD of」→「Molecules」を選択し、名前を変更した2分子を選んだ後、「Atom name」と「Atom alternate location indicator」チェックボックスを有効にし、OK ボタンをクリックするとコンソール上に RMSD の値が出力される。



```
>RMSDMol 0,1,Match=AtomName+AltLoc,Flip=No,Unit=Mol
Molecule 0 and Molecule 1 have 0.5107 A RMSD
```

Figure 3.6 RMSD 設定画面（上）と出力結果（下）

3.4 RMSD 計算（マクロ利用による複数分子の一括処理）

1つのオブジェクト内に含まれる複数分子と参照用分子の RMSD 計算を一括処理する場合、次のようなマクロを利用する。テキストエディタを用いて次の内容の「calc_rmsd.mcr」マクロファイルを作成し、「Options」→「Macro & Movie」→「Play macro」からこのファイルを指定し実行する。ダイアログが表示され、オブジェクト番号（またはオブジェクト名）と参照用分子名の入力を求められるので適宜入力し、OK ボタンをクリックすることで計算が行われる。

```
#
# A simple macro for renaming each molecule name
# and calculating RMSD with a reference molecule
#
targetobj=""
refmol=""
if targetobj==" or refmol=="
    targetobj,refmol = ShowWin Type=TextInput,
        Title="Calculate RMSD",
        Text="Please enter a target object number and reference molecule name",
        Text="Object Number for docked ligands",
        Text="Molecule Name for Reference (e.g. 0)"
console off
first='A'
mollist() = ListMol Obj (targetobj)
for i=1 to count mollist
    NameMol (mollist(i)),(chr (ord first+i-1))
    RMSDMol (refmol), (mollist(i)),Match=AtomName,Unit=Mol
console open
```

Figure 3.7 calc_rmsd マクロファイル

本マクロを実行するとまず指定オブジェクトの分子名がリネームされ (A,B,C...と順次変更)、これら分子と参照用分子との RMSD 値がコンソール上に順次出力される (本マクロでは最大で 62 の分子を一括して計算可能)。参照用分子に対しては重複がないよう分子名を指定する (この例では 0 を指定)。

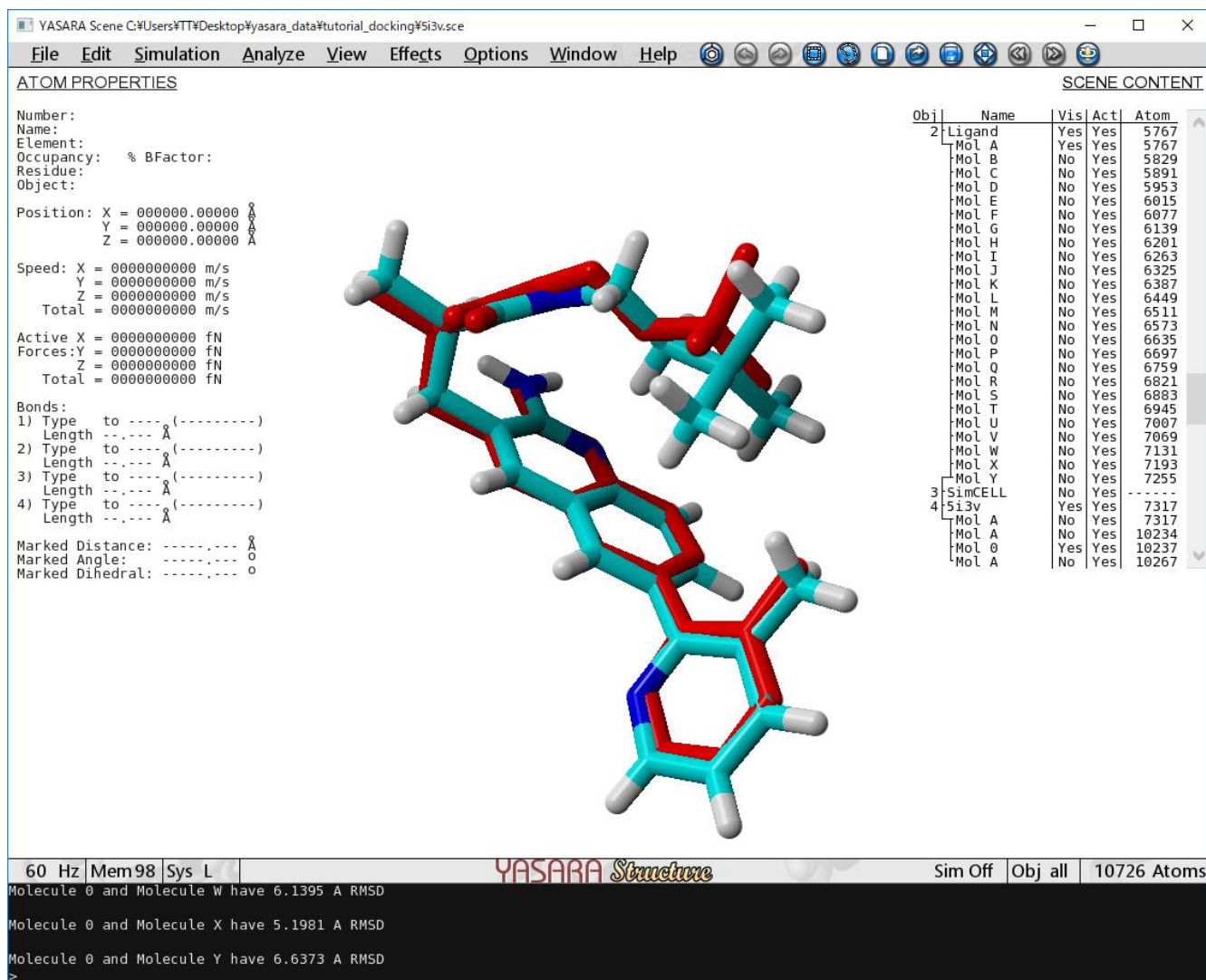


Figure 3.8 calc_rmsd マクロ実行結果

マクロファイル「calc_rmsd.mcr」は次の URL よりダウンロードすることが出来る。

YASARA マクロファイル集

https://www.affinity-science.com/yasara_macro.html

以上