

myPresto 4.204

- *sievgene* -

USER MANUAL

Version 1.0

Copyright (C) 2006-2011 National Institute of Advanced Industrial Science and Technology (AIST)

Copyright (C) 2006-2011 Japan Biological Informatics Consortium (JBIC)

本ドキュメントについて

本ドキュメントは、「*myPresto 4.2 USER MANUAL*」の別冊です。コピーライト、プログラム使用許諾条件、著者および引用文献については、「*myPresto 4.2 USER MANUAL*」の記述に準じます。

謝辞

本ソフトウェアの研究開発は、新エネルギー・産業技術総合開発機構(NEDO)、及び、経済産業省(METI)の援助によって行われました。ここに感謝の意を記します。

本ソフトウェアは、故・京極好正博士の推進する研究プロジェクトで開発されました。

目次

1	sievgene.....	5
1.1	実行方法.....	5
1.2	入力データの作成.....	5
1.2.1	コントロールファイル.....	5
1.2.1.1	PHASE> INPUT グループ.....	7
1.2.1.2	PHASE> GRID グループ.....	11
1.2.1.3	PHASE> CONF グループ.....	17
1.2.1.4	PHASE> DOCK グループ.....	19
1.2.1.5	EXE> MIN グループ.....	25
1.2.1.6	PHASE> OUTPUT グループ.....	25
2	計算例.....	27
2.1	Sample-1：受容体 - 低分子化合物ドッキング.....	27
A	入出力ファイル.....	53
A.1	ドッキングエンジンの入出力ファイル.....	53
A.1.1	フェーズの説明.....	53
A.2	入力ファイル.....	54
A.2.1	制御ファイル.....	55
A.3	出力ファイル.....	61
A.4	ログ出力（簡易出力）.....	61
B	ユーティリティ.....	69
B.1	make_point.....	69
C	乱数の選択.....	71
C.1	目的.....	71
C.2	指定方法.....	71
C.3	ログ表示.....	71

(余白)

1 sievgene

1.1 実行方法

sievgene は蛋白質 PDB ファイルとトポロジーファイル、ポケット点 PDB ファイル、リガンド MOL2 ファイルを入力として、蛋白質 - リガンドのドッキングおよびスコア評価を行い、結果をファイルに出力します。

これらの入力ファイルおよびドッキング条件は、コントロールファイルに指定します。sievgene は、標準入力よりコントロールファイルを読み込んで動作します。

```
% sievgene < control file > output
```

Sievgene は、古いバージョン (sievegene_for_dockingpose) と新しいバージョン (sievgene_for_screening) の両方を提供しています。バージョンが違くと最適なパラメータが異なります。スクリーニング計算で MTS 法を使用する場合、相互作用行列の作成と、追加のドッキング計算は同じバージョンで行う必要があります。sievegene_for_dockingpose は、2007 年の相互作用行列、sievgene_for_screening は、2009 年の相互作用行列に対応します。

1.2 入力データの作成

1.2.1 コントロールファイル

コントロールファイルは、以下のグループからなり、各グループは "QUIT" で終了します。

- PHASE> INPUT **グループ** : 主な入力ファイル名を指定します。
- PHASE> GRID **グループ** : グリッドポテンシャル生成のオプションを指定します。
- PHASE> CONF **グループ** : リガンドコンフォーマ生成のオプションを指定します。
- PHASE> DOCK **グループ** : グローバルサーチのオプションを指定します。
- EXE> MIN **グループ** : ローカルサーチのオプションを指定します。
- PHASE> OUTPUT **グループ** : 最終結果の出力を指定します。
- EXE> SIEV : コントロールファイルの終わりを示します。

ドッキング処理は、"EXE> SIEV"行を読み込んだ時点で開始します。

【注意】ドッキング処理をせずに、グリッドポテンシャルのみを発生させることができます (CONF/DOCK/MIN/OUTPUT フェーズ指定をしない方法 : 「2 計算例」参照)。この場合、入力ファイルへの「EXE> SIEV」の記述は不要です。

```

PHASE> INPUT
  LIGAND = MOL2
  NAMELI = Lig_md2.mol2 ; リガンドの座標位置ファイル
  REFERE = MOL2
  NAMERE = Lig_md.mol2 ; リガンドの参照位置ファイル
  TOPOLO = FORM
  NAMETO = Pro.tpl ; 蛋白側のトポロジー
  COORDI = PDB
  NAMECO = Pro_md.pdb ; 蛋白側の座標
  POINTC = PDB
  NAMEPO = point.pdb ; ポケット領域を示すプローブ点
  ATOMData = READ ; 原子データファイルの入力
PHASE> GRID
  GRIDPOtential = BINA ; Gridポテンシャルの入力ファイル
  NAMEGRid = grid.file ; Gridポテンシャルの入力ファイル指定
  QUIT
PHASE> CONF
  CONFLimit = 100000 ; 配座生成の試行最大数
  CONFORMernumber = 100 ; 生成する配座数(リガンドを变形)
  QUIT
PHASE> DOCK
  METHOD = FLEX
  GENERAtion = 1 ; 絞込みの回数
  SCORENumber = 10 ; 1リガンド当たり登録する上位スコア数
  QUIT
EXE> MIN
  METHOD= STEEP CPUTIM= 360000.0
  UPRATE= 1.0 DOWNRATE= 0.3
  LOOPLI= 100 UPDATE= 20
  MONITO= 100 CONVGR= 0.01D0
  LOGFOR= SHOR BESTFI= YES
  QUIT
PHASE> OUTPUT
  COORDInate = MOL2 ; リガンド座標のファイル出力形式
  NAMECOordinate = ex.cor ; 各実験でのリガンド座標ファイル名
  CANDIDatenuMber= 10 ; 各実験でのリガンド座標を出力する上位スコア数
  SCORENumber = 10 ; scoreファイルに出力する上位スコア数
  NAMESCore = ex.score ; scoreファイル名
  QUIT
EXE> SIEV

```

sievgene のコントロールファイルの例

スクリーニング、ドッキングに最適化した入力は、
 sievgene/sievgene_for_screening/input, sievgene/sievgene_for_dockingpose/input
 に、s0.inp として置いてあります。MTS/DSI 法でのドッキング計算は、この入力ファイル
 で作成されています。

コントロールファイルの各コマンド説明

必須	
省略可	
特定機能をユーザー指定するとき必須	

1.2.1.1 PHASE> INPUT グループ

INPUT グループでは、蛋白質の PDB ファイルとトポロジーファイル、蛋白質のポケット座標ファイル、リガンド座標指定ファイルの指定などのために、それらの外部ファイルの指定を行います。

【注意】各外部ファイルの書式は、巻末「A ファイル書式」を参照。

INPUT グループでの指定事項：

- (1) 蛋白質トポロジー指定
- (2) 蛋白質の座標の指定
- (3) リガンドの構造と座標の指定
- (4) RMSD 計算のためのリガンドの参照座標の指定
- (5) 蛋白質ポケット点座標の指定
- (6) Accessible Surface Area モデルの指定
- (7) リスタートの指定
- (8) 探索対象の指定
- (9) 蛋白質の原子半径の係数
- (1 0) ログ出力形式の指定
- (1 1) QUIT

(1) 蛋白質トポロジーの指定

TOPOLogy: 蛋白質トポロジーファイルの書式 ()

=NOREad ; 読み込まない(デフォルト)

=FORMAtted ; 書式つきアスキーファイル

=BINArY ; バイナリーファイル

NAMETOpology=(トポロジーファイル名、80 文字以内 。 TOPOLogy=[FORM|BINA]のとき)

(2) 蛋白質の座標の指定

蛋白質座標の指定を行います。BINArY の場合、蛋白質を構成する原子のデータの入力のため、(1)の蛋白質トポロジーの指定ファイルが必須となります。

COORDInate: 蛋白質 3 次元座標ファイルの書式 ()

=PDB ; PDB ファイルフォーマット

=BINArY ; バイナリーファイル

NAMECOordinate=(座標ファイル名、80 文字以内)

(3) リガンドの構造と座標の指定

ドッキング対象のリガンドの構造と座標を指定します。マルチフォーマット(複数の分子のファイルをつなげた形式)のファイルの場合、複数リガンドに対してドッキングを行います。

LIGANDcoordinate: リガンドファイルの書式 ()
 =MOL2 ; MOL2 形式(デフォルト)
 =PDBX ; PDBX 形式(cosgene トポロジーファイル+PDB 書式を XML で統合)
 NAMELigandordinate=(座標ファイル名、80 文字以内。())

(4) RMSD 計算のためのリガンド参照座標の指定

ドッキング後の配座と、参照座標との位置のずれ RMSD を計算することができます。正解の受容体 リガンド複合体があり、リガンドの座標が既知である場合、正解の座標を参照座標とすることでドッキングの精度を確認することができます。

REFEREcecoordinate: リガンド参照座標ファイルの書式 ()
 =NORE ; 読み込まない(デフォルト)
 =MOL2 ; MOL2 形式
 NAMEReferenceordinate=(座標ファイル名、80 文字以内。())

【注意】 REFERE= NORE とした場合、入力リガンドの座標を参照座標に設定します。

【注意】 REFERE= MOL2 かつ参照座標が入力リガンドと異なる分子である場合は、入力リガンドの座標を参照座標に設定します。

(5) 蛋白質ポケット点座標の指定

蛋白質のポケット点の座標を指定します。

POINTCoordinate : 蛋白質ポケット点座標の書式 ()
 =PDB ; PDB ファイルフォーマット
 =BINArY ; バイナリーファイル
 NAMEPoinTcoordinate=(蛋白ポケット点座標の名前、80 文字以内())

(6) Accessible Surface Area モデルの指定

sievgene では Accessible Surface Area (以降 ASA) でのポテンシャル計算を行うため、この計算の指定を行います。

ASAMethod : ASA 計算方法 ()
=PAIRwise ; Pairwise (デフォルト)
=RICHmond ; Richmond

(7) リスタートの指定

スクリーニングのジョブが中断された場合に、中断時点から再開するためのリスタート情報の入出力を指定します。

RESTARTjob : 中断されたジョブを再開する。()
=NO ; 最初から実行する (デフォルト)
=YES ; 途中から実行する
NAMERInput= (リスタートファイルの名前、80 文字以内)()
NAMEROutput= (リスタートファイルの名前、80 文字以内)()

(8) 探索対象の指定

指定した 1、2、3 個の蛋白原子が探索の対象となるように、探索対象を指定します。

SETTARget : 探索指定情報ファイルの読み込み ()
=NOREad ; 使用しない (デフォルト)
=READ ; 使用する
NAMETARget=(探索指定情報ファイルのファイル名、80 文字以内 ())

(9) 蛋白質の原子半径の係数

蛋白質の原子半径の係数を指定します。

DAMPPA : 蛋白質の原子半径の係数 ()
=1.0 ; (デフォルト)

(1 0) ログ出力形式の指定

標準出力に出力されるログの形式を指定します。

LOGFORmat : ログ出力形式

=SHORT ; 簡易出力
=MEDIum ; 警告を出力しない
=DETAil ; 通常出力 (デフォルト)

“ LOGFOR= SHOR ” の場合のログ出力例を p.60 以降に掲載しています。

(1 1) QUIT

PHASE> INPUT グループの入力の終わりを示します。

1.2.1.2 PHASE> GRID グループ

sievgene では蛋白質の原子によって発生するポテンシャルを Grid 上に分配した Grid ポテンシャルを生成し、Grid ポテンシャルとリガンド分子の間でポテンシャルの評価を行います。ここでは Grid ポテンシャル生成のパラメータの指定を行います。

GRID グループでの指定事項：

- (1) Grid ポテンシャルの入出力
- (2) レセプター原子登録範囲
- (3) Grid ポテンシャル生成時の原子半径オフセット
- (4) A.S.A.のメッシュ点生成の半径オフセット
- (5) Grid ポテンシャルのサイズのマージン
- (6) Grid ポテンシャルのスミージング回数
- (7) van der Waals 半径の係数
- (8) PB (Poisson Boltzmann) 方程式指定
- (9) PB-Grid ポテンシャルの入出力
- (10) メッシュ数
- (11) PB-Grid ポテンシャルのサイズのマージン
- (12) Debye-Huckel の遮蔽定数
- (13) 誘電率
- (14) 加速因子
- (15) CPU time の上限値
- (16) 静電場計算終了条件の閾値
- (17) Grid ポテンシャルの補間方法
- (18) QUIT

(1) Grid ポテンシャルの入出力

Grid ポテンシャルは蛋白質の形状および座標によって決定します。同一の蛋白質での Grid ポテンシャルは一定であるため、この Grid ポテンシャルを再利用することができます。

GRIDPOtential : Grid ポテンシャルの入力指定 ()

=NOREad ; 読み込まない (デフォルト)

=ASCI i ; ASCII 形式

=BINAr y ; BINARY 形式

OUTGRIdpotential : Grid ポテンシャルの出力指定 ()

=NOWRite ; 出力しない (デフォルト)

=ASCI i ; ASCII 形式

=BINArY ; BINARY 形式
 NAMEGridpotential= (Grid ポテンシャルファイルの名前、80 文字以内)()

【注意】ASCII 形式の場合はデータ変換により、完全に一致したデータとはなりません。

(2) レセプター原子登録範囲

レセプター原子を登録する範囲を指定します。

ポケット点から指定された範囲の蛋白質上の原子がレセプターとして登録されます。

PROBDistance : レセプター原子の登録範囲 () ()
 =6.5 ; (デフォルト)

(3) Grid ポテンシャル生成時の原子半径オフセット

蛋白質上の原子と Grid の格子点との距離により、この格子点が原子内か原子外かを判定します。

A.S.A. は原子表面に対して球を転がした場合の球の中心がとり得る面を表しますので、このときの球の半径を指定します。

RADVDW : van der Waals 半径オフセット () ()
 =0.6 ; (デフォルト)
 RADELE : coulomb 半径オフセット () ()
 =0.6 ; (デフォルト)

(4) A.S.A. メッシュ点生成の半径オフセット

sievgene での A.S.A. は A.S.A. 面上の等間隔に配置した点(メッシュ点)で表現します。各原子の表面を表現するために原子半径に加算するオフセットを指定します。

RADMEsh : メッシュ点オフセット () ()
 =1.6 ; (デフォルト)

(5) Grid ポテンシャルのサイズのマージン

Grid ポテンシャルのサイズはポケット点座標の x, y, z 座標の最大値-最小値で構成される直方体となります。

このサイズに対するマージンを指定します。

MARGIN : Grid ポテンシャルのマージン () ()
=8 ; (デフォルト)

(6) Grid ポテンシャルのスミージング回数

生成された Grid ポテンシャルは格子点ごとに計算しますので、隣り合う格子点と不自然に大きく異なる場合があります、この状態を緩和するために平滑化(スミージング)を行います。

スミージングの回数は大きいほどポテンシャル面は滑らかになりますが、大きすぎると一様なポテンシャルになります。

ITERATION : スミージング回数 ()
=3 ; (デフォルト)

(7) van der Waals 半径の係数

グリッドポテンシャル計算時の van der Waals 半径の係数を指定します。

DANPVW : van der Waals 半径の係数 ()
=1.0 ; (デフォルト)

(8) PB (Poisson Boltzmann) 方程式指定

グリッドポテンシャル生成時に、Poisson Boltzmann 方程式の使用を指定します。ドッキング精度、スクリーニング精度は、PB 方程式を使用しない場合のほうが精度は高いです。

USEPBG : PB グリッドポテンシャル指定 ()
=NO ; PB 方程式を使用しない (デフォルト)
=YES ; PB 方程式を使用する

(9) PB-Grid ポテンシャルの入出力

Poisson Boltzmann 方程式使用時の静電場計算の結果は蛋白質の形状および座標によって決定します。同一の蛋白質での静電場計算の結果は一定であるため、この計算結果を再利用することができます。

PBGRID : 静電場計算結果の入力指定 ()
=NOread ; 読み込まない (デフォルト)
=ASCII ; ASCII 形式
=BINARY ; BINARY 形式
OUTPBG : 静電場計算結果の出力指定 ()

```

=NOWrite      ; 出力しない (デフォルト)
=ASCIi        ; ASCII 形式
=BINArY       ; BINARY 形式
NAMEPB= ( PB ファイルの名前、80 文字以内 ) ( )

```

【注意】PBGRID、OUTPBG、NAMEPB の各オプションは USEPBG= YES の時のみ有効です。

【注意】ASCII 形式の場合はデータ変換により、完全に一致したデータとはなりません。

(1 0) メッシュ数

Poisson Boltzmann 方程式使用時の、x, y, z 方向のメッシュ数を指定します。

```

NMESHX : メッシュ数 ( x 方向 ) ( )
        =100      ; ( デフォルト )
NMESHY : メッシュ数 ( y 方向 ) ( )
        =100      ; ( デフォルト )
NMESHZ : メッシュ数 ( z 方向 ) ( )
        =100      ; ( デフォルト )

```

【注意】NMESHX、NMESHY、NMESHZ の各オプションは USEPBG= YES の時のみ有効です。

(1 1) PB-Grid ポテンシャルのサイズのマージン

PB-Grid ポテンシャルのサイズはポケット点座標の x, y, z 座標の最大値-最小値で構成される直方体となります。

このサイズに対するマージンを指定します。

```

MARGPB : PB-Grid ポテンシャルのマージン ( ) ( )
        =5.0d0    ; ( デフォルト )

```

【注意】MARGPB オプションは USEPBG= YES の時のみ有効です。

(1 2) Debye-Huckel の遮蔽定数

Poisson Boltzmann 方程式使用時の静電場計算における Debye-Huckel の遮蔽定数を指定します。

```

KAPPAV : Debye-Huckel の遮蔽定数 ( 1 / ) ( )
        =0.0d0    ; ( デフォルト )

```

【注意】KAPPAV オプションは USEPBG= YES の時のみ有効です。

(1 3) 誘電率

Poisson Boltzmann 方程式使用時の静電場計算における溶媒、蛋白質、中間領域、真空領域の誘電率をそれぞれ指定します。

DIESOL : 溶媒の誘電率 ()
=78.5d0 ;(デフォルト)
DIEPRO : 蛋白質の誘電率 ()
=4.0d0 ;(デフォルト)
DIEINT : 中間領域の誘電率 ()
=78.5d0 ;(デフォルト)
DIEVAC : 真空領域の誘電率 ()
=1.0d0 ;(デフォルト)

【注意】DIESOL、DIEPRO、DIEINT、DIEVAC の各オプションは USEPBG= YES の時のみ有効です。

(1 4) 加速因子

Poisson Boltzmann 方程式使用時の静電場計算の加速因子を指定します。

ACCELE : 加速因子 ()
=1.6d0 ;(デフォルト)

【注意】ACCELE オプションは USEPBG= YES の時のみ有効です。

(1 5) CPU time の上限値

Poisson Boltzmann 方程式使用時の静電場計算の CPU time の上限値を指定します。

CPULIM : CPU time の上限値 (s)()
=10800 ;(デフォルト)

【注意】CPULIM オプションは USEPBG= YES の時のみ有効です。

(1 6) 静電場計算終了条件の閾値

Poisson Boltzmann 方程式使用時の静電場計算終了のための閾値を指定します。

SORRE1 : 終了条件の相対閾値(SOR) ()
=1.0d-8 ;(デフォルト)

SORRE2 : 終了条件の相対閾値(SOR) ()
=1.0d-10 ;(デフォルト)

SRFAB1 : 終了条件の絶対閾値(SURF) ()
=1.0d-1 ;(デフォルト)

SRFAB2 : 終了条件の絶対閾値(SURF) ()
=1.0d-3 ;(デフォルト)

SRFRE1 : 終了条件の相対閾値(SURF) ()
=1.0d-8 ;(デフォルト)

SRFRE2 : 終了条件の相対閾値(SURF) ()
=1.0d-10 ;(デフォルト)

(注) SORRE1、SORRE2、SRFAB1、SRFAB2、SRFRE1、SRFRE2 の各オプションは
USEPBG= YES の時のみ有効です。

(1 7) Grid ポテンシャルの補間方法

Grid ポテンシャル上の任意の点における Grid ポテンシャルの補間方法を指定します。

INTERPorate : Grid ポテンシャルの補間方法 ()
=LAGRange ; Lagrange の 1 次補間 (デフォルト)
=BSPLine ; 3 次の B-Spline 補間

(1 8) QUIT

PHASE> GRID グループの入力の終わりを示します。

1.2.1.3 PHASE> CONF グループ

sievgene では入力したリガンドから回転可能な構造を検出し、この構造をランダムに回転させた新たな配座でのスクリーニングを行います。

ここでは新規配座生成のパラメータの指定を行います。

CONF グループでの指定事項：

- (1) 配座生成試行回数の指定
- (2) 生成する配座数の指定
- (3) 原子データのソート実行
- (4) van der Waals 接触判定の原子半径への係数
- (5) 構造の回転角数
- (6) 原子モデルの指定
- (7) QUIT

(1) 配座生成試行回数の指定

配座生成では回転角を乱数で決定するため、同一の配座を複数生成してしまい配座生成処理が終了しない場合があります。この状態を抑止するために配座生成試行回数の上限を指定します。

CONFLimit : 配座生成試行回数上限 ()
=10000000 ;(デフォルト)

(2) 配座する生成数の指定

生成する配座数を指定します。

COMFORMernumber : 配座生成数 ()
=100 ;(デフォルト)

(3) 原子データのソート実行

出力ファイルでの原子の並びを入力原子の並びと一致させることを指定します。

SORTATom : 原子データのソート ()
=NO ;並びを元に戻さない(デフォルト)
=YES ;並びを元に戻す

(4) van der Waals 接触判定の原子半径への係数

新規配座生成でリガンドの原子同士が van der Waals 接触する場合は、この配座を採用しません。van der Waals 接触判定で原子間距離と原子半径の和を比較する際、原子半径にこの係数をかけて比較します。

DAMPINGfactor : van der Waals 接触判定の係数 ()
=0.7 ; (デフォルト)

(5) 構造の回転角数

構造(Torsion)を回転する際にとり得る角度の個数を指定します。デフォルトの 3 の場合は 120° の単位で構造が回転します。

PHASETorsion : 構造の回転角数 ()
=3 ; (デフォルト)

(6) 原子モデルの指定

リガンド分子の原子モデルを指定します。従来の原子モデルである「all atom model」および、-CH, -CH₂, -CH₃ を '?C' に置換する「united atom モデル」のいずれかを指定できます。

AToMMoDeL : 原子モデル
= ALL atom model ; all atom モデル (デフォルト)
= UNITed atom model ; united atom モデル

(7) QUIT

PHASE> CONF グループの入力の終わりを示します。

1.2.1.4 PHASE> DOCK グループ

グローバルサーチでのパラメータ指定およびスコア計算のパラメータ指定を行います。

DOCK グループでの指定事項：

- (1) ドッキング方法の指定
- (2) 蛋白側接合点発生方法
- (3) 絞込み回数
- (4) 接合面数
- (5) 結合面の辺の長さ
- (6) 結合面の原子タイプ適合度
- (7) スコア係数
- (8) ポケット内原子判定距離
- (9) ポケット中心指定
- (10) 蛋白とリガンドの水素結合の評価
- (11) ドッキング時のリガンド-OH 基の回転
- (12) ドッキング時の蛋白質側鎖の回転
- (13) ドッキング後のリガンド座標の微調整
- (14) ローカルサーチ対象数
- (15) ドッキング対象原子の切り替え
- (16) エネルギーの見積もり
- (17) 各分子のドッキング計算時間の上限
- (18) QUIT

(1) ドッキング方法の指定

sievgene では入力したリガンドの配座を変更しない RIGID ドッキングと、配座を変えながらドッキングを繰り返す FLEXIBLE ドッキングの 2 種類が指定できます。

```
METHODofdocking : ドッキング方法指定 ( )  
=FLEXible      ; FLEXIBLE ドッキング ( デフォルト )  
=RIDId        ; RIGID ドッキング
```

(2) 蛋白側接合点発生方法

sievgene では接合面を 3 点の構成する三角形として規定しています。

ドッキングの際の、蛋白側の接合点の発生方法を選択します。

PROteinSURface : 蛋白側接合点発生方法

=ELECTrostatic ; 蛋白の Accessible surface 上の静電場の極大・極小点および電位 = 0 の点を採用する方法 (デフォルト)

=HYDRogen ; 蛋白の Accessible surface 上の静電場の極大・極小点および CH₄ プローブのポテンシャル面の極大点を採用する方法

(3) 絞込み回数

探索範囲を絞込む回数を指定します。

GENERationnumber : 絞込み回数 ()

=5 ; (デフォルト)

(4) 接合面数

sievgene では接合面を 3 原子の構成する三角形として規定しています。

ドッキングの際にリガンドの接合面を選択する個数を指定します。

NUMCONformer : 配座生成数 ()

=100 ; (デフォルト)

(5) 接合面の辺の長さ

接合面の数を満たす三角形の辺の長さを指定します。sievgene では与えられた辺の範囲から接合面の個数を満たす辺の長さを計算し、接合面を決定します。

LOWMIN : 辺の長さの下限側の最小値 () ()

=2.5 ; (デフォルト)

LOWMAX : 辺の長さの下限側の最大値 () ()

=4.0 ; (デフォルト)

UPRMIN : 辺の長さの上限側の最小値 () ()

=7.5 ; (デフォルト)

UPRMAX : 辺の長さの上限側の最大値 () ()

=10.0 ; (デフォルト)

(6) 結合面の原子タイプ適合度

sievene では原子の特性により接合に関連するかどうかの重みを設定しています。

この適合度を調節することで、不要な接合判定が減少しドッキング時間を減少させることができます。

1 から 5 までの指定が可能で、値が大きいほど接合面が限定されます。

MATCHING : 結合面の原子タイプ適合度 ()
=1 ;(デフォルト)

(7) スコア係数

sievene でのスコアは A.S.A.、coulomb、水素結合、異方性を考慮した水素結合、van der Waals の 5 つのポテンシャルの和で表現されます。

各ポテンシャルに重みを設定したい場合はスコア係数を変更します。。

WETVDW : van der Waals ポテンシャル係数 ()
=1.0 ;(デフォルト)

WETASA : A.S.A. ポテンシャル係数 ()
=1.0 ;(デフォルト)

WETHYD : 水素結合ポテンシャル係数 ()
=1.0 ;(デフォルト)

WETANH : 異方性を考慮した水素結合ポテンシャル係数 ()
=1.0 ;(デフォルト)

WETELE : coulomb ポテンシャル係数 ()
=1.0 ;(デフォルト)

(8) ポケット内原子判定距離

リガンド原子がポケット内に存在する判定でのポケット中心点からの距離を指定します。

RADIUS : ポケット内原子判定距離 () ()
=6.0 ;(デフォルト)

(9) ポケット中心指定

蛋白質のポケット中心点を指定します。

指定されない場合、ポケット点の midpoint がポケット中心に仮定されます。

(9 - 1) ポケット中心座標指定

ポケット中心の絶対座標を指定します。

POCKCX : ポケット中心 X 座標
= 999.9 ; (デフォルト)
POCKCY : ポケット中心 Y 座標
= 999.9 ; (デフォルト)
POCKCZ : ポケット中心 Z 座標
= 999.9 ; (デフォルト)

(9 - 2) ポケット中心座標指定

ポケット中心の蛋白原子を指定します。

POCKET : ポケット中心原子 ID
=0 ; (デフォルト)

(1 0) 蛋白とリガンドの水素結合の評価

蛋白質とリガンドの、異方性を考慮した水素結合の評価を指定します。

EVALHB : 異方性を考慮した水素結合の評価 ()
=YES ; 評価する (デフォルト)
=NO ; 評価しない

(1 1) ドッキング時のリガンド-OH 基の回転

ドッキング時の、リガンド-OH 基の末端の回転を指定します。

ROTLOH : リガンド-OH 基の回転 ()
=YES ; 回転する (デフォルト)
=NO ; 回転しない

(1 2) ドッキング時の蛋白質側鎖の回転

ドッキング時の、蛋白質の水素結合可能な側鎖の回転を指定します。

ROTPSC : 蛋白質側鎖の回転 ()
=YES ; 回転する

=NO ; 回転しない (デフォルト)

(13) ドッキング後のリガンド座標の微調整

ドッキング後のリガンド座標の微調整における、座標をずらす回数を指定します。

MOVNUM : 座標をずらす回数 ()
=10 ; (デフォルト)

(14) ローカルサーチ対象数

ローカルサーチ対象数を指定します。

CANDID : ローカルサーチ対象数 ()
=30 ; (デフォルト)

(15) ドッキング対象原子の切り替え

蛋白とリガンドの重ね合わせの際に、重ね合わせの対象外となる原子を計算対象から除外するかしないかを指定します。

DOCKSP : ドッキング対象原子の切り替え ()
=NORM ; 対象外となる原子を除外しない (デフォルト)
=FAST ; 除外する

(16) エネルギーの見積もり

入力原子に対して、ドッキングを行わずに Minimize 計算のみを行います。

OPTIME : エネルギー見積もり ()
=OPT ; ドッキングを行う (デフォルト)
=ENE ; ドッキングを行わない

(17) 各分子のドッキング計算時間の上限

1分子あたりのドッキング計算時間に上限を設けます。制限時間を越えると計算を打ち切り、途中の計算結果を記録します。Multi-mol2 file の場合は、1分子あたりの制限時間です。

CPUTIM : ドッキング時間の制限値(秒) ()
=30 ; (デフォルト)

(1 8) QUIT

PHASE> DOCK グループの入力の終わりを示します。

1.2.1.5 EXE> MIN グループ

ローカルサーチでのパラメータ指定およびスコア計算のパラメータ指定を行います。

MIN グループでの指定事項および指定方法は cosgene の EXE> MIN グループと同じです。

ユーザマニュアル「cosgnene」の章の「EXE> MIN グループ」の本項を参照。

1.2.1.6 PHASE> OUTPUT グループ

OUTPUT グループでは、上位 conformer 座標の PDB ファイル、スコアファイルの出力指定およびクラスタリング機能の指定を行います。

OUTPUT グループでの指定事項：

- (1) 上位スコア conformer 座標の PDB ファイル出力
- (2) 上位スコアの出力
- (3) 化合物順位の出力
- (4) クラスタリング機能
- (5) QUIT

(1) 上位 conformer 座標の PDB ファイル出力

COORDinate：上位スコア conformer 座標の PDB ファイル出力の書式 ()

=NOWrite ; 出力しない(デフォルト)

=MOL2 ; MOL2 ファイル形式

=PDBX ; PDBX 形式(cosgene トポロジーファイル+PDB 書式を XML で統合)

=PDB ; PDB 形式

CANDIDatenumber：上位スコア conformer 数 ()

=30 ; (デフォルト)

NAMECOordinate=(上位 conformer 座標ファイル名、80 文字以内。())

【注意】COORDI= MOL2 は、LIGAND= MOL2 のときのみ有効です。

(2) 上位スコアの出力

SCORENumber：出力上位スコア数 ()

=30 ; (デフォルト)

NAMEScore=(上位スコアファイル名、80 文字以内。())

(3) 化合物順位の出力

入力リガンドファイルに記載されている分子ごとのドッキング最上位のスコアを比較し、上位の分子名およびスコア情報を「化合物順位」として出力します。

COMRNK : 出力上位化合物数
=1 ; (デフォルト)

【注意】COMRNK オプションでの指定が入力リガンドファイルに記載されている分子数より大きい場合、「化合物順位」は入力リガンドファイルに記載されている分子数分を出力します。

(4) クラスタリング機能

多様なドッキングポーズを得るため、ドッキングポーズ間の RMSD を指標に構造クラスタリングを適用し、指定したクラスタ群のポーズで、最も良いスコアのポーズを出力します。

CLUSTEr ing : クラスタリングの実行
=NO ; 実行しない(デフォルト)
=YES ; 実行する

CLUster ingMEThod : クラスタリング方法

=NEAR
=FURTest
=MEDian
=CENTroid
=AVERage
=WARD

NCLUStEr : クラスタ数
=1 ; (デフォルト)

【注意】クラスタ数指定は DOCK フェーズでの出力候補数(CANDID オプションで指定)をクラスタリング対象構造数とし、OUTPUT フェーズでの出力構造数をクラスタ数とします。

(5) QUIT

PHASE> OUTPUT グループの入力の終わりを示します。

2 計算例

2.1 Sample-1: 受容体 - 低分子化合物ドッキング

蛋白質などの受容体に、低分子化合物をドッキングさせる手法を説明します。1ai5/1c1e は、それぞれ受容体・低分子化合物のデータを含んだディレクトリーです。各ディレクトリーに移動してジョブを実行してください。また、allrun.sh にて、2つの複合体についてテスト計算ができます。いずれの場合も出力結果は、各受容体・低分子化合物のデータを含んだディレクトリーに保存されます。

(1) グリッドポテンシャルの生成からドッキングまで一括で行なう方法

まず、INPUT フェーズの設定について説明します。ドッキングさせるべき低分子は、mol2 ファイル書式(<http://www.tripos.com/custResources/mol2Files/>)で準備してください。分子モデルは、水素原子を含む全原子モデルで作成し、原子電荷は、必ず記入してください。mol2 ファイル書式のうち、<MOLECULE>,<ATOM>,<BOND>のセクションがあれば十分です。低分子に関して、PDB ファイルと tpl ファイルがあれば、myPresto 付随のツール"tpl2mol2"で、mol2 ファイルを生成できます。また、フリーの babel/openBabel などでも mol2 ファイルを準備できます。

【注意】 sievgene では、金属錯体を扱うことができません。

蛋白質の準備ですが、水素を含む全原子モデルで準備してください。全原子の PDB ファイルと電荷の記載されたトポロジーファイルを準備します。これらは myPresto-tplgene などでも準備できます。

リガンドは、ファイル形式(LIGAND=MOL2)とファイル名(NAMELI)を指定してください。ドッキング後の配座と、参照座標との位置のずれ RMSD を計算することができます。正解の受容体 リガンド複合体があり、リガンドの座標が既知である場合、正解の座標を参照座標とすることでドッキングの精度を確認することができます。参照座標はそのファイル形式(REFERE=MOL2)とファイル名(NAMERE)を指示してください。受容体側は、トポロジーファイルの書式(TOPOLO=FORM)とファイル名(NAMETO)、座標の書式(COORDI=PDB)とファイル名(NAMECO)を指定してください。

受容体ポケットの位置は、プローブ点の集合で指定し、サンプルでは、point.pdb という PDB 書式での点の集合で与えられています。リガンドの構造探索空間とポテンシャルグリッドは、このプローブ点の集合に対し、x,y,z 軸方向に次のグリッドフェーズで指定する、±

PROBDI/MARGIN の範囲に設定されます。既知リガンドとの複合体座標があれば、そのリガンドの座標をそのまま用いれば良いです。既知リガンドの情報がない場合、make_point.f を用いて PDB 書式でのプローブ点の集合を作成することができます。

コンパイルの仕方：

```
% f90 make_point.f ?o make_point.exe
```

使い方：

make_point とタイプします。標準入力から順次、受容体のファイル名、プローブ点を発生させる球状領域の半径、中心とすべき原子の番号を指定すれば、point.pdb という固定の名前で、PDB 書式でのプローブ点の集合を出力します。

```
% make_point.exe
```

この point.pdb は、より実際に合うようにエディターなどで編集されると良いでしょう。プローブ点の集合は、そのファイル書式(POINTC)と、ファイル名(NAMEPO)を指定します。

sievgene では、ドッキング結果の構造について、蛋白質がわの特定の原子の近傍半径 6 以内に、ドッキングしたリガンドの原子が何個含まれるのかを表示する機能がありますが、下記例では、指定しない(SETTAR=NORE)ことにしています。蛋白質側の原子の半径は、
 $(\text{蛋白質側の原子の半径}) = (\text{ダンピング係数}) \times (\text{デフォルトの設定半径})$
 になっており、ダンピングの係数 (DAMPPA) で調整できます。

GRID の生成フェーズでは、あらかじめ生成したグリッドポテンシャルを読み込む場合 (GRIDPO=BINA)と、新たに生成する場合 (GRIDPO=NORE)、その場合グリッドポテンシャルをファイル出力する場合 (OUTGRI=BINA)、出力しない場合 (OUTGRI=NOWR)があります。

リガンド探索の空間として、幾何学ハッシングテーブルを生成する範囲を指定しますが、これは INPUT フェーズで指定したプローブ点(NAMEPO)を基準に、各点を中心に半径 PROBDI の範囲で指定されます。同じく、グリッドポテンシャルを生成する範囲も、プローブ点 (NAMEPO)を基準に、x,y,z 方向に ±MARGIN の範囲で指定されます。PROBDI を大きくとると、構造探索範囲が広がり、計算時間が増大します。最初にプローブ点の集合を決める時点で、必要最小限の最適な選び方にすることが大切です。グリッドの大きさは 60x60x60 に固定されているため、MARGIN を大きくとると、グリッド幅が広がって精度が低下します。グリッド幅は、0.2~0.35 程度にとり、0.5 を超えない大きさにしてください。グリッドポテンシャルは、ITERAT で指定した回数だけ数値的にスムージングされます。van der Waals ポテンシャル及び、静電ポテンシャルは、蛋白質の界面を基準に内外に分割され、外部領域では通常のポテンシャル、内部領域では滑らかな仮想ポテンシャルが適用されます。こ

のときの、界面である accessible surface のプローブ半径は、それぞれ、RADVDW/RADELE で指定されます。RADVDW/RADELE を小さくとると、本来蛋白質の内部の領域でも原子間隔の広いところが外部と判定されます。0.6 程度が最適です。この時の界面生成での蛋白質原子のダンピング係数は DAMPVW で指定されます。幾何学ハッシングテーブル作成時には、蛋白質表面に疎水性プローブ、親水性正電荷プローブ、親水性負電荷プローブを転がしてそれぞれの結合しやすい点の集合を作成し、それらの点を結ぶ三角形の辺をハッシングテーブルに格納しますが、そのときの点の間隔は RADMESH で指定します。RADMESH は、1.2~1.6 程度、通常 1.4 で良い結果が得られます。

静電場は USEPBG=NO の場合、 $\epsilon = 4R$ とした式で与えられ、USEPBG=YES の場合、Poisson-Boltzmann 方程式を self-consistent boundary(SCB)法で解いた電場が適用されます。現在、SCB 法は速度・精度とも $\epsilon = 4R$ とした式に劣っています。

CONF フェーズでは、リガンドの配座生成の方法を指定します。配座生成は、分子の環を除く直鎖部分の 1 重結合の 2 面角を回転することで行い、ランダムサーチが行なわれます。ランダムサーチでの試行回数は上限を CONFLI 回とし、一度に生成する配座数は最大 CONFOR 個とします。2 面角は 360/PHASET で指定した回転角で回転します。例では PHASET=3 なので、120 度ずつ回転します。配座生成時に入力原子を tree 状のグラフとなるように並べ替えますが、結果の出力時には、元の原子の順番に戻すことができます (SORTAT=YES)。生成した配座について、分子内衝突の判定を行い、分子内で原子間衝突のある配座を除外しますが、このときの van der Waals 半径のダンピング係数は DAMPIN で指定します。配座生成は 2 面角のみのラフな回転で生成するため、分子内衝突が起き易く、DAMPIN は 0.7 程度にとるとよいでしょう。ROTTER は、-OH 基のようなリガンドの端を回転させるかどうかの指定です。ROTTET で-OH 基を回転させなくても、次の DOCK フェーズでドッキング後に-OH 基を回転させて構造探索することができます。

DOCK フェーズでは、幾何学ハッシングによる受容体 低分子のラフなドッキングを行ないます。リガンド側をフレキシブルにしたフレキシブルドッキング (METHOD=FLEX) と、リガンドを入力配座に固定したリジッドドッキング (METHOD=RIGID) の両方ができます。CONF フェーズで生成した CONFOR 個の配座のドッキングを試してから、もっともらしいドッキング位置に探索空間を絞り、再度、CONF フェーズで異なる CONFOR 配座を発生してドッキングする操作を GENERAT 回繰り返します。従って、ドッキングで試行される配座の総数は最大 CONFORxGENERAT 配座になります。ハッシュ表と参照するリガンド原子間の長さには、制限をつけます。最短の長さは、LOWMIN 以上 LOWMAX 以下、最長の長さは、UPRMIN 以上 UPRMAX 以下です。この長さは、1 配座について、ハッシュ表と参照する回数が NUMCONF 回程度になるように上記の範囲内で調整します。ハッシュ表には、受容体側の三角形をなす各点が H ドナーか、H アクセプターの原子近傍にあるかどうかの情報が記載されており、リガンド側原子は、H ドナーか、H アクセプターかが判別されています。これらの受容体とリガンドの 3 点の重ね合わせを行なうとき、何点で H ドナーと H アクセプターとが対をなすかどうか

が計算されます。H ドナーと H アクセプターの完全一致は 2 点、不完全一致では 1 点、H ドナーと H ドナー、H アクセプターと H アクセプターのような不適合は 0 点となっています。これら適合度の 3 点の合計が、MATCHING 以上の時に、構造重ね合わせを採用します。MATCHING は、0 ~ 6 の範囲の整数で指定します。親水性が高く、結合モードが多くの水素結合を含む場合は、MATCHING を多く、ポケットが疎水性で水素結合をほとんど含まない場合は MATCHING を少なくとってください。

WETVDW/WETASA/WETELE/WETHYD/WETAHB は、総スコアを計算するときの、van der Waals 相互作用項、ASA 項、静電項、水素結合項、水素結合の方向を考慮した水素結合項の重みの指定です。EVALHB は、水素結合の方向を考慮した水素結合の計算の指定、ROTLOH は、ドッキング後に、-OH 基を回転させて構造探索をするかどうかを指定します。ROTPSC は、ドッキング後に蛋白質側の側鎖を回転させて、蛋白質側の構造変化も考慮したフレキシブルドッキングを行なうかどうかを指定します。MOVNUM は、リガンドドッキング後に、x, y, z 軸方向に 0.5 の範囲でランダムに構造探索を行います。その回数の指定です。

MIN フェーズでは、リガンドドッキングの後、もっともらしい構造について最急勾配法 (METHOD=STEEP) または、共役勾配法 (METHOD=CONJ) で、エネルギー最小化を行なう指定です。このフェーズでの指定は cosgene の MIN フェーズと同じです。

OUTPUT フェーズでは、ドッキング結果の出力を指定します。ドッキングスコア上位の CANDID 個について、ファイル出力形式 (COORDI=MOL2) で指定した形式で、出力ファイル名 (NAMECO) に座標を出力します。ドッキングスコアそのものは、SCOREN で指定した個数を、NAMESC で指定したファイルに出力します。

sievgene の実行 :

入力ファイルを all.inp とします。

```
% sievgene < all.inp
```

入力ファイル sievgene_for_dockingpose の場合

```

PHASE> INPUT
LIGAND = MOL2
NAMELI = Lig_es_1.mol2 ; リガンドの座標位置ファイル
REFERE = MOL2
NAMERE = lig_ref.mol2 ; リガンドの参照位置ファイル
TOPOLO = FORM
NAMETO = Pro.tpl ; 蛋白側のトポロジー
COORDI = PDB
NAMECO = Pro_md.pdb ; 蛋白側の座標
POINTC = PDB
NAMEPO = point.pdb ; ポケット領域を示すプローブ点

SETTAR = NORE
DAMPPA = 1.0d0
QUIT
;
; grid generation and Hash table generation
;
PHASE> GRID
GRIDPOtential = NORE ; BINA ; Grid ポテンシャル入力
NAMEGRid = grid.file ; Grid ポテンシャルファイル名
OUTGRIDpotential = NOWR ; Grid ポテンシャル出力
PROBDist = 6.5 ; プローブ点から PROBDIST 内のレセプター原子を対象にする。
MARGIN = 6.5 ; Grid 生成時の、探索ポケット周辺にとるマージン
ITERAT = 3 ; Grid ポテンシャルのスムージングの繰り返し回数
RADVDW = 0.6 ; vDW ポテンシャルの内外領域の界面の設定
RADELE = 0.6 ; クーロンポテンシャルの内外領域の界面の設定
RADMESH = 1.4 ; mesh 点生成のプローブ半径

DAMPVW = 0.99d0

USEPBG = NO ; no use of PB

PBGRID = BINA ; for PB
OUTPBG = NOWR ; for PB
NAMEPB = #NAMEPB# ; for PB
NMESHX = 100 ; for PB
NMESHY = 100 ; for PB
NMESHZ = 100 ; for PB
MARGPB = 5.0d0 ; for PB
KAPPAV = 0.0d0 ; for PB
DIESOL = 78.5d0 ; for PB
DIEPRO = 4.0d0 ; for PB
DIEINT = 78.5d0 ; for PB
DIEVAC = 1.0d0 ; for PB
ACCELE = 1.6d0 ; for PB
CPUTIM = 10800 ; for PB
QUIT
;

```

(つづく)

(つづき)

```

PHASE> CONF
  CONFLimit      = 100000 ; 配座生成の試行最大数
  CONFORMernumber = 100   ; 生成する配座数(リガンドを変形)
  SORTATom       = YES    ; 出力ファイルの原子順を変える
  DAMPING        = 0.7    ; 配座候補の原子間距離のダンピングファクター
  PHASETorsion   = 3      ; conformer 生成時の torsion 回転候補の個数
  ROTTER         = YES
  QUIT

PHASE> DOCK
  METHOD          = FLEX
  GENERAtion     = 1      ; 絞込みの回数
  NUMCONFomer    = 1000 ; リガンドの接合面三角形数
  MATCHING       = 4      ; 3 角形の原子タイプの適合度
  LOWMIN         = 2.5    ; Hash 表と比較する原子間距離の下限の最小値
  LOWMAX         = 3.5    ; Hash 表と比較する原子間距離の下限の最大値
  UPRMIN         = 5.0    ; Hash 表と比較する原子間距離の上限の最小値
  UPRMAX         = 12.0   ; Hash 表と比較する原子間距離の上限の最大値

  WETVDW= 1.0 ; 総合スコア計算時の各項の重み
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  RADIUS= 6.0 ; リガンド分子ドッキングの後 標的原子(座標)から RADIUS内
               ; に存在するリガンドの原子の数を数える。

  EVALHB = NO
  WETANH = 1.0d0
  ROTLOH = NO
  ROTPSC = NO
  MOVNUM = 10
  QUIT

EXE> MIN
  METHOD= STEEP      CPUTIM = 360000.0
  UPRATE= 1.0       DOWNRATE= 0.3
  LOOPLI= 100      UPDATE = 20
  MONITO= 100      CONVGR = 0.1D0
  CUTMET= RESA     CUTLEN = 22.0D0
  DIEFUN= DIST     DIEVAL = 4.0D0
  LOGFOR= SHOR
  BESTFI= YES

  QUIT

PHASE> OUTPUT
  COORDInate      = MOL2 ; リガンド座標のファイル出力形式
  NAMECOordinate = ex.cor ; 各実験でのリガンド座標ファイル名
  CANDIDatenumbeR= 3     ; 各実験でのリガンド座標を出力する上位スコア数
  SCORENumber     = 3     ; score ファイルに出力する上位スコア数
  NAMESCore       = ex.score ; score ファイル名
  QUIT

EXE> SIEV

```


入力ファイル sievgen_for_screening の場合

```
PHASE> INPUT
LIGAND = MOL2
NAMELI = Lig_es_1.mol2 ; リガンドの座標位置ファイル
REFERE = MOL2
NAMERE = lig_ref.mol2 ; リガンドの参照位置ファイル
TOPOLO = FORM
NAMETO = Pro.tpl ; 蛋白側のトポロジー
COORDI = PDB
NAMECO = Pro_md.pdb ; 蛋白側の座標
POINTC = PDB
NAMEPO = point.pdb ; ポケット領域を示すプローブ点

SETTAR = NORE
DAMPPA = 1.0d0
QUIT
;
; grid generation and Hash table generation
;
PHASE> GRID
GRIDPOtential = NORE ; BINA ; Grid ポテンシャル入力
NAMEGRid = grid.file ; Grid ポテンシャルファイル名
OUTGRIdpotential = NOWR ; Grid ポテンシャル出力
PROBDist = 6.5 ; プローブ点から PROBDIST 内のレセプター原子を対象にする。
MARGIN = 6.5 ; Grid 生成時の、探索ポケット周辺にとるマージン
ITERAT = 3 ; Grid ポテンシャルのスムージングの繰り返し回数
RADVDW = 0.6 ; vDW ポテンシャルの内外領域の界面の設定
RADELE = 0.6 ; クーロンポテンシャルの内外領域の界面の設定
RADMESH = 1.4 ; mesh 点生成のプローブ半径

DAMPVW = 0.99d0

USEPBG = NO ; no use of PB

QUIT
;
PHASE> CONF
CONFLimit = 100000 ; 配座生成の試行最大数
CONFORMernumber = 100 ; 生成する配座数(リガンドを变形)
SORTATom = YES ; 出力ファイルの原子順を変える
DAMPING = 0.7 ; 配座候補の原子間距離のダンピングファクター
PHASETorsion = 3 ; conformer 生成時の torsion 回転候補の個数
ROTTER = YES
QUIT
```

(つづく)

(つづき)

```
PHASE> DOCK
METHOD      = FLEX      PROSUR = HYDR
GENERATION  = 1        ; 絞込みの回数
NUMCONFomer = 1000    ; リガンドの接合面三角形数
MATCHING    = 4        ; 3角形の原子タイプの適合度
LOWMIN      = 2.5     ; Hash表と比較する原子間距離の下限の最小値
LOWMAX      = 3.5     ; Hash表と比較する原子間距離の下限の最大値
UPRMIN      = 5.0     ; Hash表と比較する原子間距離の上限の最小値
UPRMAX      = 12.0    ; Hash表と比較する原子間距離の上限の最大値

WETVDW= 1.0 ; 総合スコア計算時の各項の重み
WETASA= 1.0
WETELE= 3.2 ; sievgene_for_dockingposeと重みが異なることに注意
WETHYD= 1.2 ; sievgene_for_dockingposeと重みが異なることに注意
RADIUS= 6.0 ; リガンド分子ドッキングの後 標的原子(座標)から RADIUS内
              ; に存在するリガンドの原子の数を数える。

EVALHB = NO
WETANH = 1.0d0
ROTLOH = YES
ROTPSC = NO
MOVNUM = 100 CANDID= 10
QUIT

EXE> MIN
METHOD= STEEP      CPUTIM = 360000.0
UPRATE= 1.0        DOWNRATE= 0.3
LOOPLI= 50         UPDATE = 20
MONITO= 50         CONVGR = 0.1D0
CUTMET= RESA       CUTLEN = 22.0D0
DIEFUN= DIST       DIEVAL = 16.0D0
LOGFOR= SHOR

QUIT

PHASE> OUTPUT
COORDinate   = MOL2      ; リガンド座標のファイル出力形式
NAMECOordinate = ex.cor  ; 各実験でのリガンド座標ファイル名
CANDIDatenumbe= 3       ; 各実験でのリガンド座標を出力する上位スコア数
SCORENumber  = 3        ; scoreファイルに出力する上位スコア数
NAMESCore    = ex.score  ; scoreファイル名
QUIT

EXE> SIEV
```

実行結果出力例（標準出力）

実行結果の標準出力の末尾は、以下ようになります。全スコア(SCORE)、 G スコア(dG)、 hit-optimized スコア(HIT)、 MTS スコア(MTS)、 回転可能結合数(rotNum)、 ASA 項(ASA)、 静電項(ELE)、 水素結合項(HYD)、 van der Waals 項(VDW)、 リガンド分子の accessible surface の何%が受容体との結合に埋まっているか(SURFACE)、 参照座標との水素原子を除いた重ね合わせのずれ(RMSD、) (RMSD)が表示されます。また、各フェーズに要した CPU 時間などが表示されます。

```

INFORMATION> AFTER MINIMIZE RANKING
INFORMATION> AFTER MINIMIZE RANKING :EXPERIMENT ID=1
INFORMATION> LOCAL SCORE RANKING          3

COMPOUND NAME : conf1.mol2
FILE NAME     : Lig_es_1.mol2

      SCORE(/100) dG   HIT   MTS   rotNum  ASA     ELE   HYD   VDW   SURFACE  RMSD
@  1  -2.70  -7.33 -71.60 -260.69   1  -232.43 -2.73 -33.19 -1.62  84.50  1.00
@  2  -2.66  -7.16 -69.10 -255.77   1  -233.92 -0.91 -29.32 -1.58  85.80  1.09
@  3  -2.65  -7.19 -69.57 -255.85   1  -227.70 -2.29 -33.02 -1.70  85.04  0.97

REFERENCE COORDINATE DATA
-233.23 -210.82  1.23 -23.64

INFORMATION> ALL EXPERIMENT FINISHED

INFORMATION> COMPOUND RANKING          1

      COMPOUND NAME                SCORE(/100) dG   HIT   MTS
1 conf1.mol2                       -2.70  -7.33 -71.60 -260.69

TOTAL EXPERIMENT NUMBER :          1
TOTAL TRIAL NUMBER      :          904
TOTAL CPU TIME(S)      : 1416.448

```

(つづく)

(つづき)

INPUT	:	21.42383
GRID	:	1369.997
CONF	:	3.9062500E-03
DOCK	:	0.1435547

実行結果出力例（スコアファイル）

NAMESC で指定したスコアファイルの出力は以下のとおりです。各項の定義は標準出力と同じです。最初に、幾何学ハッシュによるドッキング結果、次にエネルギー最小化後のドッキング結果、最後にもう一度、最終結果を表示します。最終結果は、エネルギー最小化を行っていないければ幾何学ハッシュによるドッキング結果を、エネルギー最小化を行っていたらエネルギー最小化後のドッキング結果になります。

```

INFORMATION> BEFORE MINIMIZE RANKING :EXPERIMENT ID=1
INFORMATION> LOCAL SCORE RANKING          3

      SCORE(/100)  dG      HIT      MTS      rotNum  ASA      ELE      HYD      VDW      SURFACE  RMSD
1   -2.29   -6.23  -53.03  -223.13    1   -205.27   3.91  -27.76   0.00   0.00   5.86
2   -2.26   -6.17  -52.03  -220.84    1   -202.19   4.00  -28.23   0.00   0.00   6.03
3   -2.19   -5.89  -53.12  -211.43    1   -199.92   2.81  -22.01   0.00   0.00   1.01

REFERENCE COORDINATE DATA
-233.23 -210.82   1.23  -23.64

INFORMATION> AFTER MINIMIZE RANKING :EXPERIMENT ID=1
INFORMATION> LOCAL SCORE RANKING          3

COMPOUND NAME : conf1.mol2
FILE NAME     : Lig_es_1.mol2
      SCORE(/100)  dG      HIT      MTS      rotNum  ASA      ELE      HYD      VDW      SURFACE  RMSD
@  1   -2.70   -7.33  -71.60  -260.69    1   -232.43  -2.73  -33.19  -1.62   84.50   1.00
@  2   -2.66   -7.16  -69.10  -255.77    1   -233.92  -0.91  -29.32  -1.58   85.80   1.09
@  3   -2.65   -7.19  -69.57  -255.85    1   -227.70  -2.29  -33.02  -1.70   85.04   0.97

REFERENCE COORDINATE DATA
-233.23 -210.82   1.23  -23.64

INFORMATION> COMPOUND RANKING          1

      COMPOUND NAME          SCORE(/100)  dG      HIT      MTS
1  conf1.mol2          -2.70   -7.33  -71.60  -260.69

```

実行結果出力例（座標ファイル）

OUTPUT フェーズNAMECOによって指定されるドッキング後のリガンドの座標出力は以下のようになります。

(a) PDB 形式による出力 (COORDI= PDB のとき)

ファイル先頭には、全スコアと各項の結果、以下、リガンドの座標出力が行なわれますが、残基名は LGD に固定されています。この座標を、入力受容体座標と直接結合すれば、受容体 リガンド複合体の座標となります。

REMARK	conf1.mol2			1	of	3		
REMARK							SCORE =	-269.9679
REMARK							dG-SCORE =	-7.3270
REMARK							HIT-OPTIMIZED SCORE =	-71.6039
REMARK							MTS SCORE =	-260.6920
REMARK							ASA =	-232.4291
REMARK							ELE =	-2.7309
REMARK							HYD =	-33.1914
REMARK							VDW =	-1.6165
REMARK							RMSD =	1.0006
ATOM	1	H	LGD	1	14.483	37.096	34.845	
ATOM	2	C	LGD	1	13.739	37.114	35.656	
ATOM	3	H	LGD	1	14.027	36.387	36.431	
ATOM	4	C	LGD	1	13.766	38.538	36.271	
ATOM	5	O	LGD	1	12.974	39.529	36.222	
ATOM	6	O	LGD	1	14.768	38.853	36.954	
ATOM	7	C	LGD	1	12.454	36.628	35.048	
ATOM	8	C	LGD	1	11.423	35.866	35.673	
ATOM	9	H	LGD	1	11.561	35.490	36.683	
ATOM	10	C	LGD	1	10.199	35.595	34.972	
ATOM	11	H	LGD	1	9.431	35.008	35.449	
ATOM	12	C	LGD	1	9.988	36.111	33.689	
ATOM	13	H	LGD	1	9.043	35.956	33.187	
ATOM	14	C	LGD	1	11.060	36.870	33.105	
ATOM	15	C	LGD	1	12.313	37.009	33.714	
ATOM	16	H	LGD	1	13.156	37.375	33.155	
ATOM	17	N	LGD	1	10.893	37.607	31.942	
ATOM	18	O	LGD	1	11.430	38.662	31.685	
ATOM	19	O	LGD	1	10.134	37.042	31.184	

(b) mol2 形式による出力 (COORD1= MOL2 のとき)

ファイル先頭には、全スコアと各項の結果、以下、リガンドの座標出力が行なわれます。
リガンドの情報は、<MOLECULE>、<ATOM>、<BOND>セクションが記載されます。

```
# conf1.mol2          1    of    3
#                   SCORE = -269.9679
#                   dG-SCORE =  -7.3270
# HIT-OPTIMIZED SCORE = -71.6039
#                   MTS SCORE = -260.6920
#                   ASA = -232.4291
#                   ELE =  -2.7309
#                   HYD = -33.1914
#                   VDW =  -1.6165
#                   RMSD =   1.0006
@<TRIPOS>MOLECULE
conf1.mol2
  19  19  0  0  0
SMALL
NO_CHARGES

@<TRIPOS>ATOM
  1 H      14.4832  37.0962  34.8446 H      0 <1>      0.0018
  2 C      13.7387  37.1141  35.6561 C.3    0 <1>     -0.0913
  3 H      14.0267  36.3870  36.4311 H      0 <1>      0.0018
  4 C      13.7659  38.5378  36.2714 C.2    0 <1>      0.9634
  5 O      12.9736  39.5293  36.2221 O.co2   0 <1>     -0.8700
  6 O      14.7681  38.8530  36.9543 O.co2   0 <1>     -0.8700
  7 C      12.4544  36.6285  35.0477 C.ar    0 <1>      0.0677
  8 C      11.4228  35.8656  35.6730 C.ar    0 <1>     -0.1545
  9 H      11.5610  35.4899  36.6831 H      0 <1>      0.1531
 10 C      10.1993  35.5948  34.9716 C.ar    0 <1>     -0.1906
 11 H       9.4310  35.0081  35.4491 H      0 <1>      0.1429
 12 C       9.9882  36.1107  33.6891 C.ar    0 <1>     -0.1982
 13 H       9.0435  35.9563  33.1867 H      0 <1>      0.1532
```

(つづく)

(つづき)

14	C	11.0604	36.8703	33.1052	C.ar	0 <1>	0.0189
15	C	12.3134	37.0092	33.7138	C.ar	0 <1>	-0.1494
16	H	13.1560	37.3747	33.1546	H	0 <1>	0.2269
17	N	10.8933	37.6067	31.9419	N.pl3	0 <1>	0.7824
18	O	11.4300	38.6616	31.6850	O.co2	0 <1>	-0.4940
19	O	10.1339	37.0415	31.1844	O.co2	0 <1>	-0.4940
@<TRIPOS>BOND							
1	1	2	1				
2	2	3	1				
3	2	4	1				
4	2	7	1				
5	4	5	1				
6	4	6	2				
7	7	8	ar				
8	7	15	ar				
9	8	9	1				
10	8	10	ar				
11	10	11	1				
12	10	12	ar				
13	12	13	1				
14	12	14	ar				
15	14	15	ar				
16	14	17	1				
17	15	16	1				
18	17	18	1				
19	17	19	2				
#MOLECULE END							

(2) グリッドポテンシャルの生成と、ドッキングを別々に行なう方法

標的受容体に対し、多数の化合物をドッキングする場合、標的受容体のグリッドポテンシャルを生成してファイルに保存しておき、化合物のドッキング計算をあらかじめ生成しておいたグリッドポテンシャルを用いて行なうのが効率的です。

【グリッドポテンシャルのみを発生し、ファイル保存する】

INPUT フェーズ、GRID フェーズの指定は(1)での指定と同様で、CONF/DOCK/MIN/OUTPUT フェーズの指定は不要です。出力するグリッドポテンシャルのファイル名(NAMEGR)と、ファイル書式(OUTGRI)を指定します。

sievgene の実行 :

入力ファイルを grid_make.inp とします。

```
% sievgene < grid_make.inp
```

【注意】グリッドポテンシャルのみを発生させる (CONF/DOCK/MIN/OUTPUT フェーズ指定をしない) 場合、入力ファイルへの「EXE> SIEV」の記述は不要です。
この場合に「EXE> SIEV」を指定すると、グリッドポテンシャル出力の後、以下のようなエラーメッセージを出力します。

```
INFORMATION> WRITE GRID POTENTIAL BINARY FORM:  
FILE NAME :grid_file
```

```
-----  
Error!! : FILE NOT FOUND: input ligand reference mol2 file: FILE NAME=""
```

入力ファイル (grid_make.inp)

```
PHASE> INPUT
  TOPOLO = FORM
  NAMETO = Pro.tpl      ; 蛋白質側のトポロジー
  COORDI = PDB
  NAMECO = Pro_md.pdb  ; 蛋白質側の座標
  POINTC = PDB
  NAMEPO = point.pdb   ; ポケット領域を示すプローブ点

  SETTAR = NORE
  DAMPPA = 1.0d0
  QUIT
;
; gird generation and Hash table generation
;
PHASE> GRID
  GRIDPOtential = NORE      ; Grid ポテンシャル入力
  NAMEGRid      = grid.file ; Grid ポテンシャルファイル名
  OUTGRIdpotential = BINA   ; Grid ポテンシャル出力
  PROBDist = 6.5 ; プローブ点から PROBDIST 内のレセプター原子を対象にする。
  MARGIN = 6.5 ; Grid 生成時の、探索ポケット周辺にとるマージン
  ITERAT = 3 ; Grid ポテンシャルのスムージングの繰り返し回数
  RADVDW = 0.6 ; vDW ポテンシャルの内外領域の界面の設定
  RADELE = 0.6 ; クーロンポテンシャルの内外領域の界面の設定
  RADMESh = 1.4 ; mesh 点生成のプローブ半径

  DAMPVW = 0.99d0

  USEPBG = NO ; no use of PB

  PBGRID = BINA ; for PB
  OUTPBG = NOWR ; for PB
  NAMEPB = pb.file ; for PB
  NMESHX = 100 ; for PB
  NMESHY = 100 ; for PB
  NMESHZ = 100 ; for PB
  MARGPB = 5.0d0 ; for PB
  KAPPAV = 0.0d0 ; for PB
  DIESOL = 78.5d0 ; for PB
  DIEPRO = 4.0d0 ; for PB
  DIEINT = 78.5d0 ; for PB
  DIEVAC = 1.0d0 ; for PB
  ACCELE = 1.6d0 ; for PB
  CPUTIM = 10800 ; for PB
  QUIT
```

【既に生成したグリッドポテンシャルを利用して、ドッキングのみを行なう】

基本的な指定は(1)と変わりません。GRID フェーズにおいて、入力グリッドファイル名(NAMEGR)、グリッドファイル書式指定(GRIDPO)をします。グリッドポテンシャルを出力しないので、OUTGRI=NOWR を指定します。

sievgene の実行 :

入力ファイルを restart_run.inp とします。

```
% sievgene < restart_run.inp
```

出力ファイルなどは、(1)と変わりません。

入力ファイル (restart_run.inp)

```

PHASE> INPUT
LIGAND = MOL2
NAMELI = Lig_es_1.mol2 ; リガンドの座標位置ファイル
REFERE = MOL2
NAMERE = lig_ref.mol2 ; リガンドの参照位置ファイル
TOPOLO = FORM
NAMETO = Pro.tpl ; 蛋白側のトポロジー
COORDI = PDB
NAMECO = Pro_md.pdb ; 蛋白側の座標
POINTC = PDB
NAMEPO = point.pdb ; ポケット領域を示すプローブ点

SETTAR = NORE
DAMPPA = 1.0d0
QUIT
;
; gird generation and Hash table generation
;
PHASE> GRID
GRIDPOtential = BINA ; Grid ポテンシャル入力
NAMEGRid = grid_file ; Grid ポテンシャルファイル名
OUTGRIdpotential = NOWR ; Grid ポテンシャル出力
PROBDist = 6.5 ; プローブ点から PROBDIST 内のレセプター原子を対象にする。
MARGIN = 6.5 ; Grid 生成時の、探索ポケット周辺にとるマージン
ITERAT = 3 ; Grid ポテンシャルのスムージングの繰り返し回数
RADVDW = 0.6 ; vDW ポテンシャルの内外領域の界面の設定
RADELE = 0.6 ; クーロンポテンシャルの内外領域の界面の設定
RADMESh = 1.4 ; mesh 点生成のプローブ半径

DAMPVW = 0.99d0

USEPBG = NO ; no use of PB

PBGRID = BINA ; for PB
OUTPBG = NOWR ; for PB
NAMEPB = #NAMEPB# ; for PB
NMESHX = 100 ; for PB
NMESHY = 100 ; for PB
NMESHZ = 100 ; for PB
MARGPB = 5.0d0 ; for PB
KAPPAV = 0.0d0 ; for PB
DIESOL = 78.5d0 ; for PB
DIEPRO = 4.0d0 ; for PB
DIEINT = 78.5d0 ; for PB
DIEVAC = 1.0d0 ; for PB
ACCELE = 1.6d0 ; for PB
CPUTIM = 10800 ; for PB
QUIT
;

```

(つづく)

(つづき)

```

PHASE> CONF
  CONFLimit      = 100000 ; 配座生成の試行最大数
  CONFORMernumber = 100   ; 生成する配座数(リガンドを変形)
  SORTATom       = YES    ; 出力ファイルの原子順を変える
  DAMPING        = 0.7    ; 配座候補の原子間距離のダンピングファクター
  PHASETorsion   = 3      ; conformer 生成時の torsion 回転候補の個数
  ROTTER         = YES
  QUIT

PHASE> DOCK
  METHOD          = FLEX
  GENERAtion     = 1      ; 絞込みの回数
  NUMCONFomer    = 1000 ; リガンドの接合面三角形数
  MATCHING       = 4      ; 3 角形の原子タイプの適合度
  LOWMIN         = 2.5    ; Hash 表と比較する原子間距離の下限の最小値
  LOWMAX         = 3.5    ; Hash 表と比較する原子間距離の下限の最大値
  UPRMIN         = 5.0    ; Hash 表と比較する原子間距離の上限の最小値
  UPRMAX         = 12.0   ; Hash 表と比較する原子間距離の上限の最大値

  WETVDW= 1.0 ; 総合スコア計算時の各項の重み
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  RADIUS= 6.0 ; リガンド分子ドッキングの後 標的原子(座標)から RADIUS内
               ; に存在するリガンドの原子の数を数える。

  EVALHB = NO
  WETANH = 1.0d0
  ROTLOH = NO
  ROTPSC = NO
  MOVNUM = 10
  QUIT

EXE> MIN
  METHOD= STEEP      CPUTIM = 360000.0
  UPRATE= 1.0       DOWNRATE= 0.3
  LOOPLI= 100       UPDATE = 20
  MONITO= 100       CONVGR = 0.1D0
  CUTMET= RESA      CUTLEN = 22.0D0
  DIEFUN= DIST      DIEVAL = 4.0D0
  LOGFOR= SHOR
  BESTFI= YES

  QUIT

PHASE> OUTPUT
  COORDInate      = MOL2      ; リガンド座標のファイル出力形式
  NAMECOordinate = ex.cor     ; 各実験でのリガンド座標ファイル名
  CANDIDatenumbe r= 10       ; 各実験でのリガンド座標出力する上位スコア数
  SCORENumber     = 10       ; score ファイルに出力する上位スコア数
  NAMESCore       = ex.score  ; score ファイル名
  QUIT

EXE> SIEV

```

2.2 オプションの組み合わせ (sievgene_for_dockingpose の場合)

sievgene では、入力ファイルで指定するオプションの組み合わせが実行時間やドッキングの精度に大きく影響します。

組み合わせ方法は多岐にわたりますが、ここでは代表的な 3 例について説明します。

- (1) fast version : 実行時間を最優先する計算方法
- (2) moderate version : 実行時間、ドッキング精度とも重視する計算方法
- (3) precise version : ドッキング精度を最優先する計算方法

これらの 3 例では、下表のオプションの組み合わせにより、目的に合った計算を行っています。

フェーズ	オプション	指定方法		
		(1) fast version	(2) moderate version	(3) precise version
CONF フェーズ	ATMMDL	UNIT	ALL	ALL
DOCK フェーズ	DOCKSP	FAST	NORM	NORM
	CANDID	10	100	100
	MATCHI	3	2	0
	LOWMIN	2.5	2.5	1.0
	LOWMAX	3.5	3.5	1.2
	UPMIN	5.0	5.0	8.0

(1) fast version

この方法は、他の方法と比べてドッキング精度はやや落ちるものの、極めて短い時間でドッキング計算を可能とするものです。

実行時間を優先するため、原子モデルは united atom model を使用し (ATMMDL= UNIT)、蛋白とリガンドの重ね合わせの際に、重ね合わせの対象外となる原子を計算対象から除外する計算方法を使用します (DOCKSP= FAST)。

また、ローカルサーチの対象数を 10 とすることで (CANDID= 10)、ローカルサーチの実行時間も短縮します。

更に、結合面の原子タイプ適合度 3 とし (MATCH= 3)、接合面の辺の長さの範囲をデフォルトよりやや狭めることで (LOWMIN= 2.5, LOWMAX= 3.5, UPMIN= 5.0)、実行時間を短縮します。

入力ファイル

```

PHASE> INPUT
  LIGAND = MOL2
  NAMELI = Lig_es_1.mol2 ; リガンドの座標位置ファイル
  REFERE = MOL2
  NAMERE = lig_ref.mol2 ; リガンドの参照位置ファイル
  TOPOLO = FORM
  NAMETO = Pro.tpl ; 蛋白側のトポロジー
  COORDI = PDB
  NAMECO = Pro_md.pdb ; 蛋白側の座標
  POINTC = PDB
  NAMEPO = point.pdb ; ポケット領域を示すプローブ点

  ASAMET = PAIR ; ASA 計算方法
  SETTAR = NORE
  DAMPPA = 1.0d0
  QUIT
;
;
; gird generation and Hash table generation
;
;
PHASE> GRID
  GRIDPotential = BINA ; Grid ポテンシャル入力
  NAMEGRid = grid_file ; Grid ポテンシャルファイル名
  OUTGRIdpotential = NOWR ; Grid ポテンシャル出力
  PROBDist = 6.5 ; プローブ点から PROBDIST 内のレセプター原子を対象にする。
  MARGIN = 6.5 ; Grid 生成時の、探索ポケット周辺にとるマージン
  ITERAT = 3 ; Grid ポテンシャルのスムージングの繰り返し回数
  RADVDW = 0.6 ; vDW ポテンシャルの内外領域の界面の設定
  RADELE = 0.6 ; クーロンポテンシャルの内外領域の界面の設定
  RADMESH = 1.4 ; mesh 点生成のプローブ半径

  DAMPVW = 0.99d0

  USEPBG = NO ; no use of PB

  PBGRID = BINA ; for PB
  OUTPBG = NOWR ; for PB
  NAMEPB = pb.file ; for PB
  NMESHX = 100 ; for PB
  NMESHY = 100 ; for PB
  NMESHZ = 100 ; for PB
  MARGPB = 5.0d0 ; for PB
  KAPPAV = 0.0d0 ; for PB
  DIESOL = 78.5d0 ; for PB
  DIEPRO = 4.0d0 ; for PB
  DIEINT = 78.5d0 ; for PB
  DIEVAC = 1.0d0 ; for PB
  ACCELE = 1.6d0 ; for PB
  CPUTIM = 10800 ; for PB
  QUIT

```

(つづく)

(つづき)

```

PHASE> CONF
  ATMMDL           = UNIT           ; 原子モデル指定
  CONFLimit        = 100000         ; 配座生成の試行最大数
  CONFORMernumber  = 100            ; 生成する配座数(リガンドを变形)
  SORTATom        = YES            ; 出力ファイルの原子順を変える
  DAMPING          = 0.7           ; 配座候補の原子間距離のダンピングファクター
  PHASETorsion     = 3              ; conformer 生成時の torsion 回転候補の個数
  ROTTER          = YES
  QUIT

PHASE> DOCK
  METHOD           = FLEX ; ドッキング方法 ( rigid | flexible )
  DOCKSP         = FAST ; ドッキング方法 ( normal | fast )
  CANDID         = 10 ; ローカルサーチ対象数
  GENERAtion     = 1 ; 絞込みの回数
  NUMCONFomer    = 1000 ; リガンドの接合面三角形数
  MATCHING       = 3 ; 3 角形の原子タイプの適合度
  LOWMIN         = 2.5 ; Hash 表と比較する原子間距離の下限の最小値
  LOWMAX         = 3.5 ; Hash 表と比較する原子間距離の下限の最大値
  UPRMIN         = 5.0 ; Hash 表と比較する原子間距離の上限の最小値
  UPRMAX         = 12.0 ; Hash 表と比較する原子間距離の上限の最大値

  WETVDW= 1.0 ; 総合スコア計算時の各項の重み
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  WETANH= 1.0
  RADIUS= 6.0 ; リガンド分子ドッキングの後 標的原子(座標) から RADIUS 内
               ; に存在するリガンドの原子の数を数える。

  EVALHB = NO
  ROTLOH = NO
  ROTPSC = NO
  MOVNUM = 10
  QUIT

EXE> MIN
  METHOD=  STEEP           CPUTIM = 360000.0
  UPRATE= 1.0             DOWNRATE= 0.3
  LOOPLI= 100             UPDATE = 20
  MONITO= 100             CONVGR = 0.1D0
  CUTMET=  RESA           CUTLEN = 22.0D0
  DIEFUN=  DIST           DIEVAL = 4.0D0
  LOGFOR=  SHOR
  BESTFI=  YES

  QUIT

PHASE> OUTPUT
  COORDInate      = MOL2           ; リガンド座標のファイル出力形式
  NAMECOordinate  = ex.cor         ; 各実験でのリガンド座標ファイル名
  CANDIDatenumbe r= 3              ; 各実験でのリガンド座標を出力する上位スコア数
  SCORENumber     = 3              ; score ファイルに出力する上位スコア数
  NAMESCore       = ex.score       ; score ファイル名
  QUIT

EXE> SIEV

```

(2) moderate version

この方法は、ドッキング精度、実行時間の双方を重視する計算方法です。

ドッキング精度を保つため、原子モデルは all atom model を使用し (ATMMDL= ALL)、蛋白とリガンドの重ね合わせの際に、重ね合わせの対象外となる原子を計算対象から除外しない計算方法を使用します (DOCKSP= NORM)。

また、ローカルサーチの対象数を 100 とすることでも (CANDID= 100) ドッキング精度を維持します。

一方、結合面の原子タイプ適合度 2 とし (MATCH= 2)、接合面の辺の長さの範囲をデフォルトよりやや狭めることで (LOWMIN= 2.5, LOWMAX= 3.5, UPMIN= 5.0) 実行時間を短縮します。

入力ファイル

```

PHASE> INPUT
LIGAND = MOL2
NAMELI = Lig_es_1.mol2 ; リガンドの座標位置ファイル
REFERE = MOL2
NAMERE = lig_ref.mol2 ; リガンドの参照位置ファイル
TOPOLO = FORM
NAMETO = Pro.tpl ; 蛋白側のトポロジー
COORDI = PDB
NAMECO = Pro_md.pdb ; 蛋白側の座標
POINTC = PDB
NAMEPO = point.pdb ; ポケット領域を示すプローブ点

ASAMET = PAIR ; ASA 計算方法
SETTAR = NORE
DAMPPA = 1.0d0
QUIT
;
; gird generation and Hash table generation
;
PHASE> GRID
GRIDPOtential = BINA ; Grid ポテンシャル入力
NAMEGRid = grid_file ; Grid ポテンシャルファイル名
OUTGRIdpotential = NOWR ; Grid ポテンシャル出力
PROBDIst = 6.5 ; プローブ点から PROBDIST 内のレセプター原子を対象にする。
MARGIN = 6.5 ; Grid 生成時の、探索ポケット周辺にとるマージン
ITERAT = 3 ; Grid ポテンシャルのスムージングの繰り返し回数
RADVDW = 0.6 ; vDW ポテンシャルの内外領域の界面の設定
RADELE = 0.6 ; クーロンポテンシャルの内外領域の界面の設定
RADMESH = 1.4 ; mesh 点生成のプローブ半径

DAMPVW = 0.99d0

USEPBG = NO ; no use of PB

```

(つづく)

(つづき)

```

PHASE> CONF
  ATMDL      = ALL      ; 原子モデル指定
  CONFLimit  = 100000   ; 配座生成の試行最大数
  CONFORMernumber = 100 ; 生成する配座数(リガンドを变形)
  SORTATom   = YES     ; 出力ファイルの原子順を変える
  DAMPING    = 0.7     ; 配座候補の原子間距離のダンピングファクター
  PHASETorsion = 3     ; conformer 生成時の torsion 回転候補の個数
  ROTTER     = YES
  QUIT

PHASE> DOCK
  METHOD      = FLEX ; ドッキング方法 ( rigid | flexible )
  DOCKSP     = NORM ; ドッキング方法 ( normal | fast )
  CANDID     = 100 ; ローカルサーチ対象数
  GENERAtion = 1   ; 絞込みの回数
  NUMCONFomer = 1000 ; リガンドの接合面三角形数
  MATCHING   = 2   ; 3 角形の原子タイプの適合度
  LOWMIN     = 2.5 ; Hash 表と比較する原子間距離の下限の最小値
  LOWMAX     = 3.5 ; Hash 表と比較する原子間距離の下限の最大値
  UPRMIN     = 5.0 ; Hash 表と比較する原子間距離の上限の最小値
  UPRMAX     = 12.0 ; Hash 表と比較する原子間距離の上限の最大値

  WETVDW= 1.0 ; 総合スコア計算時の各項の重み
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  WETANH= 1.0
  RADIUS= 6.0 ; リガンド分子ドッキングの後 標的原子(座標) から RADIUS内
               ; に存在するリガンドの原子の数を数える。

  EVALHB = NO
  ROTLOH = NO
  ROTPSC = NO
  MOVNUM = 10
  QUIT

EXE> MIN
  METHOD=  STEEP      CPUTIM = 360000.0
  UPRATE=  1.0      DOWNRATE= 0.3
  LOOPLI= 100       UPDATE  = 20
  MONITO= 100       CONVGR  = 0.1D0
  CUTMET=  RESA     CUTLEN  = 22.0D0
  DIEFUN=  DIST     DIEVAL  = 4.0D0
  LOGFOR=  SHOR
  BESTFI=  YES

  QUIT

PHASE> OUTPUT
  COORDInate = MOL2      ; リガンド座標のファイル出力形式
  NAMECOordinate = ex.cor ; 各実験でのリガンド座標ファイル名
  CANDIDatenuMber= 3     ; 各実験でのリガンド座標を出力する上位スコア数
  SCORENumber = 3       ; score ファイルに出力する上位スコア数
  NAMESCore   = ex.score ; score ファイル名
  QUIT

EXE> SIEV

```

(3) precise version

この方法は、他の方法と比べて長い実行時間を要するものの、極めて精度の高いドッキング計算を可能とするものです。

ドッキング精度を保つため、原子モデルは all atom model を使用し (ATMMDL= ALL)、蛋白とリガンドの重ね合わせの際に、重ね合わせの対象外となる原子を計算対象から除外しない計算方法を使用します (DOCKSP= NORM)。

また、ローカルサーチの対象数を 100 とすることでも (CANDID= 100) ドッキング精度を維持します。

さらに、結合面の原子タイプ適合度 0 とし (MATCH= 0)、接合面の辺の長さの範囲をデフォルトより広くすることで (LOWMIN= 1.0, LOWMAX= 1.2, UPMIN= 8.0) ドッキング精度の向上を図ります。

入力ファイル

```

PHASE> INPUT
  LIGAND = MOL2
  NAMELI = Lig_es_1.mol2 ; リガンドの座標位置ファイル
  REFERE = MOL2
  NAMERE = lig_ref.mol2 ; リガンドの参照位置ファイル
  TOPOLO = FORM
  NAMETO = Pro.tpl ; 蛋白側のトポロジー
  COORDI = PDB
  NAMECO = Pro_md.pdb ; 蛋白側の座標
  POINTC = PDB
  NAMEPO = point.pdb ; ポケット領域を示すプローブ点

  ASAMET = PAIR ; ASA 計算方法
  SETTAR = NORE
  DAMPPA = 1.0d0
  QUIT
;
; gird generation and Hash table generation
;
PHASE> GRID
  GRIDPOTential = BINA ; Grid ポテンシャル入力
  NAMEGRid = grid_file ; Grid ポテンシャルファイル名
  OUTGRIdpotential = NOWR ; Grid ポテンシャル出力
  PROBDist = 6.5 ; プローブ点から PROBDIST 内のレセプター原子を対象にする。
  MARGIN = 6.5 ; Grid 生成時の、探索ポケット周辺にとるマージン
  ITERAT = 3 ; Grid ポテンシャルのスムージングの繰り返し回数
  RADVDW = 0.6 ; vDW ポテンシャルの内外領域の界面の設定
  RADELE = 0.6 ; クーロンポテンシャルの内外領域の界面の設定
  RADMESH = 1.4 ; mesh 点生成のプローブ半径
  DAMPVW = 0.99d0

  USEPBG = NO ; no use of PB

```

(つづく)

(つづき)

```

PHASE> CONF
  ATMDL      = ALL      ; 原子モデル指定
  CONFLimit  = 100000   ; 配座生成の試行最大数
  CONFORMernumber = 100 ; 生成する配座数(リガンドを变形)
  SORTATom   = YES     ; 出力ファイルの原子順を変える
  DAMPING    = 0.7     ; 配座候補の原子間距離のダンピングファクター
  PHASETorsion = 3     ; conformer 生成時の torsion 回転候補の個数
  ROTTER     = YES
  QUIT

PHASE> DOCK
  METHOD      = FLEX ; ドッキング方法 ( rigid | flexible )
  DOCKSP     = NORM ; ドッキング方法 ( normal | fast )
  CANDID     = 100 ; ローカルサーチ対象数
  GENERAtion = 1   ; 絞込みの回数
  NUMCONFomer = 1000 ; リガンドの接合面三角形数
  MATCHING   = 0   ; 3 角形の原子タイプの適合度
  LOWMIN     = 1.0 ; Hash 表と比較する原子間距離の下限の最小値
  LOWMAX     = 1.2 ; Hash 表と比較する原子間距離の下限の最大値
  UPRMIN     = 8.0 ; Hash 表と比較する原子間距離の上限の最小値
  UPRMAX     = 12.0 ; Hash 表と比較する原子間距離の上限の最大値

  WETVDW= 1.0 ; 総合スコア計算時の各項の重み
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  WETANH= 1.0
  RADIUS= 6.0 ; リガンド分子ドッキングの後 標的原子(座標) から RADIUS 内
               ; に存在するリガンドの原子の数を数える。

  EVALHB = NO
  ROTLOH = NO
  ROTPSC = NO
  MOVNUM = 10
  QUIT

EXE> MIN
  METHOD=  STEEP      CPUTIM = 360000.0
  UPRATE=  1.0      DOWNRATE= 0.3
  LOOPLI=  100      UPDATE  = 20
  MONITO=  100      CONVGR  = 0.1D0
  CUTMET=  RESA     CUTLEN  = 22.0D0
  DIEFUN=  DIST     DIEVAL  = 4.0D0
  LOGFOR=  SHOR
  BESTFI=  YES

  QUIT

PHASE> OUTPUT
  COORDInate = MOL2      ; リガンド座標のファイル出力形式
  NAMECOordinate = ex.cor ; 各実験でのリガンド座標ファイル名
  CANDIDatenuMber= 3     ; 各実験でのリガンド座標を出力する上位スコア数
  SCORENumber = 3       ; score ファイルに出力する上位スコア数
  NAMESCore   = ex.score ; score ファイル名
  QUIT

EXE> SIEV

```

A 入出力ファイル

A.1 ドッキングエンジンの入出力ファイル

ドッキングエンジンは、以下の目的でファイルの入出力を行います。

- (1)シミュレーション条件の指定
- (2)シミュレーション状態の保存
- (3)シミュレーション結果の出力

これらのファイルを総称して入出力ファイルと呼びます。

A.1.1 フェーズの説明

ドッキングエンジンは、以下のフェーズでファイルの入出力を行います。

(1) INPUT フェーズ

系のトポロジー、座標、シミュレーション条件などのファイルの入力。

(2) OUTPUT フェーズ

シミュレーション後のスコア情報、リガンド Conformer の座標のファイル出力。

(3) GRID フェーズ

蛋白質ポケット部分のグリッドポテンシャルの生成。

(4) CONF フェーズ

リガンドの Conformer の生成。

(5) DOCK フェーズ

系のドッキングシミュレーション。

(6) MIN フェーズ

系のポテンシャルエネルギーの最小化。

A.2 入力ファイル

ドッキングエンジンの入力ファイルを以下に示します。

項番	ファイル名称	対象	用途
#1	制御ファイル	-	ドッキングエンジンの制御
#2	トポロジーファイル	全フェーズ	蛋白質のトポロジー指定
#3	座標ファイル	全フェーズ	蛋白質原子、ポケット点の座標指定
#4	リスタートファイル	全フェーズ	リスタート情報の指定
#5	探索対象指定ファイル	全フェーズ	探索対象にする蛋白質原子の指定
#6	Gridポテンシャル ファイル	GRID	Gridポテンシャルの指定
#7	PB-Gridポテンシャル ファイル	GRID	Poisson-Boltzmann Gridポテンシャルの 指定

A.2.1 制御ファイル

対象フェーズ：全フェーズ

用途：構造探索エンジンの実行フェーズ制御とフェーズに指定するパラメータを指定します。

特記事項：

実数の指数を示す文字は "D" でなければなりません。

実数指数の指定例)

```
CPUTIM = 60.0D0
```

書式：

```
[実行フェーズ指定行 [フェーズ内パラメータ指定行]... パラメータ終了行 ]...
```

実行フェーズ指定行：実行するフェーズを以下の文字列で指定します。

```
INPUT フェーズ = " PHASE> INPUT "
OUTPUT フェーズ = " PHASE> OUTPUT "
GRID フェーズ = " PHASE> CONF "
CONF フェーズ = " PHASE> OUTPUT "
MINimize フェーズ = " EXE> MIN "
実行終了行 = " EXE> SIEV "
```

パラメータ終了行：フェーズ内パラメータ指定行の終了を以下の形式で指定します。

```
" QUIT "
```

フェーズ内パラメータ指定：各フェーズのパラメータを以下の形式で指定します。

```
キーワード "=" 値
```

キーワードは英字 6 文字で構成され、値はキーワードに対応し、選択型、実数型、整数型、文字型の 4 種類があります。

キーワード指定例)

```
UNITAN = 30           ; 整数型パラメータ
NAMEAN = aa.ana      ; 文字型パラメータ
BINCL0 = NO         ; 選択型パラメータ
CPUTIM = 60.000     ; 実数型パラメータ
```

以下、各フェーズのキーワードと値を示す。英字の大文字が有効となります。

また、「内容」列の括弧内は、「値」列の形式に従い、次の項目を記述します。

- ・ 選択型 : 大文字は指定文字、アンダーライン付はデフォルト値
- ・ 整数型、実数型、文字型 : デフォルト値

A.2.1.1 INPUT フェーズ

項番	項目	キーワード	値	内容
#1	蛋白 トポロジー	<u>TOPOLO</u>	選択型	ファイルの読み込みと形式 (<u>NORE</u> FORM BINA)
#2		<u>NAMETO</u>	文字型	ファイル名("")
#3	蛋白原子 座標	<u>COORDI</u>	選択型	ファイルの読み込みと形式 (<u>PDB</u> BINA)
#4		<u>NAMECO</u>	文字型	ファイル名("")
#5	リガンド 座標	<u>LIGAND</u>	選択型	ファイルの読み込みと形式 (<u>PDBX</u> <u>MOL2</u>)
#6		<u>NAMELI</u>	文字型	ファイル名("")
#7	リガンド 参照座標	<u>REFERE</u>	選択型	ファイルの読み込みと形式 (<u>NORE</u> <u>MOL2</u>)
#8		<u>NAMERE</u>	文字型	ファイル名("")
#9	蛋白 ポケット点 座標	<u>POINTC</u>	選択型	ファイルの読み込み (<u>PDB</u> BINA)
#10		<u>NAMEPO</u>	文字型	ファイル名("")
#11	ASA の指定	<u>ASAMET</u>	選択型	ASA の計算方法 (<u>PAIR</u> RICH)
#12	探索対象 指定	<u>SETTAR</u>	選択型	ファイルの読み込みと形式 (<u>NORE</u> READ)
#13		<u>NAMETA</u>	文字型	ファイル名("")
#14	蛋白質原子情報	<u>DAMPPA</u>	実数型	蛋白質の原子半径の係数(1.0)
#15	ログ出力	<u>LOGFOR</u>	選択型	ログファイルの出力形式 (<u>SHOR</u> <u>MEDI</u> <u>DETA</u>)

A.2.1.2 GRID フェーズ

項番	項目	キーワード	値	内容
#1	Grid-Potential の入出力	<u>GRIDPO</u>	選択型	ファイルの読み込みと形式 (<u>NORE</u> ASCII BINA)
#2		<u>OUTGRI</u>	選択型	ファイルの書き込みと形式 (<u>NOWR</u> ASCII BINA)
#3		<u>NAMEGR</u>	文字型	ファイル名("")
#4	レセプター 登録範囲	<u>PROBDI</u>	実数型	ポケット点とレセプター原子の 距離上限(6.5)
#5	Grid-Potential の生成	<u>RADVW</u>	実数型	vdW の境界距離補正(0.6)
#6		<u>RADELE</u>	実数型	静電の境界距離補正 (0.6)
#7		<u>DANPVW</u>	実数型	vdW 半径の係数 (1.0)
#8	メッシュ生成	<u>RADMES</u>	実数型	mesh 点生成のプロープ半径(1.6)
#9	Grid-Potential の範囲補正	<u>MERGIN</u>	実数型	GridPotential の範囲のマージン (0.0)
#10	スムージング 回数	<u>ITERAT</u>	整数型	スムージング繰り返し回数(3)
#11	PB-Grid ポテンシャル	<u>USEPBG</u>	選択型	Poisson Boltzmann 方程式の 使用 (<u>NO</u> YES)
#12		<u>PBGRID</u>	選択型	ファイルの読み込みと形式 (<u>NORE</u> ASCII BINA)
#13		<u>OUTPBG</u>	選択型	ファイルの書き込みと形式 (<u>NOWR</u> ASCII BINA)
#14		<u>NAMEPB</u>	文字型	ファイル名("")
#15		<u>NMESHX</u>	整数型	メッシュ数(x方向) (100)
#16		<u>NMESHY</u>	整数型	メッシュ数(y方向) (100)
#17		<u>NMESHZ</u>	整数型	メッシュ数(z方向) (100)
#18		<u>MARGPB</u>	実数型	PB-Gridポテンシャルの マージン (5.0d0)
#19		<u>KAPPAV</u>	実数型	Debye-Huckel の遮蔽定数 (0.0)
#20		<u>CPULIM</u>	実数型	CPU time の上限値 (10800)
#21	静電場計算終了 条件の閾値	<u>SORRE1</u>	実数型	終了条件の相対閾値(SOR)
#22		<u>SORRE2</u>	実数型	終了条件の相対閾値(SOR)
#23		<u>SRFAB1</u>	実数型	終了条件の絶対閾値(SURF)

#24		<u>SRFAB2</u>	実数型	終了条件の絶対閾値(SURF)
#25		<u>SRFRE1</u>	実数型	終了条件の相対閾値(SURF)
#26		<u>SRFRE2</u>	実数型	終了条件の相対閾値(SURF)

A.2.1.3 CONF フェーズ

項番	項目	キーワード	値	内容
#1	conformer 生成	<u>CONFLI</u>	整数型	conformer 生成の最大試行回数 (10000000)
#2		<u>CONFOR</u>	整数型	生成する配座数 (100)
#3	原子データ のソート	<u>SORTAT</u>	選択型	原子データのソート実行 (<u>NO</u> YES)
#4	原子間 vdW 距離下限の 係数	<u>DAMPIN</u>	実数型	配座の原子間 vdW 距離の ダンピングファクター (0.7)
#5	torsion 回転数	<u>PHASET</u>	整数型	torsion 回転の候補数 (6)
#6	原子モデル	<u>ATMMDL</u>	選択型	原子モデルの指定 (<u>ALL</u> UNIT)
#7	C-C 構造の 回転	<u>ROTC-C</u>	選択型	C-C の回転 (<u>NO</u> 、YES)

A.2.1.4 DOCK フェーズ

項番	項目	キーワード	値	内容
#1	ドッキング方法	<u>METHOD</u>	選択型	ドッキング方法 (<u>FLEX</u> <u>RIGI</u>)
#2		<u>PROSUR</u>	選択型	蛋白側接合点発生方法 (<u>ELEC</u> <u>HYDR</u>)
#3		<u>GENERA</u>	整数型	絞込みの回数(5)
#4		<u>NUMCON</u>	実数型	表示する上位スコア数(30)
#5		<u>MATCHI</u>	整数型	結合面の原子タイプ適合度(5)
#6		<u>LOWMIN</u>	実数型	結合面の辺の下限最小値(2.5)
#7		<u>LOWMAX</u>	実数型	結合面の辺の下限最大値(4.0)
#8		<u>UPRMIN</u>	実数型	結合面の辺の上限最小値(7.5)
#9		<u>UPRMAX</u>	実数型	結合面の辺の上限最大値(10.0)
#10		<u>INTERP</u>	選択型	Gridポテンシャルの補間方法 (<u>LAGR</u> <u>BSPL</u>)
#11		<u>EVALHB</u>	選択型	蛋白質とリガンドの、異方性を考慮した水素結合の評価 (<u>YES</u> <u>NO</u>)
#12		<u>ROTLOH</u>	選択型	リガンド-OH基の回転 (<u>YES</u> <u>NO</u>)
#13		<u>ROTPSC</u>	選択型	水素結合可能な蛋白質側鎖の回転 (<u>YES</u> <u>NO</u>)
#14		<u>MOVNUM</u>	整数型	座標をずらす回数(10)
#15		<u>CANDID</u>	整数型	ローカルサーチ対象数(30)
#16		<u>DOCKSP</u>	整数型	ドッキング対象原子の切り替え (<u>NORM</u> <u>FAST</u>)
#17		<u>OPTIME</u>	整数型	ドッキング計算の有無の指定 (<u>OPT</u> <u>ENE</u>)
#18		<u>CPUTIME</u>	実数型	1分子あたりのドッキング計算時間の上限値(秒)(30.0)
#19	スコア計算	<u>WETVDW</u>	実数型	スコア計算時のvdWの係数(1.0)
#20		<u>WETASA</u>	実数型	スコア計算時のASAの係数(1.0)
#21		<u>WETELE</u>	実数型	スコア計算時の静電の係数(1.0)
#22		<u>WETHYD</u>	実数型	スコア計算時の水素結合の係数(1.0)

#23		<u>WETANH</u>	実数型	蛋白質とリガンドの、異方性を考慮した水素結合の係数(1.0)
#24		<u>RADIUS</u>	実数型	結合原子のカウント距離上限(6.0)
#25	ポケット 情報	<u>POCKCX</u>	実数型	ポケット中心 X 座標(999.9)
#26		<u>POCKCY</u>	実数型	ポケット中心 Y 座標(999.9)
#27		<u>POCKCZ</u>	実数型	ポケット中心 Z 座標(999.9)
#28		<u>POCKET</u>	整数型	ポケット中心原子 ID (0)
#29	グローバル サーチ種別	GLOMET	選択型	乱数/剛体での疑似 MINIMIZE (RANDom、RIGId)
#30		GLOMAX	整数型	疑似 MINIMIZE の繰り返し上限(10)
#31		TRASTP	実数型	1ステップでの平行移動量(1.0)
#32		ROTSTP	実数型	1ステップでの回転角(1.0°)

A.2.1.5 MIN フェーズ

ユーザマニュアル「cosgnene」の章の「EXE> MIN グループ」の本項を参照ください。

A.2.1.6 OUTPUT フェーズ

項番	項目	キーワード	値	内容
#1	上位 conformer の 座標ファイル出力	<u>COORDI</u>	選択型	ファイル出力と形式 (<u>NOWR</u> MOL2 PDBX PDB)
#2		<u>CANDID</u>	整数型	上位スコア conformer 数(30)
#3		<u>NAMECO</u>	文字型	ファイル名(" ")
#4	上位スコアの出力	<u>SCOREN</u>	整数型	出力上位スコア数(30)
#5		<u>NAMESC</u>	文字型	上位スコアファイル名(" ")
#6	化合物順位の出力	<u>COMRnk</u>	整数型	出力上位化合物数(1)
#7	クラスタリング 実行	<u>CLUSTE</u>	選択型	クラスタリングの実行 (<u>NO</u> YES)
#8	クラスタリング 方法	<u>CLUMET</u>	選択型	クラスタリング実行方法 (<u>NEAR</u> FURT MEDI CENT AVER WARD)
#9	クラスタ数指定	<u>NCLUST</u>	整数型	上位スコア conformer 数(1)

A.3 出力ファイル

ドッキングエンジンの出力ファイルを以下に示します。

項番	ファイル名称	出力フェーズ	用途
#1	上位 conformer 座標 ファイル	OUTPUT	最終結果の上位 conformer のスコア情報 および原子情報の出力
#2	上位スコアファイル	OUTPUT	幾何学ハッシュによるドッキング結果、 エネルギー最小化後のドッキング結果、 および最終結果の出力

A.4 ログ出力（簡易出力）

sievgene で大規模なスクリーニング計算を行う際に、ログファイルによってディスク容量を圧迫することを防ぐために簡易出力形式を使用することができます。

簡易出力形式では、以下の情報を出力する。

- ・制御ファイルの INPUT フェーズで指定した情報
- ・「INFORMATION>」で始まる行
- ・ドッキング結果情報
- ・化合物順位情報

出力例

ここでは、以下の条件で sievgene を実行したときのログ出力例を示します。

- ・入力リガンドファイルに 3 分子の情報が記載されている。
- ・入力リガンド分子名は「compound.1」「compound.2」「compound.3」。
- ・「簡易出力」を指定（制御ファイルの INPUT フェーズに “LOGFOR= SHOR” を記載）
- ・化合物順位を 3 位まで出力（制御ファイルの OUTPUT フェーズに “COMRNK= 3” を記載）

```
*****
*
*          sievgene(06/08/01)          *
*
* Japan Biological Information Research Center *
*      Aomi 2-41-6, Koto-ku, Tokyo 135-0064 *
*              JAPAN                    *
*
*          v3.100: Aug. 1, 2006.        *
*
*****

+++++
+
+ INFORMATION> sievgene(06/08/01)      +
+      READ INPUT SPECIFICATION      +
+
+
+++++
INFORMATION> INPUT
    1) PROTEIN DATA
      1-1) TOPOLOGY FILE
          FORMAT      :asci i
          NAME        :Pro.tpl

      1-2) COORDINATE FILE
          FORMAT      :asci i
          NAME        :Pro_md.pdb

      1-3) POCKET POINT FILE
          FORMAT      :asci i
          NAME        :point.pdb
```

(つづく)

(つづき)

```
2) LIGAND DATA
  2-1) INPUT TYPE  :FILE
        FORMAT     :MOL2
        NAME       :Lig_es_1.mol2

  2-2) REFERENCE COORDINATE FILE
        FORMAT     :no read
        NAME       :lig_ref.mol2

3) ATOM DATA
  3-1) ATOM DATA FILE
        NAME       :

4) ASA DATA
        METHOD     :pairwise
        NAME       :

5) LOG FORMAT     :short

INFORMATION> INPUT POCKET CORRDATE FILE
INFORMATION> INPUT PROTEIN DATA FILE
INFORMATION> INPUT FORMATTED TOPOLOGY FILE
INFORMATION> GRID
INFORMATION> GRID SIZE
              NUMBER OF GRID :          60
INFORMATION> READ GRID POTENTIAL:
INFORMATION> CONF
INFORMATION> DOCK
INFORMATION> OUTPUT
INFORMATION> GENERATE PROTEIN POCKET
INFORMATION> POCKET CENTER COORDINATE
```

(つづく)

(つづき)

```
INFORMATION> SCREENING NUMBER :          1
INFORMATION> REFERENCE COORDINATE NOT READ
INFORMATION> INPUT LIGAND FILE
INFORMATION> LIGAND ANALYZE
INFORMATION> LIGAND ANALYZE
INFORMATION> LIGAND ACCEPTOR, DONAR
INFORMATION> GLOBAL SEARCH START
INFORMATION> CONFORMER NUMBER =          1 GENERATION =          1
INFORMATINO> CREATE NEXT CONFORMER
INFORMATION> LIGAND ANALYZE
INFORMATINO> NEW CONFOMER NUMBER =          1
INFORMATION> LOCAL SEARCH START
INFORMATION> LOCAL SEARCH FINISHED

INFORMATION> AFTER MINIMIZE RANKING
INFORMATION> AFTER MINIMIZE RANKING :EXPERIMENT ID=1
INFORMATION> LOCAL SCORE RANKING          3

COMPOUND NAME : compound.1
FILE NAME      : Lig_es_1.mol2

      SCORE(/100)  dG      HIT      MTS      rotNum  ASA      ELE      HYD      VDW      SURFACE  RMSD
@  1  -0.88  -1.56  -54.24  -144.36  2    -79.27  0.06  -5.57  -3.21  59.75  23.90
@  2  -0.83  -1.32  -48.36  -138.20  2    -78.95  2.65  -2.95  -3.48  30.09  21.50
@  3  -0.74  -1.14  -42.48  -121.03  2    -68.41  2.31  -4.44  -3.03  32.53  21.31
REFERENCE COORDINATE DATA
      999.90  999.90   0.00   0.00

INFORMATION> WRITE RESTART FILE :

INFORMATION> SCREENING NUMBER :          2
INFORMATION> REFERENCE COORDINATE NOT READ
INFORMATION> INPUT LIGAND FILE
```

(つづく)

(つづき)

```

INFORMATION> LIGAND ANALYZE
INFORMATION> LIGAND ANALYZE
INFORMATION> LIGAND ACCEPTOR, DONAR
INFORMATION> GLOBAL SEARCH START
INFORMATION> CONFORMER NUMBER =          1 GENERATION =          1
INFORMATINO> CREATE NEXT CONFORMER
INFORMATION> LIGAND ANALYZE
INFORMATINO> NEW CONFOMER NUMBER =          0
INFORMATION> LOCAL SEARCH START
INFORMATION> LOCAL SEARCH FINISHED

INFORMATION> AFTER MINIMIZE RANKING
INFORMATION> AFTER MINIMIZE RANKING :EXPERIMENT ID=2
INFORMATION> LOCAL SCORE RANKING          3

COMPOUND NAME : compound.2
FILE NAME      : Lig_es_1.mol2

      SCORE(/100)  dG      HIT      MTS      rotNum  ASA      ELE      HYD      VDW      SURFACE  RMSD
@  1  -0.85  -2.82  -46.82  -133.34   0   -73.47   3.45  -12.84  -2.07   52.35  18.89
@  2  -0.83  -3.04  -57.82  -117.15   0   -47.97  -9.78  -22.66  -2.62   89.72  21.78
@  3  -0.80  -2.49  -41.75  -129.95   0   -75.81   6.07  -7.66  -2.38   69.76  21.03

REFERENCE COORDINATE DATA
      999.90  999.90   0.00   0.00

INFORMATION> WRITE RESTART FILE :
INFORMATION> SCREENING NUMBER :          3
INFORMATION> REFERENCE COODINATE NOT READ
INFORMATION> INPUT LIGAND FILE
INFORMATION> LIGAND ANALYZE
INFORMATION> LIGAND ANALYZE
INFORMATION> LIGAND ACCEPTOR, DONAR

```

(つづく)

(つづき)

```

INFORMATION> GLOBAL SEARCH START
INFORMATION> CONFORMER NUMBER =          10 GENERATION =          1
INFORMATINO> CREATE NEXT CONFORMER
INFORMATION> LIGAND ANALYZE
INFORMATINO> NEW CONFOMER NUMBER =          10
INFORMATION> LOCAL SEARCH START
INFORMATION> LOCAL SEARCH FINISHED

INFORMATION> AFTER MINIMIZE RANKING
INFORMATION> AFTER MINIMIZE RANKING :EXPERIMENT ID=3
INFORMATION> LOCAL SCORE RANKING          3

COMPOUND NAME : compound.3
FILE NAME      : Lig_es_1.mol2

      SCORE(/100)  dG      HIT      MTS      rotNum  ASA      ELE      HYD      VDW      SURFACE  RMSD
@  1  -0.80  -1.56  -51.29  -123.35  2      -61.45  -3.55  -11.95  -2.99  71.44  23.76
@  2  -0.70  -1.28  -43.42  -108.09  2      -54.88  -1.84  -10.98  -2.46  71.36  22.87
@  3  -0.69  -1.29  -43.40  -104.32  2      -50.97  -2.96  -12.43  -2.73  67.01  22.64

REFERENCE COORDINATE DATA
      999.90  999.90   0.00   0.00
INFORMATION> ALL EXPERIMENT FINISHED

INFORMATION> COMPOUND RANKING          3

      COMPOUND NAME          SCORE(/100)  dG      HIT      MTS
1  compound.1          -0.88  -1.56  -54.24  -144.36
2  compound.2          -0.85  -1.81  -46.55  -133.07
3  compound.3          -0.80  -1.56  -51.29  -123.35

```

(つづく)

(つづき)

TOTAL EXPERIMENT NUMBER	:	3
TOTAL TRIAL NUMBER	:	23135
TOTAL CPU TIME(S)	:	19.09570
INPUT	:	6.637695
GRID	:	2.7344227E-02
CONF	:	1.2695789E-02
DOCK	:	1.643554
MIN	:	10.13281

(余白)

B ユーティリティ

B.1 make_point

受容体ポケットの位置は、プローブ点の集合で指定し、PDB 書式で与えます。既知リガンドの情報がない場合、make_point.f を用いて PDB 書式でのプローブ点の集合を作成することができます。

入力データ

- (1) 受容体のファイル名
- (2) プローブ点を発生させる球状領域の半径
- (3) 中心とする原子の番号

使用例

標準入力から順次、受容体のファイル名、プローブ点を発生させる球状領域の半径、中心とすべき原子の番号を指定すると、point.pdb という固定の名前で、PDB 書式でのプローブ点の集合を出力します。

```
% make_point.exe
```

【注意】 sievgene の入力情報として、そのファイル書式 (POINTC= PDB) と、ファイル名 (NAMEP0= point.pdb) を指定します。

【注意】 作成した point.pdb は、より実際に合うようにエディターなどで編集すると良いでしょう。

(余白)

C 乱数の選択

C.1 目的

sievgene で使用する乱数はコンパイラベンダが提供している組み込み手続き "random_number" を使用していますが、使用するコンパイラにより結果が異なるという問題が発生します。

そのため、ユーザが使用する乱数として sievgene でのオリジナルとコンパイラの組み込み手続きのいずれかを選択します。

C.2 指定方法

Makefile のコンパイルオプションで、使用する乱数を指定します。

"-D_FORTRAN_RANDOM" 指定あり : コンパイラシステムが提供する乱数(random_number)を使用します。

指定なし(デフォルト) : myPresto システムのオリジナル乱数を使用します。

C.3 ログ表示

下記のように sievgene の実行開始時のログに使用する乱数の種別を表示します。

```
*****
*
*          sievgene(08/07/24)
*
* Biomedical Information Research Center
*   2-3-26 Aomi, Koto-ku, Tokyo 135-0064
*           JAPAN
*
*          v4.204 : Aug  1 2011.
*
*****

RANDOM NUMBER METHOD : FORTRAN RANDOM LIBRARY
```

"-D_FORTRAN_RANDOM" 指定の sievgene のログ

```
RANDOM NUMBER METHOD : SIEVGENE ORIGINAL RANDOM
```

コンパイル指定なしの sievgene のログ

myPresto 4.204

- sievgene -