myPresto 4.2

- sievgene -

USER MANUAL

Version 1.0

Copyright (C) 2006-2010 National Institute of Advanced Industrial Science and Technology (AIST)
Copyright (C) 2006-2010 Japan Biological Informatics Consortium (JBIC)

About this document

This document is a separate volume of the "*myPresto* 4.2 USER MANUAL". Copyrights, program licensing terms, the authors, and works cited are as stated in the "*myPresto* 4.2 USER MANUAL".

Acknowledgments

This software was developed with the help of the New Energy and Industrial Technology Development Organization (NEDO) and the Ministry of Economy, Trade, and Industry (METI). We would like to express our gratitude to these organizations for their assistance.

This software was developed as part of a research project advanced by the late Dr. Yoshimasa Kyogoku.

Table of Contents

1 sievgene 5

1.1 Method of Execution 5

1.2 Creating Input Data 5

1.2.1 Control File 5

1.2.1.1 PHASE> INPUT GROUP 7

1.2.1.2 PHASE> GRID GROUP 10

1.2.1.3 PHASE> CONF GROUP 16

1.2.1.4 PHASE> DOCK GROUP 18

1.2.1.5 EXE> MIN GROUP 24

1.2.1.6 PHASE> OUTPUT GROUP 24

2 Sample Calculation 27

2.1 Sample-1 : Receptor ? low-molecular compound docking 27

2.2 Option combinations 44

A Input/Output Files 51

A.1 Input/output files of docking engine 51

A.1.1 Explanation of phases 51

A.2 Input files 52

A.2.1 Control file 53

A.3 Output files 61

B Utilities 63

B.1 make_point 63

(Blank)

1 sievgene

1.1 Method of Execution

Using protein PDB, topology, pocket point PDB, and ligand MOL2 files as input, sievgene executes protein-ligand docking, produces an evaluation score, and outputs the results as files.

These input files and docking conditions are specified in a control file. sievgene reads in the control file by means of standard input and operates accordingly.

```
% sievgene < control_file > output
```

1.2 Creating Input Data

1.2.1 Control File

The control file consists of the following groups. Each group is ended using "QUIT".

- **PHASE> INPUT GROUP** : Specifies the main input file names.
- **PHASE> GRID GROUP** : Specifies options for grid potential generation.
- **PHASE> CONF GROUP** : Specifies options for ligand conformer generation.
- **PHASE> DOCK GROUP** : Specifies global search options.
- **EXE> MIN GROUP** : Specifies local search options.
- **PHASE> OUTPUT GROUP** : Specifies the output of the final results.
- **EXE> SIEV** : Indicates the end of the control file.

Docking begins when the **"EXE> SIEV"** line is read.

【Note】A grid potential can be generated without executing docking (method whereby the CONF, DOCK, MIN and OUTPUT phases are not specified: see "2 Calculation Example") . In this case, "EXE> SIEV" is not needed in the input file.

```
PHASE> INPUT
  LIGAND  =    MOL2
  NAMELI  =    Lig_md2.mol2 ; Ligand coordinate position file
  REFERE  =    MOL2
  NAMERE  =    Lig_md.mol2  ; Ligand reference position file
  TOPOLO  =    FORM
  NAMETO  =    Pro.tpl      ; Protein topology
  COORDI  =    PDB
  NAMECO  =    Pro_md.pdb   ; Protein coordinates
  POINTC  =    PDB
  NAMEPO  =    point.pdb    ; Probe points indicating pocket region
  ATOMDAta     = READ       ; Input of atom data file
PHASE> GRID
  GRIDPOtential = BINA            ; Grid potential input file
  NAMEGRid = grid.file            ; Specification of grid potential input file
  QUIT
PHASE> CONF
  CONFLImit        = 100000   ; Maximum number of conformer generation trials
  CONFORmernumber = 100  ; Number of generated conformers (changing the ligand conformation)
  QUIT
PHASE> DOCK
  METHOD       =    FLEX
  GENERAtion   = 1      ; Number of generations for narrowing search
  SCORENumber  = 10     ; Number of high scores registered per ligand
  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        ; Output format of ligand coordinate file
  NAMECOordinate = ex.cor      ; Name of ligand coordinate file for each experiment
  CANDIDatenumber= 10          ; Number of high scores for which ligand coordinates are
                                 output in each experiment.
  SCORENumber    = 10          ; Number of high scores output in the score file.
  NAMESCore      = ex.score    ; Score file name
  QUIT
EXE> SIEV
```

Example of sievgene control file

Explanation of control file commands

Required	◎
Can be omitted	△
Required when user specifies a certain function	○

1.2.1.1 PHASE> INPUT GROUP

The input group specifies external input files such as the PDB file and topology file of the protein, the pocket coordinate file of the protein, and the ligand coordinate specification file.

【Note】For the formats of the external files, see "A. File Formats" at the end of this manual.

Items specified in the INPUT group:

(1) Protein topology
(2) Protein coordinates
(3) Ligand structure and coordinates
(4) Ligand reference coordinates for RMSD calculation
(5) Protein pocket point coordinates
(6) Accessible Surface Area model
(7) Restart
(8) Search target
(9) Damping factor of atomic radius of protein

(1) Specification of protein topology

TOPOLOgy: Format of protein topology file (△)
=NOREad ; Not read (default)
=FORMAtted ; Formatted ASCII file
=BINAry ; Binary file
NAMETOpology=(Topology file name, 80 characters or less. When TOPOLOgy=[FORM|BINA])

(2) Specification of protein coordinates

Specify the protein coordinates. When BINAry is specified, the protein topology file specified in (1) is required for input of the data of the atoms forming the protein.

COORDInate: Format of 3-dimensional protein coordinate file (◎)
=PDB ; PDB file format
=BINAry ; Binary file
NAMECOordinate=(Coordinate file name, 80 characters or less)

(3) Specification of ligand structure and coordinates

Specify the structure and coordinates of the ligand to be docked. When a multi-format (a format linking the files of multiple molecules) file is specified, multiple ligands are docked.

```
LIGANDcoodinate: Format of ligand file (◎)
    =MOL2          ; MOL2 format (default)
    =PDBX          ; PDBX format (cosgene topology file + PDB format integrated by XML)
NAMELIgandordinate=(Coordinate file name, 80 characters or less. (◎) )
```

(4) Specification of ligand reference coordinates for RMSD calculation

```
REFEREncecoodinate: Format of ligand reference coordinate file (◎)
    =MOL2          ; MOL2 format (default
    =PDBX          ; PDBX format (cosgene topology file + PDB format integrated by XML)
NAMEREferenceordinate=(Coordinate file name, 80 characters or less. (◎))
```

(5) Specification of protein pocket point coordinates

Specify the protein pocket point coordinates.

```
POINTCoordinate : Format of protein pocket point coordinates (◎)
    =PDB           ; PDB file format
    =BINAry        ; Binary file
NAMEPOintcoordinate=(Name of protein pocket point coordinates, 80 characters or less (◎) )
```

(6) Specification of Accessible Surface Area model

Specify the method used in sievgene to calculate the potential at the Accessible Surface Area (hereafter ASA).

```
ASAMEThod : ASA calculation method (△)
    =PAIRwise      ; Pairwise (default)
    =RICHmond      ; Richmond
```

(7) Restart specification

Specify input/output of information for restart from the point of interruption when a screening job is interrupted.

RESTARtjob: Restart an interrupted job. (△)
=NO ; Execute the job from the beginning (default)
=YES ; Execute the job from where it was interrupted
NAMERInput= (Name of restart file, 80 characters or less) (△)
NAMEROutput= (Name of restart file, 80 characters or less) (△)

(8) Specification of search target

Specify the search target so that the specified 1, 2, or 3 protein atoms are targets of the search

SETTARget: Read the search specification information file (○)
=NOREad ; Do not use (default)
=READ ; Use
NAMETArget=(file name of search specification information file, 80 characters or less (○))

(9) Damping factor of atomic radius of protein

Specify the coefficient of the atomic radius of the protein.

DAMPPA: Damping factor of the atomic radius of the protein (△)
=1.0 ; (default)

(10) QUIT

Indicates the end of input of the PHASE> INPUT GROUP.

1.2.1.2 PHASE> GRID GROUP

The potential caused by the atoms of the protein is generated in seivgene as a grid potential distributed over a grid, and the potential is evaluated between the grid potential and the ligand. The parameters for generation of the grid potential are specified in this group.

Items specified in the GRID group：

（1）Input/output of grid potential
（2）Receptor atom registration distance
（3）Atomic radius offset for generation of grid potential
（4）Radius offset for generation of ASA mesh nodes
（5）Size margin of grid potential
（6）Number of grid potential smoothing repetitions
（7）van der Waals radius damping factor
（8）PB (Poisson Boltzmann) equation
（9）Input/output of PB-Grid potential
（10）Mesh number
（11）Size margin of PB-Grid potential
（12）Debye-Huckel shield constant
（13）Dielectric constant
（14）Acceleration factor
（15）Upper limit of CPU time

（1）Input/output of grid potential

The grid potential is determined by the shape of the protein and its coordinates. The grid potential in the same protein is always the same, and thus the grid potential can be reused.

```
GRIDPOtential：Grid potential input specification (△)
    =NOREad        ; Do not read (default)
    =ASCIi         ; ASCII format
    =BINAry        ; BINARY format
OUTGRIdpotential：Grid potential output specification (△)
    =NOWRite       ; Do not output (default)
    =ASCIi         ; ASCII format
    =BINAry        ; BINARY format
```

NAMEGRidpotential= (Name of grid potential file, 80 characters or less) (○)

【Note】When ASCII format is used, the data is converted and thus the resulting data is not in complete agreement.

(2) Receptor atom registration distance

Specify the distance over which receptor atoms are registered.

Atoms on the protein within the specified distance from the pocket point are registered as receptor atoms.

PROBDIstance : Receptor atom registration distance (Å) (△)
=6.5 ; (default)

(3) Atomic radius offset for generation of grid potential

The distance between an atom on the protein and a grid node is used to determine whether the node is inside or outside the atom.

The ASA expresses the area that can be covered by the center of a sphere rolled over the surface of an atom, and the radius of this sphere is specified here.

RADVDW : van der Waals radius offset (Å) (△)
=0.6 ; (default)
RADELE : coulomb radius offset (Å) (△)
=0.6 ; (default)

(4) Radius offset for generation of ASA mesh nodes

ASA is expressed in sievgene using nodes (mesh nodes) arranged at equal intervals on the ASA surface. Specify the offset added to the atomic radius to express the surface of each atom.

RADMESh : Mesh node offset (Å) (△)
=1.6 ; (default)

(5) Size margin of grid potential

The size of the grid potential is that of the rectangular parallelepiped formed by the maximum and minimum values of the x, y, and z coordinates of the pocket point

coordinates. Specify the margin of this size.

MARGIN : Grid potential margin (Å) (△)
=8 ; (default)

(6) Number of grid potential smoothing repetitions

The generated grid potential is calculated at each node, and in some cases, the difference in the potential between neighboring nodes may be unnaturally large. To smooth these differences, smoothing is performed. The higher the number of smoothing repetitions, the smoother the potential surface becomes, however, too many repetitions will result in a uniform potential.

ITERATion : Number of smoothing repetitions (△)
=3 ; (default)

(7) van der Waals radius coefficient

Specify the van der Waals radius coefficient for calculation of the grid potential

DANPVW : van der Waals radius damping factor (△)
=1.0 ; (default)

(8) Poisson Boltzmann equation

Specify whether or not the Poisson Boltzmann equation is used when generating the grid potential.

USEPBG : Specification of PB grid potential (○)
=NO ; Do not use the PB equation (default)
=YES ; Use the PB equation

(9) Input/output of PB-Grid potential

The result of the electrostatic field calculation using the Poisson Boltzmann equation is determined by the shape and coordinates of the protein. The result of the electrostatic field calculation will always be the same for the same protein, and thus the result can be reused.

PBGRID : Specification of input of result of electrostatic field calculation (△)

=NOREad ; Do not read (default)
=ASCIi ; ASCII format
=BINAry ; BINARY format
OUTPBG : Specification of output of electrostatic field calculation result (△)
=NOWRite ; Do not output (default)
=ASCIi ; ASCII format
=BINAry ; BINARY format
NAMEPB= (PB file name, 80 characters or less) (○)

【Note】 The PBGRID, OUTPBG, and NAMEPB options are only effective when USEPBG= YES.
【Note】 When ASCII format is used, the data is converted and thus may not be in complete agreement.

(10) Mesh number

Specify the mesh number in the x, y, and z directions when the Poisson Boltzmann equation is used.

NMESHX : Mesh number (x direction) (△)
=100 ; (default)
NMESHY : Mesh number (y direction) (△)
=100 ; (default)
NMESHZ : Mesh number (z direction) (△)
=100 ; (default)

【Note】 The NMESHX, NMESHY, and NMESHZ options are only effective when USEPBG= YES.

(11) Size margin of PB-Grid potential

The size of the PB-Grid potential is that of the rectangular parallelepiped formed by the maximum values and minimum values of the x, y, and z coordinates of the pocket point coordinates.

Specify the margin of this size.

MARGPB : Margin of PB-Grid potential (Å) (△)
=5.0d0 ; (default)

【Note】 The MARGPB option is only effective when USEPBG= YES.

（12）Debye-Huckel shield constant

Specify the Debye-Huckel shield constant for calculation of the electrostatic field when the Poisson Boltzmann equation is used.

KAPPAV： Debye-Huckel shield constant（1/Å）（△）
=0.0d0 ;(default)

【Note】The KAPPAV option is only effective when USEPBG= YES.

（13）Dielectric constant

Specify the dielectric constants of the solvent, protein, middle region, and vacuum region for calculation of the electrostatic field when the Poisson Boltzmann equation is used.

DIESOL： Dielectric constant of the solvent（△）
=78.5d0 ;(default)
DIEPRO： Dielectric constant of the protein（△）
=4.0d0 ;(default)
DIEINT： Dielectric constant of the middle region（△）
=78.5d0 ;(default)
DIEVAC： Dielectric constant of the vacuum region（△）
=1.0d0 ;(default)

【Note】The DIESOL, DIEPRO, DIEINT, and DIEVAC options are only effective when USEPBG= YES.

（14）Acceleration factor

Specify the acceleration factor for calculation of the electrostatic field when the Poisson Boltzmann equation is used.

ACCELE： Acceleration factor（△）
=1.6d0 ;(default)

【Note】The ACCELE option is only effective when USEPBG= YES.

(15) Upper limit of CPU time

Specify the upper limit of CPU time for calculation of the electrostatic field when the Poisson Boltzmann equation is used.

CPULIM: Upper limit of CPU time (s) (△)
=10800 ; (default)

【Note】The CPULIM option is only effective when USEPBG= YES.

(16) Grid potential interpolation method

Specify the grid potential interpolation method at any point on the grid potential.

INTERPorate: Grid potential interpolation method (△)
=LAGRange ; 1-degree Lagrange interpolation (default)
=BSPLine ; 3-degree B-Spline interpolation

(17) QUIT

Indicates the end of input of the PHASE> GRID group.

1.2.1.3 PHASE> CONF GROUP

Rotatable torsions are detected in sievgene in the input ligand, the torsions are randomly rotated, and screening is performed using the new conformers.

This group specifies parameters for generating new conformers.

Items specified in the CONF group：

（1）Number of conformer generation trials
（2）Number of conformers generated
（3）Sorting of atom data
（4）Atomic radius damping factor for determination of van der Waals contact
（5）Torsion rotation angle number
（6）Atom model

（1）Number of conformer generation trials

The angle of rotation in conformer generation is determined by a random number, and in some cases the same conformer is generated multiple times and then generation ends. To prevent this from happening, specify a limit for the number of conformer generation trials.

CONFLimit： Limit on number of conformer generation trials (△)
=10000000 ; (dafault)

（2）Number of conformers generated

Specify the number of conformers generated.

COMFORmernumber： Number of conformers generated (△)
=100 ; (default)

（3）Sorting of atom data

Specify whether or not the order of atom data in the output file should be the same as that in the input file.

SORTATom：Sort atomic data (△)
=NO ; Do not return to original order (default)
=YES ; Return to original order

(4) Atomic radius damping factor for determination of van der Waals contact

If the atoms of a ligand in a newly generated conformer are in van der Waals contact, the conformer is not used. When comparing the distance between the atoms and the sum of their radii to determine whether or not they are in van der Waals contact, the atomic radii are multiplied by this damping factor.

DAMPINgfactor : Damping facor for determination of van der Waals contact (△)
=0.7 ; (default)

(5) Torsion rotation angle number

Specify the number of angles through which a torsion can be rotated. When the default "3" is specified, the torsion can be rotated in increments of 120o.

PHASETorsion : Torsion rotation angle number (△)
=3 ; (default)

(6) Atom model

Specify the atom model for the ligand molecule. Either the conventional atom model "all atom model" or the "united atom model", which replaces ?CH, CH2, and ?CH3 with ?C', can be specified.

AToMMoDeL : Atom model
= ALL atom model ; all atom model (default)
= UNITed atom model ; united atom model

(7) QUIT

Indicates the end of input of the PHASE> CONF group.

1.2.1.4 PHASE> DOCK GROUP

This group specifies global search parameters and score calculation parameters.

Items specified in the DOCK group：

（1）Docking method
（2）Number of generations for narrowing the search area
（3）Number of conformers generated
（4）Side length of triangle of attached atoms
（5）Atom type matching of bonding face
（6）Grid potential interpolation method
（7）Score coefficients
（8）Distance for determination of atoms in pocket
（9）Pocket center
（10）Evaluation of hydrogen bond of protein and ligand
（11）Rotation of ligand-OH group during docking
（12）Rotation of protein side chain during docking
（13）Fine adjustment of ligand coordinates after docking
（14）Number of local search candidates
（15）Switching of docking atoms
（16）Estimation of energy

(1) Docking method

Two docking methods are available in sievgene: rigid docking which does not change the conformer of the input ligand, and flexible docking which repeats docking while changing the conformer.

METHODofdocking： Specify the docking method (△)
=FLEXible ; FLEXIBLE docking (default)
=RIDId ; RIGID docking

(2) How to generate the points on the protein surface

Selected 3 atoms of compound are overlapped to the three points on the proteinsurface. This option indicates how to generate the points on the protein surface.

PROteinSURface：Specify the points on the protein for the geometrical hashing method.

=ELECtrostatic ; The points on the protein surface are the electrostatic-potential minima/maxima and the point, at which the potential is nearly zero. (default)

=HYDRogen ; The points on the protein surface are the electrostatic-potential minima/maxima. In addition, the potential surface is calculated by using CH_4 probe. And then the potential maxima are adopted.

(3) Number of generations for narrowing the search area

Specify the number of generations to narrow the search range.

GENERAtionnumber : Number of generations for narrowing the search area (△)

=5 ; (default)

(4) Number of conformers generated?

Ligand flexibility can be taken into account in sievgene. Specify and sn upper limit for the number of conformers generated. NUMCONformer : Number of conformer generations (△)

=100 ; (default)

(5) Side length of Triangle of attached atoms

Specify the triangle side length that satisfies the number of docking faces. sievgene calculates the side length that satisfies the number of bonding faces from the specified side range, and determines the bonding faces.

LOWMIN : Lower minimum value of the side length (Å) (△)

=2.5 ; (default)

LOWMAX : Lower maximum value of the side length (Å) (△)

=4.0 ; (default)

UPRMIN : Upper minimum value of the side length (Å) (△)

=7.5 ; (default)

UPRMAX : Upper maximum value of the side length (Å) (△)

=10.0 ; (default)

(6) Atom type matching of attached atoms

The degree of relevancy to attachment is set in sievgene based on the properties of the atom.

This matching can be adjusted to reduce unnecessary attachment determinations and thereby reduce docking time.

Any number from 1 to 5 can be specified, with a higher number resulting in greater restriction of triangles of attached atoms.

MATCHIng: Atom type matching of the triangle of attached atoms△)
=1 ; (default)

(7) Score weighting coefficients

Scores are expressed in sievgene as a sum of five potentials: ASA, coulomb, hydrogen bond, hydrogen bond considering anisotropy, and van der Waals.

To set a weight for each potential, change the score coefficient.

WETVDW : van der Waals potential weighting coefficient (△)
=1.0 ; (default)
WETASA : A.S.A. potential weighting coefficient (△)
=1.0 ; (default)
WETHYD: Hydrogen bond potential weighting coefficient (△)
=1.0 ; (default)
WETANH : Hydrogen potential weighting coefficient considering anisotropy (△)
=1.0 ; (default)
WETELE : coulomb potential weighting coefficient (△)
=1.0 ; (default)

(8) Distance for determination of atoms in pocket

Specify the distance from the central point of the pocket that determines whether or not a ligand atom exists inside the pocket.

RADIUS: Distance for determination of atoms in pocket (Å) (△)
=6.0 ; (default)

(9) Pocket center

Specify the center point of the protein.

If not specified, it will be assumed that the midpoint of the pocket point is the

pocket center.

(9-1) Pocket center coordinates

Specify the absolute coordinates of the pocket center.

POCKCX : Pocket center X coordinate
= 999.9 ; (default)
POCKCY : Pocket center Y coordinate
= 999.9 ; (default)
POCKCZ : Pocket center Z coordinate
= 999.9 ; (default)

(9-2) Pocket center coordinates

Specify the protein atom at the pocket center.

POCKET : Pocket center atom ID
=0 ; (default)

(10) Evaluation of hydrogen bond of protein and ligand

Specify whether or not the hydrogen bond of the protein and ligand is evaluated with consideration given to anisotropy.

EVALHB : Evaluate hydrogen bond with consideration given to anisotropy (○)
=YES ; Evaluate (default)
=NO ; Do not evaluate

(11) Rotation of ligand-OH group during docking

Specify whether or not the end of the ligand-OH group is rotated during docking.

ROTLOH : Rotate ligand-OH group (○)
=YES ; Rotate (default)
=NO ; Do not rotate

(12) Rotation of protein side chain during docking

Specify whether or not a hydrogen bondable side chain of the protein is rotated during docking.

ROTPSC : Rotate protein side chain (○)
=YES ; Rotate
=NO ; Do not rotate (default)

(13) Fine adjustment of ligand coordinates after docking

Specify the number of times the coordinates are moved during fine adjustment of the ligand coordinates after docking.

MOVNUM : Number of times coordinates are moved (○)
=10 ; (default)

(14) Number of local search candidates

Specify the number of local search candidates.

CANDID : Local search candidates (△)
=30 ; (default)

(15) Switching of docking atoms

Specify whether or not atoms that are not targets of superimposition are excluded from the calculation when a ligand is superimposed on a protein.

DOCKSP : Switching of docking atoms (△)
=NORM ; Do not exclude non-targeted atoms (default)
=FAST ; Exclude

(16) Estimation of energy

Only perform the minimize calculation on the input atoms; do not perform docking.

OPTIME : Estimation of energy (△)
=OPT ; Perform docking (default)
=ENE ; Do not perform docking

(17) CPU time limit

CPU time limit for docking of each compound. If the CPU time exceeds the limit, the results at the indicated time limit are recorded as the final results.

CPUTIM: CPU time limit for docking of each compound (second) (△)
=30.0 ;(default)

(18) QUIT

Indicates the end of PHASE> DOCK group input.

1.2.1.5 EXE> MIN GROUP

This group specifies local search parameters and score calculation parameters.

Items specified in the MIN group and specification methods are the same as in the cosgene EXE> MIN group.

See "EXE> MIN GROUP" in the "cosgene" chapter of the user manual.

1.2.1.6 PHASE> OUTPUT GROUP

The OUTPUT group specifies output of the high-score conformer coordinate PDB file and score file.

Items specified in OUTPUT group：

（1）Output of high-score conformer coordinate PDB file

（2）High score output

（1）Output of high-score conformer coordinate PDB file

COORDInate： Format of high-score conformer coordinate PDB file output (△)

=NOWRite ; Do not output (default)

=MOL2 ; MOL2 file format

=PDBX ; PDBX format (cosgene topology file + PDB format integrated by XML)

=PDB ; PDB format

CANDIDatenumber： Number of high-score conformers (△)

=30 ; (default)

NAMECOorinate=(Name of high-score conformer coordinate file, 80 characters or less (○))

【Note】COORDI= MOL2 is only effective when LIGAND= MOL2 .

（2）High score output

SCORENumber：Number of high scores output (△)

=30 ; (default)

NAMESCore=(Name of high-score file, 80 characters or less. (◎))

(3) QUIT

Indicates the end of PHASE> OUTPUT group input.

(Blank)

2 Sample Calculation

2.1 Sample-1：Receptor ? low-molecular compound docking

The procedure for docking a low-molecular compound to a protein or other receptor is described here. The "1ai5" and "1c1e" directories contain data on the receptor and low-molecular compound. Move to each directory and execute the job. "allrun.sh" can be used to perform test calculations regarding the two complexes. In each case the output results are stored in the directory that contains the receptor data or low-molecular compound data.

（1）Method for collectively performing the sequence from grid potential generation to docking

The INPUT phase settings are explained first. Prepare the low molecule to be docked in mol2 file format (http://www.tripos.com/custResources/mol2Files/). Create the molecular model using the all atom model including hydrogen, and be sure to enter the atomic charge. It is sufficient to include only the <MOLECULE>, <ATOM> and <BOND> sections of the mol2 file format. With regard to the low molecule, if the PDB file and tpl file exist, the mol2 file can be generated with the "tpl2mol2" tool included with myPresto. The mol2 file can also be created with the free babel/openBabel.

【Note】sievgene cannot handle metal complexes.

Prepare the protein using the all atom model including hydrogen. An all atom PDB file and a topology file indicating the charge are needed. These files can be created with "myPresto-tplgene" or other similar tool.

For the ligand, specify the file format (LIGAND=MOL2) and the file name (NAMELI). The deviation RSMD between the positions of the conformer after docking and the reference coordinates can be calculated. If you have a correct receptor — a ligand complex and the ligand coordinates are known, the correct coordinates can be used as the reference coordinates to verify the accuracy of docking. For the reference coordinates, specify the file format (REFERE=MOL2) and file name (NAMERE). For the receptor, specify the topology file format (TOPOLO=FORM) and file name (NAMETO) and the coordinate format (COORDI=PDB) and file name (NAMECO).

Specify the position of the receptor pocket as a collection of probe points, given in the sample as the collection of points "point.pdb" in PDB format. The ligand torsion

search space and potential grid are set with respect to this probe point collection in the x, y, and z directions within the range ±PROBDI/MARGIN specified in the following grid phase. If you have the coordinates of a compound with a known ligand, those ligand coordinates can be used. If you do not have information on a known ligand, you can create a collection of probe points in PDB format using "make_point.f".

To compile :

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

To use :

Type "make_point". From standard input, specify in succession the file name of the receptor, the radius of the spherical region that generates the probe points, and the number of the atom that will be at the center. The collection of probe points will be output in PBD format with the fixed name "point.pdb".

```
% make_point.exe
```

It is desirable to edit this "point.pdb" file with an editor to make it more closely resemble the actual state. Specify the file format "POINTC" and the file name "NAMEPO" for the probe point collection.

With regard to the structure of the docking result, sievgene has a function that indicates how many of the docked ligand atoms are within a vicinity of radius 6 Å of a specific atom of the protein. However, in the following example this is not specified (SETTAR=NORE). The radius of the protein atom is

(radius of protein atom) = (damping factor) × (default radius setting)

and can be adjusted using the damping factor (DAMPPA) .

In GRID generation phase, a previously generated grid potential can be read (GRIDPO=BINA) or the grid potential can be newly generated (GRIDPO=NORE). In the latter case, the grid potential can be output as a file (OUTGRI=BINA) or not output (OUTGRI=NOWR).

A range of generation of a geometric hash table is specified as the ligand search space, and with the probe points (NAMEPO) specified in the INPUT phase as reference points, this is specified within the range of the radius PROBDI around the points. Likewise, the range of generation of the grid potential is specified within the ± MARGIN range in the x, y, and z directions taking the probe points (NAMEPO) as reference

points. As PROBDI is made larger, the torsion search range widens and the calculation time increases. When initially deciding the probe point collection, it is important to use a selection method that will result in the minimum required number of the most suitable points. The grid size is fixed at 60 x 60 x 60, and thus if the MARGIN is increased, the grid width will spread and decrease the precision. The grid width should be from about 0.2 to 0.35Å, and should not be more than 0.5Å. The grid potential is smoothed the number of times specified in ITERAT. The van der Waals potential and the electrostatic potential are partitioned into interior and exterior at the boundary surface of the protein, with the normal potential applied outside the boundary and a smooth virtual potential applied inside. At this time, the probe radii of the accessible surfaces, which are the boundaries, are each specified in RADVDW/RADELE. When RADVDW/RADELE is decreased, wide places between atoms that are in reality inside the protein are determined to be outside. A value of approximately 0.6Å is best. The damping factor of the protein atom in boundary generation at this time is specified in DAMPVW. When creating the geometric hash table, a hydrophobic probe, a hydrophilic positive charge probe, and a hydrophilic negative charge probe are rolled over the protein surface, a collection of points where these can bind easily is created, and the sides of the triangle that joins these points are stored in the hash table. The interval between the points is specified in RADMESH. A good result is obtained when a value of about 1.2 to 1.6Å, normally 1.4Å, is specified in RADMESH.

When USEPBG=NO, the electrostatic field is given by the equation taking $\varepsilon = 4R$. When USEPBG=YES, the electric field obtained when the Poisson-Boltzmann equation is solved by the self-consistent boundary (SCB) method is applied. Currently, the SCB method is inferior in terms of both speed and accuracy to the equation taking $\varepsilon = 4R$.

The method of generating the ligand conformer is specified in the CONF phase. Conformer generation is accomplished by rotating the dihedral angle of the single bond of the linear chain segment excluding the molecular ring, carrying out a random search. The maximum number of random search trials is specified in CONFLI, and the maximum number of conformers generated at once is specified in CONFOR. The dihedral angle is rotated through the angle specified in 360/PHASET. In the example, PHASET=3 and thus each rotation is 120o. The input atoms are rearranged during conformer generation to form a tree-shaped graph, however, these can be returned to their original order when the result is output (SORTAT=YES) . It is determined whether or not collisions in the molecule of the generated conformers will occur, and any conformers that have collisions are eliminated. The damping factor for the van der

Waals radius used at this time is specified in DAMPIN. Conformers are generated by means of a rough rotation of only the dihedral angle and thus collisions within the molecule can easily occur. For this reason, a value of about 0.7 for DAMPIN is recommended. ROTTER specifies whether or not a ligand end such as an ?OH group is rotated. Even if rotation of the ?OH group is not specified in ROTTET, the ?OH group can still be rotated and torsion search performed after docking in the next DOCK phase.

In the DOCK phase, rough docking of the receptor and low molecule is performed by geometric hashing. Either flexible docking (METHOD=FLEX) , which keeps the ligand flexible, or rigid docking (METHOD=RIGID) , which fixes the ligand to an input conformer, can be performed. After dockings are attempted of the CONFOR number of conformers generated in the CONF phase, the search space is narrowed down to the most likely docking positions, and once again the CONFOR number of different conformers are generated and docking is repeated the number of times specified in GENERAT. As such, the total number of conformers for which docking is attempted is CONFORxGENERAT at the maximum. Restrictions apply to the hash table and the length between the atoms of the referenced ligand. The shortest length is from LOWMIN to LOWMAX and the longest length is from UPRMIN to UPRMAX. This length is adjusted for one conformer within the above range so that the number of references to the hash table is approximately the number in NUMCONF. The hash table indicates whether or not each of the points that form the triangle of the receptor is an H donor and whether it is in the vicinity of an H acceptor atom, and it is determined whether the ligand atoms are H donors or H acceptors. When making these 3-point combinations of receptor and ligand, the number of points forming an H donor ? H receptor pair is calculated. A complete H donor - H acceptor match is given 2 points, an incomplete match is given 1 point, and a non-match such as an H donor and H donor or H acceptor and H acceptor is given 0 points. When the total matching points of the three points is equal to or greater than MATCHING, the torsion superposition is used. Any integral number from 0 to 6 can be specified for MATCHING. When there is a high hydrophilicity and many modes of hydrogen bonding are included, specify a high MATCHING number. When the pocket is hydrophobic and almost no hydrogen bonds are included, specify a low MATCHING number.

WETVDW, WETASA, WETELE, WETHYD, and WETAHB specify the weighting during calculation of the overall score of the van der Waals reciprocal action term, the ASA term, the electrostatic term, the hydrogen bonding term, and the hydrogen bonding term with consideration given to the hydrogen bonding orientation. EVALHB specifies calculation of hydrogen bonding with consideration given to the orientation of hydrogen bonding,

and ROTLOH specifies whether or not the torsion search is performed with the ?OH radical rotated after docking. ROTPSC specifies whether or not flexible docking is performed by rotating the side chain of the protein and giving consideration to the changes in the torsion of the protein. A random torsion search is performed within a range of 0.5 Å along the x, y , and z axes after ligand docking, and MOVNUM specifies the number of times that this search is performed.

In the MIN phase, the steepest gradient method (METHOD=STEEP) or the conjugate gradient method (METHOD=CONJ) is specified for minimization of the energy of plausible torsions after ligand docking. The specification method in this phase is the same as in the MIN phase of cosgene.

Output of the docking results is specified in the OUTPUT phase. The coordinates of the candidates (CANDID) with the highest docking scores are output inthe specified file output format (COORDI=MOL2) with the specified file name (NAMECO). The docking scores are output in the file specified in NAMESC, with the number of scores output specified in SCOREN.

Execution of sievgene :

"all.inp" is specified as the input file.

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

■Input file

```
PHASE> INPUT
  LIGAND  =    MOL2
  NAMELI  =    Lig_es_1.mol2 ; Coordinate position file of ligand
  REFERE  =    MOL2
  NAMERE  =    lig_ref.mol2  ; Reference position file of ligand
  TOPOLO  =    FORM
  NAMETO  =    Pro.tpl  ; Protein topology
  COORDI  =    PDB
  NAMECO  =    Pro_md.pdb  ; Protein coordinates
  POINTC  =    PDB
  NAMEPO  =    point.pdb   ; Probe points indicating pocket region

  SETTAR = NORE
  DAMPPA = 1.0d0
  QUIT
;
; Grid generation and hash table generation
;
PHASE> GRID
  GRIDPOtential    = NORE ; BINA ; Grid potential input
  NAMEGRid         = grid.file   ; Grid potential file name
  OUTGRIdpotential = NOWR        ; Grid potential output
  PROBDIst = 6.5  ; Target receptor atoms within PROBDIST from probe point
  MARGIN   = 6.5  ; Margin around search pocket during grid generation
  ITERAT   = 3    ; Number of repetitions of grid potential smoothing
  RADVDW   = 0.6  ; Set boundary between inside and outside regions of vDW potential
  RADELE   = 0.6  ; Set boundary between inside and outside regions of coulomb potential
  RADMESh  = 1.4  ; Probe radius for mesh node generation

  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
;
```

(Continued on next page)

(Continued from next page)

```
PHASE> CONF
  CONFLImit         = 100000  ; Maximum number of conformer generation trials
  CONFORmernumber   = 100     ; Number of generated conformers (changing the ligand)
  SORTATom            = YES       ; Change order of atoms in output file
 DAMPINg          = 0.7     ; Damping factor for distance between atoms of conformer candidates
  PHASETorsion      = 3       ; Number of torsion rotation candidates for conformer generation
  ROTTER            = YES
  QUIT

PHASE> DOCK
  METHOD        = FLEX
  GENERAtion    = 1    ; Number of generations for narrowing the search area
  NUMCONFomer   = 1000 ; Number of triangles of attached atoms of ligand
  MATCHING  =  4       ; Matching of triangle atom types
  LOWMIN    =  2.5     ; Lower minimum of distance between atoms compared with hash table
  LOWMAX    =  3.5     ; Lower maximum of distance between atoms compared with hash table
  UPRMIN    =  5.0     ; Upper minimum of distance between atoms compared with hash table
  UPRMAX    =  12.0    ; Upper maximum of distance between atoms compared with hash table

  WETVDW= 1.0  ; Weight of each item when calculating overall score
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  RADIUS= 6.0  ; Count the number of ligand atoms existing within RADIUS from target
               ; atom (coordinates) after ligand molecule docking
  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       ; Output format of ligand coordinate file
  NAMECOordinate = ex.cor     ; Ligand coordinate file name in each experiment
  CANDIDatenumber= 3          ; Number of high scores for which ligand coordinates are output
in each experiment
  SCORENumber    = 3          ; Number of high scores output in score file
  NAMESCore      = ex.score   ; Score file name
  QUIT

EXE> SIEV
```

* Please check the highlighted parts. Line 2 was an omission and line 10 and 11 is

for「絞込みの回数」and「リガンドの接合面三角形数」

■Example of output of execution results (standard output)

The end of standard output of the execution results is as shown below. The overall score (SCORE), ⊿G score (dG) , hit-optimized score (HIT) , MTS score (MTS) , number of rotable bondings (rotNum) , ASA item (ASA), electrostatic item (ELE), hydrogen bonding item (HYD), van der Waals item (VDW), percentage of accessible surface of ligand molecule filled in by bonding with receptor (SURFACE), and superimposition deviation excluding hydrogen atoms superimposed on the reference coordinates (RMSD, Å) (RMSD) are shown. CPU time required for each phase is also shown.

```
INFORMATION> TOP SCORE RANKING           3

     SCORE(/100)  dG     HIT     MTS    rotNum   ASA      ELE    HYD     VDW   SURFACE   RMSD
@   1   -4.88  -12.54 -314.03  -835.76    2    -475.94    1.95   -5.18   -9.10   68.95   32.39
@   2   -4.17  -10.65 -265.81  -709.78    2    -402.52    2.15   -7.66   -9.17   68.34   32.58
@   3   -4.09  -10.78 -239.64  -708.69    2    -428.10   24.87   -5.98    0.00   68.61   32.39

     TOTAL EXPERIMENT NUMBER  :          1
     TOTAL TRIAL NUMBER       :                2052
     TOTAL CPU TIME(S)  :   290.1748
         INPUT   :   8.724608
         GRID    :   274.5596
         CONF    :  8.7890625E-03
         DOCK    :  0.1875000
         MIN     :   6.625000
```

■Example of output of execution results (score file)

Output of the score file specified in NAMESC is as follows. The definition of each item is the same as in standard output. The result of docking by geometric hash is shown first, followed by the result of docking after energy minimization, and then the final result once again. If energy minimization is not performed, the final result will be the result of docking by geometrical hash. If energy minimization is performed, the final result will be the result of docking after energy minimization.

```
INFORMATION> BEFORE MINIMIZE RANKING :EXPERIMENT ID=1
 INFORMATION> LOCAL SCORE RANKING           3
 Lig_es_1.mol2

   SCORE     ASA      ELE     HYD      VDW    SURFACE  ATOMS  RMSD    CENTER-RMSD
   -198.24 -154.26    0.94  -44.93    0.00     0.00     19     4.24    0.57
   -197.30 -154.27   -2.24  -40.79    0.00     0.00     19     3.73    0.45
   -194.54 -150.32   -2.33  -41.88    0.00     0.00     19     3.75    0.62

 REFERENCE COORDINATE DATA
   -273.00 -230.01   -1.31  -41.68

 INFORMATION> AFTER MINIMIZE RANKING :EXPERIMENT ID=1
 INFORMATION> LOCAL SCORE RANKING           3
 Lig_es_1.mol2

   SCORE     ASA      ELE     HYD      VDW    SURFACE  ATOMS  RMSD    CENTER-RMSD
   -243.92 -188.15   -3.02  -49.51   -3.24    81.35     24     3.82    0.97
   -241.53 -191.46   -2.81  -44.02   -3.24    79.16     24     3.71    0.42
   -216.30 -168.30   -2.39  -42.40   -3.21    77.32     24     4.46    0.81

 REFERENCE COORDINATE DATA
   -273.00 -230.01   -1.31  -41.68

 INFORMATION> TOP SCORE RANKING           3

     SCORE(/100)  dG      HIT      MTS     rotNum   ASA      ELE    HYD      VDW    SURFACE   RMSD
@   1   -2.44   -8.58 -147.48  -369.69     0     -188.15   -3.02  -49.51   -3.24   81.35     3.82
@   2   -2.42   -8.30 -147.41  -371.23     0     -191.46   -2.81  -44.02   -3.24   79.16     3.71
@   3   -2.16   -7.54 -130.72  -329.17     0     -168.30   -2.39  -42.40   -3.21   77.32     4.46
```

■Example of output of results (coordinate file)

Ligand coordinate output after docking, which is specified in NAMECO in the OUTPUT phase, is as follows.

(a) Output in PDB format (when COORDI= PDB)

The overall score and the result for each item appears at the top of the file, followed by output of the ligand coordinates. The residue name is fixed at LGD. If these coordinates are directly bonded with the input receptor coordinates, the coordinates of the receptor — ligand conplex are produced.

```
REMARK                SCORE =  -212.1914
REMARK             dG-SCORE =    -6.4866
REMARK  HIT-OPTIMIZED SCORE =  -130.5660
REMARK             MTS SCORE =  -336.8068
REMARK                  ASA =  -180.8419
REMARK                  ELE =    -0.3130
REMARK                  HYD =   -29.5172
REMARK                  VDW =    -1.5193
ATOM      1  H   LGD     1      15.303  36.287  35.260
ATOM      2  C   LGD     1      14.608  36.286  36.113
ATOM      3  H   LGD     1      14.919  35.506  36.825
ATOM      4  C   LGD     1      14.752  37.633  36.810
ATOM      5  O   LGD     1      14.465  38.790  36.404
ATOM      6  O   LGD     1      15.657  37.655  37.686
ATOM      7  C   LGD     1      13.292  35.849  35.562
ATOM      8  C   LGD     1      12.432  34.920  36.208
ATOM      9  H   LGD     1      12.428  34.863  37.291
ATOM     10  C   LGD     1      11.250  34.465  35.542
ATOM     11  H   LGD     1      10.487  33.970  36.134
ATOM     12  C   LGD     1      10.850  35.062  34.341
ATOM     13  H   LGD     1       9.847  34.901  33.962
ATOM     14  C   LGD     1      11.648  36.145  33.834
ATOM     15  C   LGD     1      12.873  36.525  34.406
ATOM     16  H   LGD     1      13.446  37.353  34.002
ATOM     17  N   LGD     1      11.133  36.972  32.845
ATOM     18  O   LGD     1      11.152  38.186  32.818
ATOM     19  O   LGD     1      10.491  36.289  32.063
```

(b) Output in mol2 format (when COORDI= MOL2)

The overall score and result for each item appears at the top of the file, followed by output of the ligand coordinates. Ligand information appears in the <MOLECULE>, <ATOM> and <BOND> sections.

```
#                SCORE =  -212.1914
#             dG-SCORE =    -6.4866
#  HIT-OPTIMIZED SCORE =  -130.5660
#            MTS SCORE =  -336.8068
#                  ASA =  -180.8419
#                  ELE =    -0.3130
#                  HYD =   -29.5172
#                  VDW =    -1.5193
#                 RMSD =     1.3238
@<TRIPOS>MOLECULE
conf1.mol2
   19   19    0    0    0
SMALL
NO_CHARGES

@<TRIPOS>ATOM
    1 H         14.3799   36.6888   35.0812 H         0 <1>          0.0018
    2 C         13.8677   36.8677   36.0378 C.3       0 <1>         -0.0913
    3 H         14.3697   36.3067   36.8405 H         0 <1>          0.0018
    4 C         13.9674   38.3559   36.3703 C.2       0 <1>          0.9634
    5 O         14.8891   38.6784   37.1647 O.co2     0 <1>         -0.8700
    6 O         13.2799   39.3414   35.9877 O.co2     0 <1>         -0.8700
    7 C         12.4932   36.3048   35.9160 C.ar      0 <1>          0.0677
    8 C         11.7718   36.7030   34.7716 C.ar      0 <1>         -0.1494
    9 H         12.2055   37.3845   34.0453 H         0 <1>          0.2269
   10 C         10.4658   36.1886   34.6092 C.ar      0 <1>          0.0189
   11 N          9.7449   36.5654   33.4687 N.pl3     0 <1>          0.7824
   12 O          8.6385   36.0379   33.4552 O.co2     0 <1>         -0.4940
   13 O         10.1144   37.3023   32.5800 O.co2     0 <1>         -0.4940
   14 C          9.8654   35.2992   35.5589 C.ar      0 <1>         -0.1982
   15 H          8.8570   34.9423   35.3953 H         0 <1>          0.1532
   16 C         10.5993   34.9114   36.6710 C.ar      0 <1>         -0.1906
   17 H         10.1664   34.2205   37.3834 H         0 <1>          0.1429
   18 C         11.9162   35.4213   36.8667 C.ar      0 <1>         -0.1545
   19 H         12.4792   35.1319   37.7468 H         0 <1>          0.1531
```

(Continued on next page)

(Continued from previous page)

```
@<TRIPOS>BOND
     1      1      2    1
     2      2      3    1
     3      2      4    1
     4      2      7    1
     5      4      6    1
     6      4      5    2
     7      7     18    ar
     8      7      8    ar
     9     18     19    1
    10     16     18    ar
    11     16     17    1
    12     14     16    ar
    13     14     15    1
    14     10     14    ar
    15      8     10    ar
    16     10     11    1
    17      8      9    1
    18     11     13    1
    19     11     12    2
```

（2）Method for separately performing generation of the grid potential and docking

When docking the target receptor with multiple compounds, it is efficient to generate the grid potential of the target receptor and store it in a file, and perform the compound docking calculations using the previously generated grid potential.

【Only generating the grid potential and storing it in a file】

The specifications in the INPUT phase and GRID phase are the same as in (1), and the specifications in the CONF, DOCK, MIN, and OUTPUT phases are not needed. Specify the file name and file format of the grid potential to be output.

Execution of sievgene：

"grid_make.inp" is specified for the input file.

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

【Note】When only generating the grid potential (no CONF/DOCK/MIN/OUTPUT phase specifications) , "EXE> SIEV" is not needed in the input file.
If "EXE> SIEV" is specified, an error message similar to the following will appear after output of the grid potential.

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


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

■Input file (grid_make.inp)

```
PHASE> INPUT
  TOPOLO  =    FORM
  NAMETO  =    Pro.tpl     ; Protein topology
  COORDI  =    PDB
  NAMECO  =    Pro_md.pdb  ; Protein coordinates
  POINTC  =    PDB
  NAMEPO  =    point.pdb   ; Probe points indicating pocket region

  SETTAR = NORE
  DAMPPA = 1.0d0
  QUIT
;
; Grid generation and hash table generation
;
PHASE> GRID
  GRIDPOtential    = NORE        ; Grid potential input
  NAMEGRid         = grid.file   ; Grid potential file name
  OUTGRIdpotential = BINA        ; Grid potential output
  PROBDIst = 6.5  ; Target receptor atoms within PROBDIST of probe points
  MARGIN   = 6.5  ; Margin around search pocket for grid generation
  ITERAT   = 3    ; Number of repetitions of grid  potential smoothing
  RADVDW   = 0.6  ; Set boundary between inside and outside regions of vDW potential
  RADELE    = 0.6   ; Set boundary between inside and outside regions of coulomb
potential
  RADMESh  = 1.4  ; Probe radius for mesh node generation

  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
```

【Docking only using a previously generated grid potential】

General specifications are the same as in (1). The input grid file name (NAMEGR) and grid file format (GRIDPO) are specified in the GRID phase. The grid potential is not output, and thus OUTGRI=NOWR is specified.

Execution of sievgene :

"restart_run.inp" is specified for the input file.

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

The output files are the same as in (1).

■Input file (restart_run.inp)

```
PHASE> INPUT
  LIGAND  =    MOL2
  NAMELI  =    Lig_es_1.mol2 ; Coordinate position file of ligand
  REFERE  =    MOL2
  NAMERE  =    lig_ref.mol2  ; Reference position file of ligand
  TOPOLO  =    FORM
  NAMETO  =    Pro.tpl  ; Protein topology
  COORDI  =    PDB
  NAMECO  =    Pro_md.pdb  ; Protein coordinates
  POINTC  =    PDB
  NAMEPO  =    point.pdb   ; Probe points indicating pocket region

  SETTAR = NORE
  DAMPPA = 1.0d0
  QUIT
;
; Grid generation and hash table generation
;
PHASE> GRID
  GRIDPOtential    = BINA        ; Grid potential input
  NAMEGRid         = grid_file   ; Grid potential file name
  OUTGRIdpotential = NOWR        ; Grid potential output
  PROBDIst = 6.5  ; Target receptor atoms within PROBDIST from probe point
  MARGIN   = 6.5  ; Margin around search pocket for grid generation
  ITERAT   = 3    ; Number of repetitions of grid  potential smoothing
  RADVDW   = 0.6  ; Set boundary between inside and outside regions of vDW potential
  RADELE    = 0.6  ; Set boundary between inside and outside regions of coulomb
potential
  RADMESh  = 1.4  ; Probe radius for mesh node generation

  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
```

(Continued on next page)

(Continued from previous page)

```
PHASE> CONF
  CONFLImit         = 100000  ; Maximum number of conformer generation trials
  CONFORmernumber = 100   ; Number of generated conformers (changing the ligand conformation)
  SORTATom          = YES     ; Change atom order in output file
  DAMPINg        = 0.7      ; Damping factor for distance between a toms of conformer candidates
  PHASETorsion     = 3       ; Number of torsion rotation candidates for conformer generation
  ROTTER            = YES
  QUIT
PHASE> DOCK
  METHOD        = FLEX
  GENERAtion    = 1    ; Number of generations for narrowing search
  NUMCONFomer  = 1000 ; Number of bonding surface triangles of ligand
  MATCHING  =  4       ; Matching of triangle atom types
  LOWMIN     =  2.5    ; Lower minimum of distance between atoms compared with hash table
  LOWMAX     =  3.5    ; Lower maximum of distance between atoms compared with hash table
  UPRMIN     =  5.0    ; Upper minimum of distance between atoms compared with hash table
  UPRMAX     =  12.0   ; Upper maximum of distance between atoms compared with hash table

  WETVDW= 1.0  ; Weight of each item when calculating overall score
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  RADIUS= 6.0  ; Count the number of ligand atoms existing within RADIUS from target
               ; atom (coordinates) after ligand molecule docking
  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       ; Output format of ligand coordinate file
  NAMECOordinate = ex.cor     ; Ligand coordinate file name in each experiment
  CANDIDatenumber= 10          ; Number of high scores for ligand coordinate outputin
    each experiment
   SCORENumber     = 10          ; Number ofhigh scores output in score file
   NAMESCore       = ex.score    ; Score file name
   QUIT

EXE> SIEV
```

2.2 Option combinations

In sievgene, the combination of options specified in the input file has a large effect on execution time and docking precision.

There are many different types of combinations. Three typical examples are explained below.

（1）Fast version　　：Calculation method for shortest execution time

（2）Moderate version　：Calculation method for both short execution time and high docking accuracy

（3）Precise version　　：Calculation method for highest precision

By means of the option combinations shown in the table below, these three examples enable calculation to match the objective.

Phase	Option	Specification method		
		（1） Fast version	（2） Moderate version	（3） Precise version
CONF phase	ATMMDL	UNIT	ALL	ALL
DOCK phase	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

This methods enable docking calculation in a very short time with a slightly lower docking precision than other methods.

To give priority to execution time, the united atom model is used for the atom model (ATMMDL= UNIT) and a calculation method is used that excludes atoms that are not targets of superimposition (DOCKSP= FAST).

The local search execution time is also shortened by setting the local search candidates to 10 (CANDID= 10).

In addition, the atom type matching of the bonding surface is set to 3 (MATCH=3) and the side length range of the bonding surface is narrowedfrom the default (LOWMIN= 2.5, LOWMAX= 3.5, UPMIN= 5.0) to shorten execution time.

■Input file

```
PHASE> INPUT
  LIGAND  =    MOL2
  NAMELI  =    Lig_es_1.mol2 ; Ligand coordinate position file
  REFERE  =    MOL2
  NAMERE  =    lig_ref.mol2  ; Reference position file of ligand
  TOPOLO  =    FORM
  NAMETO  =    Pro.tpl       ; Protein topology
  COORDI  =    PDB
  NAMECO  =    Pro_md.pdb    ; Protein coordinates
  POINTC  =    PDB
  NAMEPO  =    point.pdb     ; Probe point indicating pocket region

  ASAMET  =    PAIR          ; ASA calculation method
  SETTAR  =    NORE
  DAMPPA  =    1.0d0
  QUIT
;
; Grid generation and hash table generation
;
PHASE> GRID
  GRIDPOtential    = BINA        ; Grid potential input
  NAMEGRid         = grid_file   ; Grid potential file name
  OUTGRIdpotential = NOWR        ; Grid potential output
  PROBDIst = 6.5  ; Target receptor atoms within PROBDIST from probe point
  MARGIN   = 6.5  ; Margin around search pocket during grid generation
  ITERAT   = 3    ; Number of repetitions of grid potential smoothing
  RADVDW   = 0.6  ; Set boundary between inside and outside regions of vDW potential
  RADELE = 0.6 ; Set boundary between inside and outside regions of coulomb potential
  RADMESh  = 1.4  ; Probe radius for mesh node generation

  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
```

(Continued on next page)

(Continued from previous page)

```
PHASE> CONF
  ATMMDL           = UNIT      ; Specify atom model
  CONFLImit        = 100000    ; Maximum number of conformer generation trials
  CONFORmernumber  = 100       ; Number of generated conformers (transforming the ligand)
  SORTATom         = YES       ; Change atom order in output file
  DAMPINg        = 0.7    ; Damping factor for distance between atoms in conformer candidates
  PHASETorsion     = 3       ; Number of torsion rotation candidates for conformer generation
  ROTTER           = YES
  QUIT
PHASE> DOCK
  METHOD       = FLEX ; Docking method ( rigid | flexible )
  DOCKSP       = FAST ; Docking method ( normal | fast )
  CANDID       = 10   ; Number of local search candidates
  GENERAtion   = 1    ; Number of generations for narrowing search
  NUMCONFomer  = 1000 ; Number of bonding surface triangles of ligand
  MATCHING  =  3      ; Matching of triangle atom types
  LOWMIN    =  2.5    ; Lower minimum of distance between atoms compared with hash table
  LOWMAX    =  3.5    ; Lower maximum of distance between atoms compared with hash table
  UPRMIN    =  5.0    ; Upper minimum of distance between atoms compared with hash table
  UPRMAX    =  12.0   ; Upper maximum of distance between atoms compared with hash table

 WETVDW= 1.0  ; Weight of each item when calculating overall score
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  WETANH= 1.0
  RADIUS= 6.0  ; Count the number of ligand atoms existing within RADIUS from target
               ; atom (coordinates) after ligand molecule docking
  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       ; Output format of ligand coordinate file
  NAMECOordinate = ex.cor     ; Ligand coordinate file name in each experiment
  CANDIDatenumber= 3; Number of high scores for which ligand conformer is output in each
                              experiment.
  SCORENumber    = 3          ; Number of high scores output in score file
  NAMESCore      = ex.score   ; Score file name
  QUIT
EXE> SIEV
```

(2) Moderate version

This calculation method provides both high docking precision and fast execution.

To maintain docking precision, "all atom model" is used for the atom model (ATMMDL= ALL) and a calculation method is used whereby atoms that are not targets of superimposition when the protein and ligand are superimposed are not excluded (DOCKSP= NORM) .

Specifying 100 for the number of local search candidates (CANDID= 100) also maintains docking precision.

On the other hand, 2 is specified for the atom type matching of the bonding surface (MATCH= 2) and the side length range of the bonding surface is narrowed from the default (LOWMIN= 2.5, LOWMAX= 3.5, UPMIN= 5.0) to shorten the execution time.

■Input file

```
PHASE> INPUT
  LIGAND  =    MOL2
  NAMELI  =    Lig_es_1.mol2 ; Coordinate position file of ligand
  REFERE  =    MOL2
  NAMERE  =    lig_ref.mol2  ; Reference position file of ligand
  TOPOLO  =    FORM
  NAMETO  =    Pro.tpl       ; Protein topology
  COORDI  =    PDB
  NAMECO  =    Pro_md.pdb    ; Protein coordinates
  POINTC  =    PDB
  NAMEPO  =    point.pdb     ; Probe points indicating pocket region

  ASAMET  =    PAIR          ; ASA calculation method
  SETTAR  =    NORE
  DAMPPA  =    1.0d0
  QUIT
;
; Grid generation and hash table generation
;
PHASE> GRID
  GRIDPOtential    = BINA        ; Grid potential input
  NAMEGRid         = grid_file   ; Grid potential file name
  OUTGRIdpotential = NOWR        ; Grid potential output
  PROBDIst = 6.5  ; Target receptor atoms within PROBDIST from probe points
  MARGIN   = 6.5  ; Margin around search pocket during grid generation
  ITERAT   = 3    ; Number of repetitions of grid potential smoothing
  RADVDW   = 0.6  ; Set boundary between inside and outside regions of vDW potential
  RADELE   = 0.6  ; Set boundary between inside and outside regions of coulomb potential
  RADMESh  = 1.4  ; Probe radius for mesh node generation

  DAMPVW = 0.99d0

  USEPBG = NO        ; PB not used
```

(Continued on next page)

(Continued from previous page)

```
PHASE> CONF
  ATMMDL           = ALL       ; Specify atom model
  CONFLImit        = 100000    ; Maximum number of conformer generation trials
  CONFORmernumber  = 100     ; Number of generated conformers (changing conformation of ligand)
  SORTATom         = YES       ; Change order of atoms in output file
  DAMPINg         = 0.7     ; Damping factor for distance between atoms of conformer candidates
  PHASETorsion     = 3         ; Number of torsion rotation candidates for conformer generation
  ROTTER           = YES
  QUIT
PHASE> DOCK
  METHOD       = FLEX ; Docking method ( rigid | flexible )
  DOCKSP       = NORM ; Docking method ( normal | fast )
  CANDID       = 100  ; Number of local search candidates
  GENERAtion   = 1    ; Number of generations for narrowing search
  NUMCONFomer  = 1000 ; Number of bonding surface triangles of ligand
  MATCHING  =  2      ; Matching of triangle atom types
  LOWMIN    =  2.5    ; Lower minimum of distance between atoms compared with hash table
  LOWMAX    =  3.5    ; Lower maximum of distance between atoms compared with hash table
  UPRMIN    =  5.0    ; Upper minimum of distance between atoms compared with hash table
  UPRMAX    =  12.0   ; Upper maximum of distance between atoms compared with hash table

  WETVDW= 1.0  ; Weight of each item when calculating overall score
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  WETANH= 1.0
  RADIUS= 6.0  ; Count the number of ligand atoms existing within RADIUS from target
               ; atom (coordinates) after ligand molecule docking
  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       ; Output format of ligand coordinate file
  NAMECOordinate = ex.cor     ; Ligand coordinate  file name in each experiment
  CANDIDatenumber= 3         ; Number of high scores for which ligand coordinates are output
in each experiment
  SCORENumber    = 3          ; Number of high scores output in score file
  NAMESCore      = ex.score   ; Score file name
   QUIT
EXE> SIEV
```

(3) Precise version

This method requires more time than the other methods, however, it enables very high precision docking calculation.

To maintain docking precision, "all atom model" is used for the atom model (ATMMDL= ALL) and a calculation method is used in which atoms that are not targets of super imposition when the protein and ligand are superimposed are not excluded (DOCKSP= NORM) .

Specifying 100 for the number of local search candidates (CANDID= 100) also maintains docking precision.

In addition, 0 is specified for the atom type matching of the bonding surface (MATCH= 0) and the side length range of the bonding surface is widened from the default (LOWMIN= 1.0, LOWMAX= 1.2, UPMIN= 8.0) to increase precision.

■Input file

```
PHASE> INPUT
  LIGAND  =    MOL2
  NAMELI  =    Lig_es_1.mol2 ; Coordinate position file of ligand
  REFERE  =    MOL2
  NAMERE  =    lig_ref.mol2  ; Reference position file of ligand
  TOPOLO  =    FORM
  NAMETO  =    Pro.tpl       ; Protein topology
  COORDI  =    PDB
  NAMECO  =    Pro_md.pdb    ; Protein coordinates
  POINTC  =    PDB
  NAMEPO  =    point.pdb     ; Probe point indicating  pocket region

  ASAMET  =    PAIR          ; ASA calculation method
  SETTAR  =    NORE
  DAMPPA  =    1.0d0
  QUIT
;
; Grid generation and hash table generation
;
PHASE> GRID
  GRIDPOtential    = BINA        ; Grid potential input
  NAMEGRid         = grid_file   ; Grid potential file name
  OUTGRIdpotential = NOWR        ; Grid potential output
  PROBDIst = 6.5  ; Target receptor atoms within PROBDIST from probe points.
  MARGIN   = 6.5  ; Margin around search pocket during grid generation
  ITERAT   = 3    ; Number of repetitions of grid potential smoothing
  RADVDW   = 0.6  ; Set boundary between inside and outside regions of vDW potential
  RADELE   = 0.6  ; Set boundary between inside and outside regions of coulomb potential
  RADMESh  = 1.4  ; Probe radius for mesh node generation
  DAMPVW = 0.99d0

  USEPBG = NO         ; no use of PB
```

(Continued on next page)

(Continued from previous page)

```
PHASE> CONF
  ATMMDL           = ALL       ; Specify atom model
  CONFLImit        = 100000    ; Maximum number of conformer generation trials
  CONFORmernumber  = 100        ; Number of generated conformers (changing conformation of
ligand)
  SORTATom         = YES       ; Change order of atoms in output file
  DAMPINg           = 0.7        ; Damping factor for distance between atoms of conformer
candidates
  PHASETorsion      = 3            ; Number of torsion rotation candidates for conformer
generation
  ROTTER           = YES
  QUIT
PHASE> DOCK
  METHOD       = FLEX ; Docking method ( rigid | flexible )
  DOCKSP       = NORM ; Docking method ( normal | fast )
  CANDID       = 100  ; Number of local search candidates
  GENERAtion   = 1    ; Number of generations for narrowing search
  NUMCONFomer  = 1000 ; Number of bonding surface triangles of ligand
  MATCHING  =  0      ; Matching of triangle atom types
  LOWMIN    =  1.0    ; Lower minimum of distance between atoms compared with hash table
  LOWMAX    =  1.2    ; Lower maximum of distance between atoms compared with hash table
  UPRMIN    =  8.0    ; Upper minimum of distance between atoms compared with hash table
  UPRMAX    =  12.0   ; Upper maximum of distance between atoms compared with hash table

  WETVDW= 1.0  ; Weight of each item when calculating overall score
  WETASA= 1.0
  WETELE= 1.0
  WETHYD= 1.0
  WETANH= 1.0
  RADIUS= 6.0  ; Count the number of ligand atoms existing within RADIUS from target
               ; atom (coordinates) after ligand molecule docking.
  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       ; Output format of ligand coordinate file
  NAMECOordinate = ex.cor     ; Ligand coordinate file name in each experiment
  CANDIDatenumber= 3         ; Number of high scores for which ligand coordinates are output
```

A Input/Output Files

A.1 Input/output files of docking engine

File input/output is performed by the docking engine for the following purposes.

(1) Specification of simulation conditions

(2) Storing of simulation states

(3) Output of simulation results

These files are generally referred to as input/output files.

A.1.1 Explanation of phases

The docking engine performs file input/output in the following phases:

(1)INPUT phase

Input of files containing system topology, coordinates, simulation conditions, and other information.

(2) OUTPUT phase

Output of files containing score information after simulation and ligand conformer coordinates.

(3) GRID phase

Generation of grid potential of protein pocket region.

(4) CONF phase

Generation of ligand conformers.

(5) DOCK phase

System docking simulation.

(6) MIN phase

Minimization of the potential energy of the system.

A.2 Input files

The docking engine input files are shown below.

Order	File name	Applicable phase	Purpose
#1	Control file	-	Docking engine control
#2	Topology file	All phases	Specification of protein topology
#3	Coordinate file	All phases	Specification of coordinates of protein atoms and pocket point
#4	Restart file	All phases	Specification of restart information
#5	Search candidate file	All phases	Specification of protein atoms that will be search candidates
#6	Grid potential file	GRID	Specification of grid potential
#7	PB-Grid potential file	GRID	Specification of Poisson-Boltzmann grid potential

A.2.1 Control file

Applicable phase : All phases

Purpose: Controls execution of phases of the torsion search engine and specifies parameters in each phase.

Special notes :

The character that indicates a real exponent must be a "D".

```
Example of specifying a real exponent:
CPUTIM = 60.0D0
```

Format :

```
[Line specifying execution of phase  [Line specifying parameters in phase]...
```

Lines specifying execution of phases: Executed phases are specified with the following text strings.

```
INPUT phase    =  " PHASE> INPUT "
OUTPUT phase   =  " PHASE> OUTPUT "
GRID phase     =  " PHASE> CONF "
CONF phase     =  " PHASE> OUTPUT "
MINimize phase  =        " EXE> MIN "
Execution end line  =            " EXE> SIEV "
```

Parameter end line : The end of the parameter execution line in a phase is specified in the following format.

" QUIT "

Specification of phase parameters : Parameters of each phase are expressed in the following format.

```
Keyword " = " value
```

※The keyword must consist of 6 English letters, and the value specified for the keyword is one of 4 types: selection, real number, integer, or text.

```
Keyword specification example:
UNITAN = 30              ;   Integer parameter
NAMEAN = aa.ana          ;   Text parameter
BINCLO = NO              ;   Selection parameter
CPUTIM = 60.0D0          ;   Real number parameter
```

The keywords and values of each phase are shown below. The effective forms are those in capital English letters.
The "Description" column shows the following items depending on the format indicated in the "Value" column.

・Selection : The text strings in capital letters are the selections, and the underlined string is the default value.

・Integer, real number, text : Default value

A.2.1.1 INPUT phase

Item #	Item	Keyword	Value	Description
#1	Protein topology	**TOPOLO**	Selection	File read-in and format (**NORE** | **FORM** | **BINA**)
#2		**NAMETO**	Text	File name ("")
#3	Protein atom coordinates	**COORDI**	Selection	File read-in and format (**PDB** | **BINA**)
#4		**NAMECO**	Text	File name ("")
#5	Ligand coordinates	**LIGAND**	Selection	File read-in and format (**PDBX** | **MOL2**)
#6		**NAMELI**	Text	File name ("")
#7	Ligand reference coordinates	**REFERE**	Selection	File read-in and format (**PDBX** | **MOL2**)
#8		**NAMERE**	Text	File name ("")
#9	Protein Pocket point coordinates	**POINTC**	Selection	File read-in (**PDB** | **BINA**)
#10		**NAMEPO**	Text	File name ("")

#11	ASA specification	**ASAMET**	Selection	ASA calculation method (**PAIR** | **RICH**)
#12	Search candidate specification	**SETTAR**	Selection	File read-in and format (**NORE** | **READ**)
#13		**NAMETA**	Text	File name ("")
#14	Protein atom information	**DAMPPA**	Real number	Damping factor of protein atom radius (1.0)

A.2.1.2 GRID phase

Item #	Item	Keywork	Value	Description
#1	Grid-Potential input/output	**GRIDPO**	Selection	File read-in and format (**NORE** | **ASCI** | **BINA**)
#2		**OUTGRI**	Selection	File write and format (**NOWR** | **ASCI** | **BINA**)
#3		**NAMEGR**	Text	File name ("")
#4	Receptor registration range	**PROBDI**	Real number	Limit of distance between pocket point and receptor atom (6.5)
#5	Grid-Potential generation	**RADVDW**	Real number	VdW boundary distance correction (0.6)
#6		**RADELE**	Real number	Electrostatic boundary distance correction (0.6)
#7		**DANPVW**	Real number	Damping factor of vdW radius (1.0)
#8	Mesh generation	**RADMES**	Real number	Probe radius for mesh node generation (1.6)
#9	Grid-Potential range correction	**MERGIN**	Real number	Margin of grid potential range (0.0)
#10	Number of smoothings	**ITERAT**	Integer	Number of smoothing repetitions (3)
#11	PB-Grid potential	**USEPBG**	Selection	Use Poisson Boltzmann equation (**NO** | **YES**)
#12		**PBGRID**	Selection	File read-in and format (**NORE** | **ASCI** | **BINA**)

#13		OUTPBG	Selection	File write and format (NOWR | ASCI | BINA)
#14		NAMEPB	Text	File name ("")
#15		NMESHX	Integer	Mesh number (x direction) (100)
#16		NMESHY	Integer	Mesh number (y direction) (100)
#17		NMESHZ	Integer	Mesh number (z direction) (100)
#18		MARGPB	Real number	Margin of PB-Grid potential (5.0 d0)
#19		KAPPAV	Real number	Debye-Huckel shield constant (0.0)
#20		CPULIM	Real number	CPU time limit (10800)

A.2.1.3 CONF phase

Item #	Item	Keyword	Value	Description
#1	Conformer generation	CONFLI	Integer	Maximum number of conformer generation trials (10000000)
#2		CONFOR	Integer	Number of conformers generated (100)
#3	Atom data sorting	SORTAT	Selection	Sort atom data (NO | YES)
#4	Lower damping factor for vdW distance between atoms	DAMPIN	Real number	Damping factor of vdW distance between conformer atoms (0.7)

#5	Number of torsion rotations	**PHASET**	Integer	Number of torsion rotation candidates (6)
#6	Atom model	**ATMMDL**	Selection	Atom model specification (**ALL** | **UNIT**)

A.2.1.4 DOCK phase

Item #	Item	Keyword	Value	Description
#1	Docking method	METHOD	Selection	Docking method (FLEX | RIGI)
#2		GENERA	Integer	Number of search narrowing generations (5)
#3		NUMCON	Real number	Number of high scores shown (30)
#4		MATCHI	Integer	Atom type matching of bonding surface (5)
#5		LOWMIN	Real number	Lower minimum of bonding surface side (2.5)
#6		LOWMAX	Real number	Lower maximum of bonding surface side (4.0)
#7		UPRMIN	Real number	Upper minimum of bonding surface side (7.5)
#8		UPRMAX	Real number	Upper maximum of bonding surface side (10.0)
#9		INTERP	Selection	Grid potential interpolation method (LAGR | BSPL)
#10		EVALHB	Selection	Evaluation of hydrogen bonding of protein and ligand takinganisotropy into consideration. (YES | NO)
#11		ROTLOH	Selection	Rotation of ligand-OH radical (YES | NO)
#12		ROTPSC	Selection	Rotation of hydrogen-bondable protein side chain. (YES | NO)
#13		MOVNUM	Integer	Number of times coordinates are moved (10)
#14		CANDID	Integer	Number of local search candidates (30)
#15		DOCKSP	Integer	Switching of docking atoms (NORM | FAST)

#16		OPTIME	Integer	Specify whether or not there is docking calculation. (OPT | ENE)
#17		CPUTIME	Real number	Time limit for docking calculation (second) (30.0)
#18	Score calculation	WETVDW	Real number	vdW coefficient for score calculation (1.0)
#19		WETASA	Real number	ASA coefficient for score calculation (1.0)
#20		WETELE	Real number	Electrostatic coefficient for score calculation (1.0)
#21		WETHYD	Real number	Hydrogen bonding coefficient for score calculation (1.0)
#22		WETANH	Real number	Coefficient of protein and ligand hydrogen bonding taking anisotropy into consideration (1.0)
#23		RADIUS	Real number	Bonding atom count distance limit (6.0)
#24	Pocket information	POCKCX	Real number	X coordinate of pocket center (999.9)
#25		POCKCY	Real number	Y coordinate of pocket center (999.9)
#26		POCKCZ	Real number	Z coordinate of pocket center (999.9)
#27		POCKET	Integer	Pocket center atom ID (0)

A.2.1.5 MIN phase

See "EXE> MIN GROUP" in the "cosgnene" chapter of the User Manual.

A.2.1.6 OUTPUT phase

Item	Item	Keyword	Value	Description

#				
#1	Output of high-score conformer coordinate file	COORDI	Selection	File output and format (NOWR | MOL2 | PDBX | PDB)
#2		CANDID	Integer	Number of high score conformers (30)
#3		NAMECO	Text	File name (" ")
#4	Output of high scores	SCOREN	Integer	Number of high scores output (30)
#5		NAMESC	Text	High score file name (" ")

A.3 Output files

Files output by the docking engine are shown below.

Item #	File name	Output phase	Purpose
#1	High-score conformer coordinate file	OUTPUT	Output of atom information and information on scores of high-score conformers in final results.
#2	High score file	OUTPUT	Output of geometric hash docking result, docking result after energy minimization, and final result.

(Blank)

B Utilities

B.1 make_point

The position of the receptor pocket is specified by a collection of probe points and given in PDB format. If you have no previously known ligand information, make_point.f can be used to create a collection of probe points in PDB format.

Input data

(1) Name of receptor file
(2) Radius of the spherical region that generates the probe points
(3) Number of the atom that will be the center

■Example of use

From standard input, specify in succession the file name of the receptor, the radius of the spherical region that generates the probe points, and the number of the atom that will be at the center. The collection of probe points will be output in PBD format with the fixed name "point.pdb".

```
% make_point.exe
```

【Note】The file format (POINTC= PDB) and file name (NAMEPO= point.pdb) are specified as input information of sievgene.

【Note】It is useful to edit the created "point.pdb" file with an editor to make it more closely resemble the actual state.

***myPresto* 4.2**

- sievgene -