

myPresto 5.0

- *tplgeneL* -

USER MANUAL

2018/1/12

本ドキュメントについて

本ドキュメントは、「*myPresto 5.0* 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 5.0* 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、国立研究開発法人日本医療研究開発機構 (AMED) 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・京極好正博士の始められた研究の中で開発されました。

目次

1.	tplgeneL の概要	4
2.	実行方法	4
3.	入力データの作成.....	7
3.1.	tplgeneL オリジナル形式ファイル	7
3.2.	Sybyl mol2 形式ファイル	9
4.	原子タイプ定義ファイル.....	15
5.	カ場パラメータデータベースファイル.....	15
6.	フラグメントデータベースファイル.....	17
7.	環境変数	18

1. tplgeneL の概要

tplgeneL は、Sybyl mol2 ファイル形式ないし、独自ファイル形式での低分子化合物を読み込み、当該低分子化合物のトポロジーファイルと PDB 形式の原子座標ファイルを出力します。

tplgeneL は、myPresto v.5.0 では、sievgene_pack、もしくは、cosgene_pack に含まれています。本マニュアルでは、それらに含まれている tplgeneL の使用方法について説明します。

2. 実行方法

計算を行う際に参照するディレクトリ（入力ファイル用、出力ファイル用、力場 DB 用）は、環境変数に設定することができます。環境変数を設定していない場合は、カレントディレクトリを参照しますので、入力ファイル、原子タイプ定義ファイル、力場パラメータ DB ファイルを、予め実行するディレクトリにコピーしておきます。

tplgeneL の実行の際は、以下の項目を指定します。これらは、画面からの対話入力、又は、コマンドライン・オプションで指定することにより計算を行います。

■入力項目

- 入力ファイルの書式（1：tplgeneL オリジナル形式、2：Sybyl mol2 形式）
- 入力ファイル名
- 不足パラメータが存在する場合の処理方法
(1：デフォルトパラメータを使用する、2：動的にパラメータの計算を行う
3：デフォルトパラメータの使用するが、パラメータのないものは動的にパラメータの計算を行う)
- パラメータ DB ファイル名
- フラグメント DB を使用するか (yes：使用する、no：使用しない)

```
% tplgeneL  
または、  
% tplgeneL (オプション)
```

尚、sievgene_pack、および、cosgene_pack には、exec_tplgeneL.sh が含まれています。これは、スクリプトの中で環境変数を設定して、tplgeneL を起動するプログラムです。（オプションを与えた実行はできません。）

コマンドライン・オプションを用いて指定した項目は、対話入力での入力がスキップされます。オプションで指定しなかったものだけを対話的に入力することになります。

```
-ft [ 1 | 2 ]
    入力ファイル形式を指定します
        tplgeneL オリジナル入力ファイル    : 1
        Sybyl mol2 入力ファイル            : 2

-i <file>
    入力ファイル名を<file>に指定します

-d <db_file>
    パラメータ DB ファイル名を<db_file>に指定します

-r <resname>
    出力するトポロジーファイルに記述する残基名を
    <resname>に指定します(4 文字以内)

-f [ yes | no ]
    フラグメント DB を使用するか否かを指定します
        使用する    ⇒ yes あるいは y
        使用しない  ⇒ no  あるいは n

-p [ 1 | 2 ]
    不足パラメータの補填方法を選択します
        デフォルトパラメータ                : 1
        動的補填                            : 2
        デフォルトパラメータ+動的補填        : 3
```

■ オプション指定例（下線部が入力部分）（tplgeneL へのパスが設定されている場合）

```
% tplgeneL -i methanol -d prm_gaff.db -f no
```

また、オプション“-h”または、“-help”の指定により、tplgeneL の使用方法を見ることができます。

```
% tplgeneL -h
    または、
% tplgeneL -help
```

なお、予め対話入力する項目（作業指示）をファイル（作業指示ファイル:control_file）

に保存しておくことによって、`tplgeneL` を実行する度に対話入力をする手間を省略することができます。

```
% tplgeneL < control_file
```

■ 作業指示ファイル例

1	: 入力ファイルフォーマットを記述(1:original、2:mol2)。
methanol	: ファイル名を記述する(拡張子は除く)。
1	: 不足パラメータの補填方法を記述(1:default、2:動的補填)。
prm_gaff.db	: パラメータ DB ファイル名を記述する。
no	: フラグメント DB を使用するかを記述する(yes/no)。

`tplgene` で得られた高分子のトポロジーファイルと `tplgeneL` で得られたトポロジーファイルを結合することにより、`cosgene` にて、高分子-低分子複合体の MD シミュレーションを行う事ができます。

対象分子の構造に関するデータ、使用する力場、出力ファイルのディレクトリをそれぞれ分けたい場合は、それらのパスを環境変数で指定することができます。環境変数が設定されていない場合は、実行時のカレントディレクトリを参照します(「7. 環境変数」参照)。

3. 入力データの作成

tplgeneL で使用する分子の情報に関する入力ファイルは、tplgeneL オリジナル形式ファイル(bond ファイル、charge ファイル、zmat ファイル) と Sybyl mol2 形式ファイル(mol2 ファイル)の2種類に対応しています。

3.1. tplgeneL オリジナル形式ファイル

tplgeneL オリジナル入力ファイルを使用する場合には、以下の3種類のファイルを使用します。

(1) 電荷情報ファイル(XXX.charge ファイル(ここでXXXはファイル名(拡張子を除く)を示す)原子名(項目1)、元素記号(同2)、Mulliken 電荷情報(同3)、Resp 電荷情報(同4)が記述されています。

(2) 結合情報ファイル(XXX.bond ファイル)

結合している原子の番号の組(項目1、2)、結合長(同3)、結合次数(同4)が記述されています。

(3) 座標情報ファイル(XXX.zmat ファイル)

入力分子の Z-matrix 情報が記述されています。

上記1、2のファイルは必須です。3のファイルは存在する場合には、その情報がトポロジーファイルに反映されます。

■電荷情報ファイルの例

C1	C	-0.1320	-0.1320
O2	O	-0.7323	-0.7323
H3	H	0.1340	0.1340
H4	H	0.1653	0.1653
H5	H	0.1644	0.1644
H6	H	0.4006	0.4006

■結合情報ファイルの例

1	2	1.4130	0.7600
1	3	1.1160	0.9530
1	4	1.1200	0.9480
1	5	1.1200	0.9480
2	6	0.9630	0.7970

■座標情報ファイルの例

C								
O	1	1.4132350						
H	1	1.1159340	2	112.6746				
H	1	1.1195330	2	107.3658	3	121.0117	0	
H	1	1.1196880	2	107.4483	3	-121.0171	0	
H	2	.9627370	1	107.7002	5	-120.0241	0	

3.2. Sybyl mol2 形式ファイル

tplgeneL では tplgeneL オリジナル形式ファイル以外にも、Sybyl mol2 形式ファイルも使用することができます。

Sybyl mol2 ファイルには、分子情報(@<TRIPOS>MOLECULE 情報)、原子情報(@<TRIPOS>ATOM 情報)、結合情報(@<TRIPOS>BOND 情報)等が存在しますが、tplgeneL ではその内の原子情報、結合情報を使用します。原子情報、結合情報には以下の情報が含まれています。

(1) 原子情報(@<TRIPOS>ATOM 情報)

- 原子 ID : 1 からの通し番号。
- 原子名 : 原子名。1、2 文字目は元素記号です。
- 座標 : カーテシアン座標での座標。
- 原子タイプ : Sybyl の原子タイプ。
- サブ構造 ID : その原子を含むサブ構造の ID。

<tplgeneL では使用しません。>

- サブ構造名 : その原子を含むサブ構造名。

<tplgeneL では使用しません。>

- 電荷 : 各原子の電荷情報。
- status bit : Sybyl 固有のステータス情報。

<tplgeneL では使用しません。>

(2) 結合情報(@<TRIPOS>BOND 情報)

- 結合 ID : 1 からの通し番号。
- 原子 ID 1 : 結合している原子 1 の番号 (上記の原子情報の原子 ID と一致)。
- 原子 ID 2 : 結合している原子 2 の番号 (上記の原子情報の原子 ID と一致)。
- 結合タイプ : 結合のタイプ(1、2、3、am、ar、du、un、nc)。
- status bit : Sybyl 固有のステータス情報。

<tplgeneL では使用しません。>

■ Sybyl mol2 ファイルの例

```
@<TRIPOS>MOLECULE
methanol.mol2
 6 5 0 0 0
SMALL
NO_CHARGES

@<TRIPOS>ATOM
 1 C      0.7253  0.0134  0.0001 C.3  1 <1>    -0.1320
 2 O     -0.6859 -0.0645 -0.0000 O.3  1 <1>    -0.7323
 3 H      1.0981  1.0651  0.0186 H    1 <1>     0.1340
 4 H      1.0901 -0.5342  0.9059 H    1 <1>     0.1653
 5 H      1.0900 -0.5012 -0.9251 H    1 <1>     0.1644
 6 H     -1.0287  0.8351  0.0005 H    1 <1>     0.4006

@<TRIPOS>BOND
 1  1  2  1
 2  1  3  1
 3  1  4  1
 4  1  5  1
 5  2  6  1
```

【補足】 tplseneL における mol2 形式ファイル読み込み処理について**• tplseneL が参照する mol2 形式ファイルの情報**

tplseneL では、@<TRIPOS>MOLECULE 情報は参照しません。

指定された mol2 形式ファイルから、

@<TRIPOS>ATOM

:

@<TRIPOS>BOND

:

に記述された ATOM、BOND 情報のみを取得して処理を行います。したがって、

- ATOM、BOND 部分の記述において、書式にエラーがある場合、tplseneL はエラーメッセージを表示して終了します。
- ATOM、BOND 以外の@<TRIPOS>で書式にエラーがあっても、処理は継続されます。

• ATOM、BOND 項の status bit の扱いについて

MOL2 ファイル中の status bit は SYBYL 内部で指定されるものです。

有効なステータスビットは以下の様になっていますが、この項目はユーザが設定するものではありませんので、tplseneL ではこの項目のエラーチェックは行っていません。

(参考)ATOM 有効ステータスビット

DSPMOD、TYPECOL、CAP、BACKBONE、DICT、ESSENTIAL、WATER、DIRECT

(参考)BOND 有効ステータスビット

TYPECOL、GROUP、CAP、BACKBONE、DICT、INTERRES

• 結合のタイプの扱いについて

MOL2 ファイルで定義されている結合のタイプは以下に示すものがあります。

1 = single

2 = double

3 = triple

am = amide

ar = aromatic

du = dummy

un = unknown

nc = not connected

tplgeneL では、結合のタイプ"am"が指定されている場合には、内部的にはその結合を単結合として処理を行います。同様に、タイプ"ar"が指定されている場合には、内部的に芳香族結合として処理を行います。

タイプ"du"、"un"、"nc"が指定されている場合には、tplgeneL では処理を行うことができません。エラーメッセージを出力し、プログラムを終了します。

【エラーメッセージとその原因】

項番	エラーメッセージ	エラー原因
1.	ERROR> ltgReadMOParmMol2 Contents Error : <i>filename.mol2</i> Start of next line must not begin with "@" or "#", if line is continued with a back slash "\". Please check following information. (エラー付近のデータ)	継続番号の次行1文字目に "@" , "#" が存在する。
2.	ERROR> ltgReadMOParmMol2 Contents Error : <i>filename.mol2</i> It is necessary to describe sign "@" and "#" in column 1 of the line. Please check following information. (エラー付近のデータ)	行の1文字目以外に "@" , "#" が存在する。
3.	ERROR> ltgReadMOParmMol2 Contents Error : <i>filename.mol</i> Atom format is wrong. Please check following information. (エラー付近のデータ) または、 ERROR> ltgReadMOParmMol2 Contents Error : <i>filename.mol</i> Bond format is wrong. Please check following information. (エラー付近のデータ)	ATOM 項について、 継続番号のあとにデータが続いている。 1 項目め(原子 ID)が数字以外の文字含む 2 項目め(原子名)の1文字目が数字 3 項目め(x 座標)が数字、"-","."以外の文字含む 4 項目め(y 座標)が数字、"-","."以外の文字含む 5 項目め(z 座標)が数字、"-","."以外の文字含む 6 項目め(原子タイプ)の1文字目が数字 7 項目め(サブ構造 ID)に数字以外の文字含む 9 項目め(原子タイプ)が数字以外の文字含む 項目数が、6 個より少ない、または、10 個より多く記述されている。 BOND 項について、 継続番号のあとにデータが続いている。 1 項目め(結合 ID)が文字列を含む 2 項目め(原子 ID)が文字列を含む 3 項目め(原子 ID)が文字列を含む 項目数が、4 個より少ない、または、5 個より多く記述されている。
4.	ERROR> ltgReadMOParmMol2 Contents Error! The Bondtype ("結合のタイプ") that Mol2 Format does not support is found. Please check following information. (エラー付近のデータ)	BOND 項の4 項目め(結合のタイプ)が、MOL2 ファイル フォーマットで定義されていない文字列が指定され ている。
5.	ERROR> ltgReadMOParmMol2 Contents Error! The Bondtype ("結合のタイプ") that tplseneL does not support is found. Please check and modify following information. (エラー付近のデータ)	BOND 項の4 項目め(結合のタイプ)が、MOL2 ファイル フォーマットで定義されているが、tplseneL でサポ ートしていない結合のタイプ(du, un, nc) が指定さ れている。
6.	ERROR> ltgReadMOParmMol2 File Format Error : *.mol2 File Format is not correct.	ATOM 行または BOND 行がない
7.	ERROR> ltgDefineBond Contents Error! Isolated Atom ("原子名") that it has not any Bond is detected in Input File. Please check Input Files. : *.bond or *.mol2.	ATOM 項と BOND 項の整合性が取れていない ATOM 項が多い

項番	エラーメッセージ	エラー原因
8.	ERROR> ItgReadMOParmMol2 Contents Error! Bond information does not match to Atom information. Please check mol2 file " <i>filename.mol2</i> ".	ATOM 項と BOND 項の整合性が取れていない BOND 項が多い
9.	ERROR> ItgDefineBond Contents Error! The Bond Information is overlapped. : (重複している原子の組) Please confirm Input Files. : *.bond or *.mol2.	BOND 項同じ原子の組が複数回記述されている。

4. 原子タイプ定義ファイル

tplgeneL では、まず始めに計算したい分子の各々の原子の原子タイプを割り当て、次に各原子タイプの組に対応する力場のパラメータを割り当てます。

原子タイプ定義ファイルには、各原子の環境(元素記号、その原子の結合の数、結合次数、環内原子か否か、芳香族原子か否か)毎に対応する原子タイプ情報が記述されています。

tplgeneL には以下の 2 種類の力場の原子タイプ定義ファイルを用意しています。

原子タイプ定義ファイル

atomtype_gaff. db	AMBER GAFF 力場用原子タイプ割り当て規則 DB ファイル
atomtype_amber99. db	AMBER parm99 力場用原子タイプ割り当て規則 DB ファイル

5. 力場パラメータデータベースファイル

tplgeneL では、「4. 原子タイプ定義ファイル」で割り当てた原子タイプに基づいて、力場パラメータを割り当てます。

力場パラメータデータベースは、原子タイプ毎の結合パラメータ、結合角パラメータ、二面角パラメータ、インプロパー二面角パラメータ情報が記述されている"prm_XXXX. db"ファイル、及び、ファンクションパラメータ、ノンボンドパラメータが記述されている"nonbond_XXXX. db"ファイルで構成されています。

現在、力場パラメータデータベースファイルとして AMBER parm99 用のものと、AMBER GAFF 用のものを用意しています。

力場パラメータデータベースファイル

prm_gaff. db	AMBER GAFF 力場用パラメータ DB ファイル
prm_amber99. db	AMBER parm99 力場用パラメータ DB ファイル
nonbond_gaff. db	AMBER GAFF 力場用ノンボンドパラメータ DB ファイル
nonbond_amber99. db	AMBER parm99 力場用ノンボンドパラメータ DB ファイル

【補足】 AMBER GAFF パラメータについて

tplgeneL では AMBER ver. 7 GAFF (GAFF7) と AMBER ver. 8 GAFF (GAFF8) パラメータで計算を

行うことができます。GAFF7 ではほとんどの低分子で計算が可能となっています。GAFF8 は GAFF7 よりも計算できる分子は少ないが、正確な構造が計算できる事があります。

現在の仕様では GAFF7 と GAFF8 を同時に使用する事ができません。力場パラメータ DB ディレクトリ内の必要なファイルをコピーして使用して下さい。

なお、初期設定では GAFF7 の計算を行う事ができます。

GAFF8 関連ファイル

atomtype_gaff8.db、 prm_gaff8.db、 nonbond_gaff8.db

GAFF7 関連ファイル

atomtype_gaff7.db、 prm_gaff7.db、 nonbond_gaff7.db

例) ファイルのコピーについて

以下のコマンドで、必要な各ファイルのコピーを行ってください。

GAFF8 パラメータを使用する場合

```
cp prm_gaff8.db prm_gaff.db
cp atom_type_gaff8.db atom_type_gaff.db
cp nonbond_gaff8.db nonbond_gaff.db
```

GAFF7 パラメータを使用する場合

```
cp prm_gaff7.db prm_gaff.db
cp atom_type_gaff7.db atom_type_gaff.db
cp nonbond_gaff7.db nonbond_gaff.db
```


6. フラグメントデータベースファイル

tplgeneL では、「5. 力場データベースファイル」で説明した AMBER 由来のパラメータを割り当てる他に、分子の一部をフラグメントと見なし、そのフラグメント部分にユーザ独自のパラメータを割り振ることができます。

フラグメントデータベースファイルには、登録されているフラグメント（フラグメントブロック）毎に、以下の情報が書かれています。

- (1) フラグメント部分の原子パラメータ情報
- (2) フラグメント部分の結合パラメータ情報
- (3) フラグメント部分の結合角パラメータ情報
- (4) フラグメント部分の二面角パラメータ情報
- (5) フラグメント部分のインプロパー二面角パラメータ情報

フラグメントデータベースファイルを使用する場合には、上記の項目のうち、1、2は必須となっています。3～5は登録されている場合にはそのパラメータを使用します。

tplgeneL には以下の 2 種類の力場のフラグメントデータベースファイルを用意しています。

フラグメントデータベースファイル

frg_gaff.db	AMBER GAFF 力場用フラグメントデータベースファイル
frg_amber99.db	AMBER parm99 力場用フラグメントデータベースファイル

7. 環境変数

対象分子の構造に関するデータ、使用する力場、出力ファイルを保存するディレクトリをそれぞれ分けたい場合には、それらのパスを環境変数で指定することができます。

環境変数は以下の 3 種類を設定することができます。環境変数を設定しない場合には、実行時のカレントディレクトリを参照します。

■環境変数について

環境変数名	説明
TPLL_INPUT_PATH	: tplgeneL 入力ファイル用ディレクトリ (パスを含めて記述する)
TPLL_OUTPUT_PATH	: tplgeneL 出力ファイル用ディレクトリ (同上)
TPLL_DB_PATH	: tplgeneL 力場パラメータ DB 用ディレクトリ (同上)

(設定例)

tplgeneL 力場パラメータ DB 用ディレクトリを"/home/user01/myPresto/tplgeneL/DB"とする場合。

環境変数の設定は、使用しているシェルの設定方法に従います。

(下線部が入力部分です)

(1) bash の場合

```
% export TPLL_DB_PATH=/home/user01/myPresto/tplgeneL/DB
```

(2) csh の場合

```
% setenv TPLL_DB_PATH /home/user01/myPresto/tplgeneL/DB
```

※環境変数で設定するディレクトリへのパスが固定される場合は、ログインスクリプト(.bashrc, .cshrc) や専用スクリプトに書き込んでおくと便利です。

なお、現在使用しているシェルは以下のコマンドで確認できます。

```
% ps
```

以上。