

# *myPresto 4.2*

- *VCOL* -

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)

## 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	はじめに	5
1.1	システムの概要	5
1.2	システムの構成	6
2	母核・側鎖構造登録部	7
2.1	入力情報	7
2.2	処理フロー	8
3	仮想分子構造作成部	10
3.1	処理フロー	10
3.2	母核・側鎖の Z-matrix の結合方法	11
3.3	母核・側鎖の BOND 情報の結合方法	13
4	使用方法	15
4.1	配座生成ツール(configene)	15
4.1.1	入力	15
4.1.2	出力	16
4.2	仮想分子作成ツール	17
4.2.1	入力	17
4.3.2	出力	18

(余白)

# 1 はじめに

## 1.1 システムの概要

Virtual Combinatorial Library 構築システム (VCOL) は、母核構造および側鎖構造のデータファイル (mol2 形式) から、新規の仮想分子構造のデータファイル (mol2 形式) を生成するシステムです。本システムは、次の 2 つの処理部分から構成されています。

### ( 1 ) 母核・側鎖構造登録部

母核構造および側鎖構造のデータファイル (mol2 形式) から、その構造情報を取り出してデータ化 (Z-matrix、BOND 情報) し、母核・側鎖構造データベースを作成します。

### ( 2 ) 仮想分子構造作成部

母核・側鎖構造データベースから、母核構造とその置換基およびカテゴリー別に分けた側鎖構造セットを選択することで、指定置換基部分をセット内の各々の側鎖構造で置換した仮想分子を生成し、データファイル (mol2 形式) として出力します。

#### NOTE

母核・側鎖構造データベースは XML 形式のフラットファイルとし、DBMS は使用しません。

## 1.2 システムの構成

本システムは、母核・側鎖構造登録部、仮想分子構造作成部の、2つの処理部分で構成されます。以下に本システムの構成を示します。

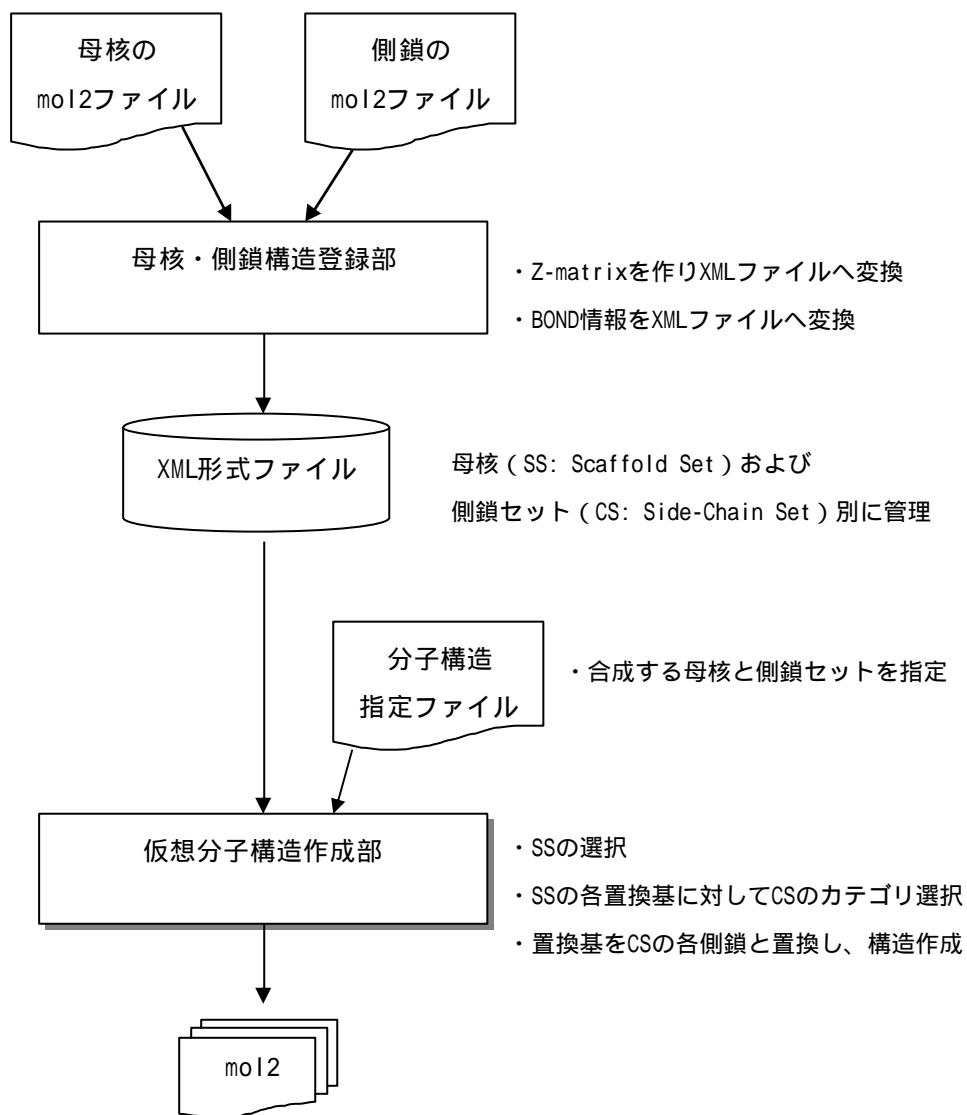


図1 . システム構成

## 2. 母核・側鎖構造登録部

### 2.1 入力情報

母核および側鎖の登録時のファイルは以下の形式で準備します。

#### (1) 母核構造

母核構造の入力ファイルは、置換基が結合しうる部分をダミー原子「R」として作成した mol2 ファイル形式とします。その際、原子タイプは R1, R2, ..., Rn を使用します。

複数の置換基 (ダミー原子 : R) を設定した場合、置換基は出現順序に従い、仮想分子合成プログラムの入力である側鎖セットの順序で置換されます。

#### (2) 側鎖構造

側鎖構造の入力ファイルは、母核が結合しうる部分をダミー原子「X」として作成した mol2 ファイル形式とする。その際、原子タイプは X1 を使用する。

複数のダミー原子「X」は使用できないこととする。

## 2.2 処理フロー

母核構造および側鎖構造のデータファイル (mol2 形式) から、その構造情報を取り出してデータ化 (Z-matrix、BOND 情報) し、母核・側鎖構造データベースを作成します。

入力ファイル名、部分構造情報 (構造名、側鎖名) を入力します。

入力した mol2 ファイルから、confgene により Z-matrix と bond 情報のファイルを作成します。

confgene が出力した Z-matrix と bond 情報、および、部分構造情報を含む XML ファイルを作成し、出力します。その際、カタログ番号を発行し XML ファイルに出力します。

作成された XML ファイルを指定ディレクトリに格納します。

ログとして登録されたファイル名と、カタログ番号を出力します。

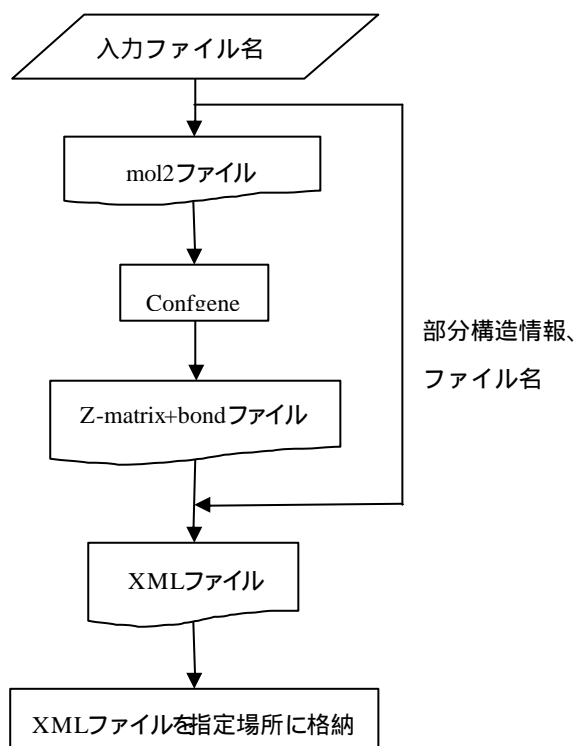


図 2 . 処理フロー



```
<vcl>
<id>
    4
</id>
<cs_id>
cs_002_01
</cs_id>
<set_id>
cs_002
</set_id>
<file>
    INgzmat/CS/cs_002_01.gzmat
</file>
<gzmat>
X1      0 0.0000000  0  0.0000000  0  0.0000000
C       1 1.7548858  0  0.0000000  0  0.0000000
H       2 1.1222162  1 109.7693253  0  0.0000000
H       2 1.0809097  1 112.0950546  3 120.3639145
H       2 1.1065120  1 110.2582626  3 -116.9284515
#
    1      1      2      1
    2      2      5      1
    3      2      4      1
    4      2      3      1
</gzmat>
</vcl>
```

図3 .XML形式ファイル例

### 3. 仮想分子構造作成部

#### 3.1 処理フロー

母核・側鎖構造データベースから、母核構造とその置換基および側鎖セットを選択することで、指定置換基部分をセット内の各々の側鎖構造で置換した仮想分子を生成し、データファイル (mol2 形式) として出力します。

母核とその置換基および置換基に結合する側鎖セット情報(SS\_id, CS\_id)を入力します。入力の SS\_id、CS\_id から対象となる XML ファイルを検索します。

対象となる SS\_id の XML ファイルより置換基の数を取得し、入力ファイルの CS\_id の数 (同一 ID の場合もあり) と比較し、置換基より側鎖セット数が少ない場合、エラーを出力し、処理を中止します。

入力ファイルの指定どおりの組合せに従い、XML ファイルから母核、側鎖の Z-matrix を取得し、連結処理を行います (下記 3.2 節を参照)。

XML ファイルから各母核・側鎖の BOND 情報を取得し、連結処理を行います (下記 3.3 節を参照)。

連結した Z-matrix ファイルから原子座標を設定し、BOND 情報を加えた形での mol2 ファイルを出力します。

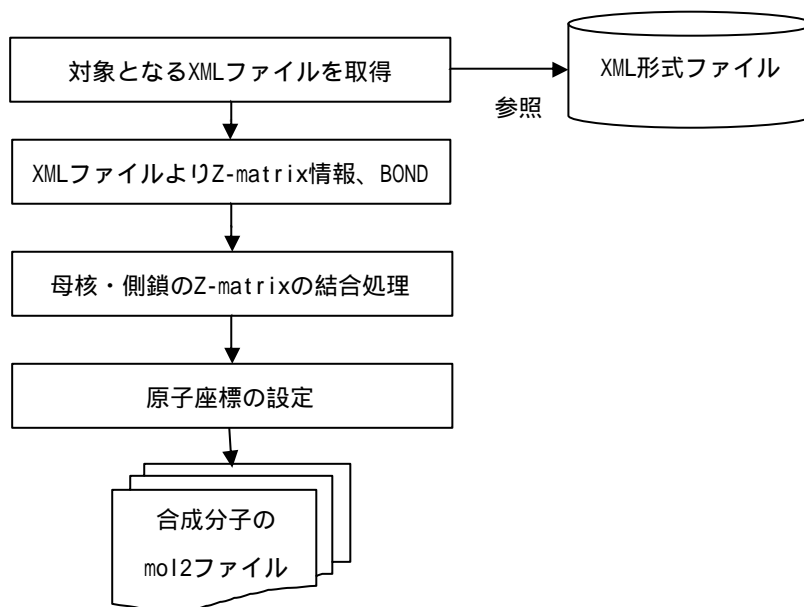


図 4 . 処理フロー

### 3.2 母核・側鎖の Z-matrix の結合方法

以下の方法に基づき、母核および側鎖の Z-matrix から、母核の置換基を側鎖の構造に置き換えた Z-matrix を作成します ( 図 4 (1) ~ (3) 参照 )。

母核の元素数から 2 引いた数を、側鎖の各元素 ID=1,2 以外の ID に加えます。

母核 Z-matrix より R1 の原子 ID ( A ) および、R1 が結合する原子 ID ( C ) を取得します。

側鎖 Z-matrix の原子 ID= 1 を上記 で取得した ( C ) の原子 ID に置き換え、原子 ID= 2 を ( A ) の原子 ID に置き換えます。

R1 の結合角 ( D )、二面角 ( E ) を成す原子の ID を取得します。

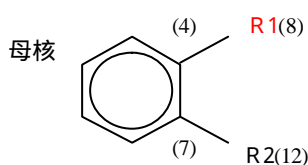
側鎖 Z-matrix の 2 行目の ( D )、( E ) の原子 ID を上記 の値に置き換えます。

側鎖 Z-matrix の 3 行目の ( E ) の原子 ID を上記 の ( D ) の原子 ID の値に置き換えます。

母核の R1 の行を、側鎖の 2 行目で置き換え、母核の最終行以降に側鎖の 3 行目以降を追加します。

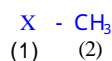
さらに Rn ( n=2,3, … ) があれば、結合させる側鎖 Z-matrix との組合せで、上記 ~ の処理を行います。

( Rn がなくなるまで繰り返します。 )



1	C	0	0	0
2	H	1	0	0
3	C	1	2	0
4	C	1	2	3
5	C	4	1	3
6	C	3	1	4
7	C	5	4	1

側鎖



1	X	0	0	0
2	C	1	0	0
3	H	2	1	0

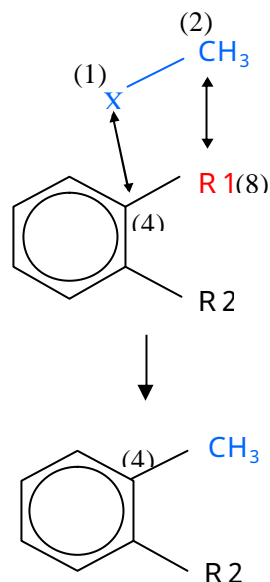


図 5 (1) 母核、側鎖の Z-matrix の例

図 5 (2) 母核、側鎖の連結



### 3.3 母核・側鎖の BOND 情報の結合方法

以下の方法に基づき、母核および側鎖の BOND 情報から、母核の置換基を側鎖の構造に置き換えた BOND 情報を作成します（図 5 (1) ~ (2) 参照）。

母核の元素数から 2 引いた数を、側鎖の各元素 ID=1,2 以外の ID に加えます。

母核 Z-matrix より R1 の原子 ID および、R1 が結合する原子 ID を取得します。

（上記 3.2 節の の処理結果を利用します。）

側鎖 Z-matrix の原子 ID=1 を上記 で取得した (C) の原子 ID に置き換え、原子 ID=2 を (A) の原子 ID に置き換える。

母核の最終行以降に側鎖の 2 行目以降を追加し、原子 ID でソートを掛けます。

さらに Rn (n=2,3, ...) があれば、結合させる側鎖の BOND 情報との組合せで、上記 ~ の処理を行います。

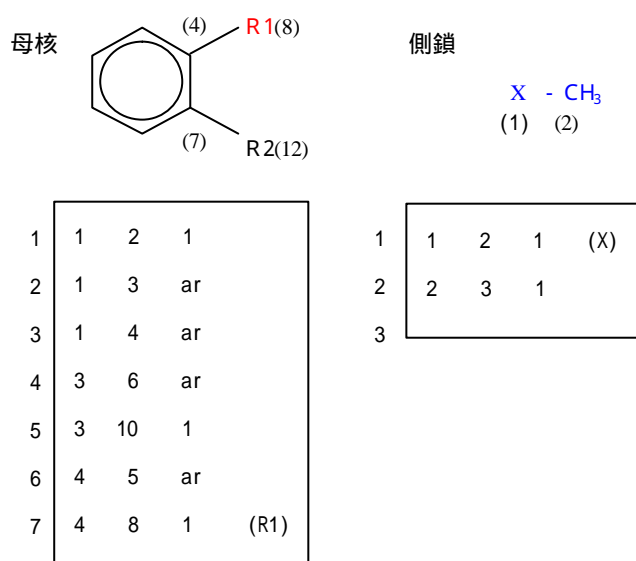


図 6 (1) 母核、側鎖の BOND の例



## 4 使用方法

### 4.1 配座生成ツール(confgene)

mol2 形式のファイルを読み込み、分子を構成する原子の座標を bond 距離、angle 角、torsion 角の 3 つのパラメータで表現する Z-matrix フォーマットファイルを作成します。

複数の配座を指定した場合は、回転可能な torsion を検出し、この torsion を回転させた場合の配座での Z-matrix ファイルを生成します。

#### 4.1.1 入力

##### (1)制御ファイル

制御ファイルは 4 行で、それぞれ以下の項目を記述します。

- (1-1)mol2 ファイル名
- (1-2)生成配座数(1 の場合は、torsion は回転させない)
- (1-3)torsion の回転候補角数
- (1-4)総電荷の値("a"の場合は自動で計算を行う)

制御ファイル例)

LIG.mol2	;	入力 mol2 ファイル名
1	;	生成配座数
3	;	torsion を回転候補角 (360° / 3 = 120° ずつ回転が行われる)
a	;	総電荷を自動計算

##### (2) mol2 ファイル

"CH<sub>3</sub>"の mol2 ファイル例)

```
@<TRIPOS>MOLECULE
m1
      4      3      1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 C1      5.2414  -3.0575  -0.0645 C.3      1 RES      0.2900
  2 H11     5.5684  -3.0437  -1.1215 H      1 RES      0.0000
  3 H12     5.6242  -3.9411   0.4265 H      1 RES      0.0000
  4 H13     5.7021  -2.1572   0.4219 H      1 RES      0.0000
@<TRIPOS>BOND
  1      1      2  1
  2      1      3  1
  3      1      4  1
@<TRIPOS>SUBSTRUCTURE
  1 RES      1
```

## 4.1.2 出力

ファイル名"zmat.dat"のファイルに以下のフォーマットで出力されます。

表中の n は原子数で、m は bond 数を示します。

項番	行	列	内容	型	備考
#1	1 ~ n	1	原子名	文字型	-
#2		2	bond 相手の原子の行	整数	先頭原子は 0
#3		3	bond 距離	実数	先頭原子は 0.0
		4	angle 相手の原子の行	整数	1,2 番目原子は 0
		5	angle 角	実数	1,2 番目原子は 0.0
		6	torsion 相手の原子の行	整数	1~3 番目原子は 0
		7	torsion 角	実数	1~3 番目原子は 0.0
	n+1	1	bond 開始識別行	"#"	-
	n+2 ~ n+m+1	1	bond の連番	整数	-
		2	bond の原子 1 の行	整数	-
#4		3	bond の原子 2 の行	整数	-
		4	bond の結合次数	実数	-

"CH<sub>3</sub>"の Z-matrix ファイル例)

C	1	1.7548858	0	0.0000000	0	0.0000000
H	2	1.1222162	1	109.7693253	0	0.0000000
H	2	1.0809097	1	112.0950546	3	120.3639145
H	2	1.1065120	1	110.2582626	3	-116.9284515
#						
	1	1	2	1		
	2	1	3	1		
	3	1	4	1		



## 4.2 仮想分子作成ツール

母核構造と側鎖構造の Z-matrix ファイルを読み込み、二つを合成した仮想分子の mol2 ファイルを出力する。

母核構造に置換基が複数存在する場合、複数の側鎖構造を合成する。

### 4.2.1 入力

#### (1)制御ファイル

制御ファイルの行数は(1-1), (1-2)がそれぞれ 1 行で、(1-3)が置換する側鎖の個数分必要です。

#### (1-1)mol2 ファイルの出力ディレクトリ

出力ファイル名は母核と側鎖の名称を組み合わせて生成します。

#### (1-2)母核の Z-matrix ファイル名

母核の置換対象原子は原子名"R"を指定します。

#### (1-3)側鎖の Z-matrix ファイル名

側鎖のダミー原子は原子名"Xn"を指定します。(n は任意の文字列)

制御ファイル例)

output/mol2	; 出力ディレクトリ名
scaford.zmat	; 母核の Z-matrix ファイル名
side1.zmat	; 側鎖 1 の Z-matrix ファイル名
side2.zmat	; 側鎖 2 の Z-matrix ファイル名

### 4.3.2 出力

出力される合成分子の mol2 ファイル名は以下の名称になります。

母核名+"\_"+側鎖 1 名+...+"\_"+側鎖 n 名

合成分子の mol2 ファイル例)

ss\_01\_1.zmat と cs\_001\_01.zmat を合成した ss\_01\_1\_cs\_001\_01.mol2 ファイル

```
@<TRIPOS>MOLECULE
ss_01_1_cs_001_01
  17  18  2
SMALL
NO_CHARGES
@<TRIPOS>ATOM
  1  H  0.0000  0.0000  0.0000  H  1  SS  0.0000
  2  C  1.0736  0.0000  0.0000  C  1  SS  0.0000
  3  C  1.7789  1.2162  0.0000  C  1  SS  0.0000
  4  H  1.2509  2.1590  0.0412  H  1  SS  0.0000
  5  C  3.1795  1.1458 -0.0346  C  1  SS  0.0000
  6  C  4.1586  2.2592 -0.0426  C  1  SS  0.0000
  7  H  4.0404  2.8543  0.9024  H  1  SS  0.0000
  8  H  4.0063  2.9218 -0.8996  H  1  SS  0.0000
  9  C  5.4272  1.5305 -0.0832  C  1  SS  0.0000
 10  F  6.7350  2.3229 -0.0762  F  2  C01 0.0000
 11  C  5.2833  0.2174 -0.1184  C  1  SS  0.0000
 12  H  6.0502 -0.5698 -0.1651  H  1  SS  0.0000
 13  C  3.8517 -0.0497 -0.1037  C  1  SS  0.0000
 14  C  3.1274 -1.2559 -0.0963  C  1  SS  0.0000
 15  H  3.6540 -2.2315 -0.1411  H  1  SS  0.0000
 16  C  1.7315 -1.2217 -0.0819  C  1  SS  0.0000
 17  H  1.1885 -2.1543 -0.0988  H  1  SS  0.0000
@<TRIPOS>BOND
  1  1  2  1
  2  2 16  2
  3  2  3  1
  4 16 17  1
  5 14 16  1
  6 14 15  1
  7 13 14  2
  8  5 13  1
  9 11 13  1
 10  3  5  2
 11  5  6  1
 12  3  4  1
 13  6  8  1
 14  6  7  1
 15  6  9  1
 16  9 10  1
 17  9 11  2
 18 11 12  1
@<TRIPOS>SUBSTRUCTURE
  1  SS  1
  2  C01 10
```

*myPresto 4.2*

- VCOL -