

AMBERの残基データベースへの新規残基追加支援システム開発

佐々 和洋^{a*}, 宇野 健^b, 林 治尚^a, 山名 一成^a, 中野 英彦^a

^a 姫路工業大学大学院工学研究科物質系工学専攻, 兵庫県姫路市書写 2167

^b 広島県立大学経営学部経営情報学科, 広島県庄原市七塚町 562

*e-mail: c4kaz@chem.eng.himeji-tech.ac.jp

(Received: June 3, 2003; Accepted for publication: July 18, 2003; Published on Web: November 20, 2003)

AMBER (Assisted Model Building with Energy Refinement) は生体関連分子のシミュレーションに適しているが、AMBER 内の残基データベースに登録されていない新規残基部位を有する DNA 分子などのシミュレーションを行うためには、モジュールのひとつである PREP を用いて新規に登録しておかなければならない。しかし、その入力ファイルの形式は繁雑であるため、手作業にて作成するのは容易ではない。そこで本研究では、これまで開発してきた“Modrast-P”を機能拡張し、視覚的に確認しながら対話的に PREP 入力ファイルを作成する機能を追加した。これにより、簡便かつ効率的に PREP 入力ファイルが作成可能となった。

キーワード: DNA, AMBER, PREP, Modrast-P, 分子動力学

1 緒言

DNA や RNA の塩基配列は、それら遺伝子から作成されるタンパク質と密接な関係にあることがよく知られている。近年では、ヒトゲノム計画に代表されるように、DNA および RNA の塩基配列の特定、それらの塩基配列によって得られる情報や、遺伝子とタンパク質との相互作用の解明など、遺伝子にまつわる様々な研究がおこなわれ、機能性を有した遺伝子やそれに準ずる物質の開発がおこなわれている。

我々も種々の機能性を有した DNA の合成実験をおこなっている。また、それと並行して、AMBER[1, 2] を用いたコンピュータシミュレーションによる DNA の構造や挙動の解析等も行っている。

しかし、様々な機能性を持つことを目的として合成された DNA には、何らかの化学的な修飾が施されている事が多い。そういう化学修飾部位のほとんどは AMBER の残基データベース(以降 AMBER データベースとする)に登録されておらず、それらを含む物

質のシミュレーションを行うためには、その化学修飾に対する結合データや各原子の持つ電荷などの情報を新たに登録する必要が生じる。AMBER データベースへの登録は AMBER の機能のひとつである“PREP”を用いるが、その際必要となる入力ファイルは、様々なデータを規定のフォーマットに従い作成しておかなければならない。しかし、新規化学修飾部位の原子数が増加するほど入力データ量も増加するため、手作業による作成は容易ではない。

そこで、以前より我々が開発している、DNA やタンパク質などに対応した分子構造表示システム Modrast-P[3, 4] への機能拡張という形で、新規化学修飾部位の登録作業を容易にすることを目的としたシステムの開発を行った。このシステムは、Modrast-P の分子構造表示機能と原子認識機能などを利用し、新規化学修飾部位の構造やその一つ一つの原子を視覚的に確認しながら、選択式かつ対話形式による一連の作業により、PREP 入力ファイルを作成するものである。

2 開発環境および動作環境

開発環境は C 言語 (ANSI C, gcc-2.95.2, 同 2.96) で、動作環境は XFree86 が備わっている UNIX OS を搭載したマシンである。

また本システムは、PREP の機能を備えている AMBER 4.1 以降、現行の AMBER 7 にも対応している。

3 AMBER と化学修飾 DNA

AMBER とは、UCSF の Kollman らによって作成された分子動力学シミュレーションソフトであり、生体高分子に適した独自のパラメータを持ち、数多くの研究に用いられている。AMBER には、エネルギー極小化計算、分子動力学計算、自由エネルギー摂動法計算などの種々のパッケージがあり、シミュレーション条件や初期配置の決定などにおいても、それぞれの目的に応じた種々のモジュールが用意されている。

その中で、シミュレーションの初期条件の設定を行うモジュールである“LEaP”は、シミュレーションターゲットとなる分子の構造を作成するために、原子や結合の変更や残基の付加が可能であり、ターゲット分子の周囲環境を構築するために、溶媒やカウンターイオンの挿入などが実行できるものである。

しかし、AMBER のデータベースに登録されていない部位を含む分子の場合では、LEaP モジュールの機能の一つである“PREP”を利用して、その部分の情報を新規部位として AMBER のデータベースに予め登録しておく必要がある。

PREP を用いた新規部位の登録には、データベースに関する制御や、空間を定義するためのダミー原子の取り扱いなどの制御、および新たに登録する部位の各原子に対する便宜上の原子名や座標、形式電荷、原子間の結合情報などのデータを、入力ファイルとして所定のフォーマットに従い用意しておかねばならない。

4 Modrast-P

Modrast-P は我々が開発している分子構造表示システムであり、Protein Data Bank(PDB)[5] など様々なフォーマットに対応している。構造表示はラスタ形式で表示され、Wireframe Model、Ball-and-Stick Model および Space-Filling Model を基本表示形態とし、その他に

も Tubular Model 等の表示が可能である。

また、Modrast-P には切断図形表示機能が備わっており、1 平面および 2 平面による切断や、部分切断や円柱型にくりぬいた分子構造を表示することができ、タンパク質など高分子の内部構造やカップリングの様子などを視覚的に確認するの非常に有効である。また、Modrast-P には切断図形表示機能が備わっており、1 平面および 2 平面による切断や、部分切断や円柱型にくりぬいた分子構造を表示することができ、タンパク質など高分子の内部構造やカップリングの様子などを視覚的に確認するの非常に有効である。

この他、AMBER の計算結果のアニメーション機能や原子間距離計算機能、全エネルギーデータ等の抽出機能などを有し、シミュレーション結果の解析に有効なソフトウェアである。

5 PREP 入力ファイル

PREP 入力ファイルは、Table 1 の書式に従って作成する必要がある。

1 ~ 2 行目は、AMBER のデータベースファイルに関して指定する部分であり、1 行目には新たなデータベースファイル生成の有無、データの保存法、力場の制御に対するフラグ (書式 3I) を入力しなければならず、2 行目には新たにデータベースを作成する場合や、以前に作成したデータベースを使用する場合はそのファイル名 (書式 A80) を入力しなければならない。

3 行目からは、PREP 入力ファイルと一連の新規残基登録において用いられるタイトル (書式 20A4) を、次には出力ファイル名 (書式 A80) を設定し、さらには AMBER のデータベースへ登録される残基名と座標データの形式および出力ファイルを binary 出力するかどうかの指定をする (書式 2A,I)。6 行目は 8 行目以降の原子構造データのフォーマット指定と残基登録時においてダミー原子を無視するかどうかの処理、ダミー原子を示す Atom Type の名称、ダミー原子を消去する際の消去法を指定し (書式 4A)、7 行目は Cutoff 距離を示す (書式 F)。

8 行目からは新規残基の各原子に対する種々の情報を記述する必要がある、1 行内に原子の通し番号、原子の Unique Atom Name、Atom Type、Topological Type (tree symbol)、座標、形式電荷の値を示す (書式 I,3A,(3I),4f)。

Table 1. Format of Prep input file

	書式	内容
1 行目	3I	データベースに関する機能制御、新たなデータベースファイル生成の有無、データの保存法、力場の制御
2 行目	A80	新たにデータベースを作成する場合や、以前作成したデータベースを使用する場合のファイル名
3 行目	20A4	入力ファイルのタイトル
4 行目	A80	出力ファイル名
5 行目	2A,I	AMBER に登録した際の残基名、座標データ種、出力フォーマット
6 行目	4A	原子構造データの形式、ダミー原子に対する処理、ダミー原子の固有名、ダミー原子の消去法
7 行目	F	Cutoff 距離
8 行目以降	I,3A,(3I),4F	原子の通し番号、原子の固有名、原子タイプ名、ツリー構造、座標値、形式電荷値
LOOP 以降	2A	原子間の結合情報

LOOP CLOSING EXPLICIT 以降は原子間の結合情報を表し、結合する各原子の Unique Atom Name を示す (書式 2A)。

最後に Done および Stop で PREP 入力ファイルの終了を宣言する。

6 プログラムの機能

今回作成した PREP 入力ファイル作成機能は、Modrast-P を機能拡張する形でコーディングした。

入力ファイルを作成するには次のような手順で行う。

6.1 新規残基構造の作成

まずは、AMBER のデータベースに登録する新規残基の構造を作成する。

我々の開発している Modrast-P には、現状では新たな分子や残基構造を作成する機能が備わっていないため、Molda[6] などの分子モデリングソフトを用いて、新規残基の構造を予め生成しておく。

6.2 Modrast-P での操作 - 分子軌道計算

Modrast-P から、分子モデリングソフトにより出力されたファイル (PDB 形式もしくは MOL 形式) を指

定して読み込むことで、新規残基の原子座標および結合情報などの必要なデータが取得される。続いて、Modrast-P のメニューより、分子軌道計算 MOPAC[7] の実行を選択すると、座標や結合データなど必要なものを用いて MOPAC 入力ファイルを一時ファイルとして作成し、MOPAC の実行が行われる。計算終了後、MOPAC 出力ファイルより、新規残基の最適化構造や各原子が持つ形式電荷の値など、PREP の入力ファイルに必要なデータが自動的に取得される。

6.3 Modrast-P での操作 - 対話形式でのデータ入力

分子軌道計算の終了後、Figure 1 のように対話形式での PREP のデータ入力モードとなる。Modrast-P の分子構造表示機能によりディスプレイに表示された残基構造で原子を確認しつつ、必要なデータを対話形式で入力することにより、自動的に PREP 入力ファイルが作成される。

7 実行例

本プログラムを用い、ピレン修飾残基の PREP 入力ファイルを作成し、PREP 機能を使用して登録を行った。以下に手順を示す。

```

kterm
Input title > Amine - Uridine unite atom
Input Name of Residue data file > PYR
Input Pass of Residue data file > /pub/home1/c4kei/exa6/pyr-uri/prep1.in

Input Atom Element for No.1
  1: C  2: H  3: N  4: O  5: others > 1

Input Unique Name for No.1 > CA2

Input Atom Type for No.1
  C : sp2 C carbonyl group
  CA: sp2 C pure aromatic (benzene)
  CB: sp2 aromatic C, 5&6 membered ring junction
  CC: sp2 aromatic C, 5 memb. ring HIS
  CK: sp2 C 5 memb.ring in purines
  CM: sp2 C pyrimidines in pos. 5 & 6
  CN: sp2 C aromatic 5&6 memb.ring junct.(TRP)
  CQ: sp2 C in 5 memb.ring of purines between 2 N
  CR: sp2 arom as CQ but in HIS
  CT: sp3 aliphatic C
  CV: sp2 arom. 5 memb.ring w/1 N and 1 H (HIS)
  CW: sp2 arom. 5 memb.ring w/1 N-H and 1 H (HIS)
  C*: sp2 arom. 5 memb.ring w/1 subst. (TRP)
Command > CT

Input Tree Structure Type for No.1 >
  M : Main Chain Atom
  S : Side Chain Atom
  B : Branch Atom
  E : Chain End Atom
Command > E

```

Figure 1. PREP input file making screen of Modrast-P

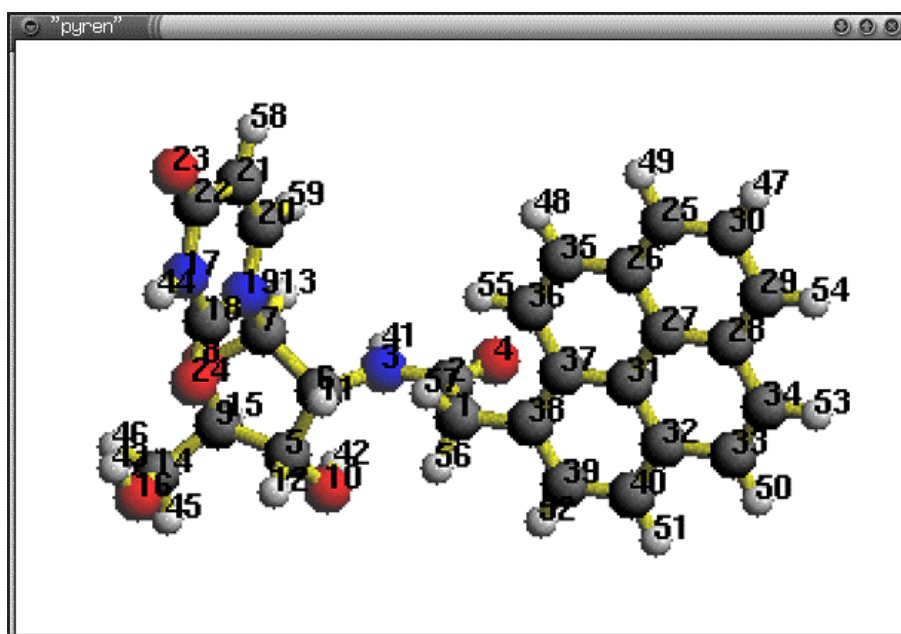


Figure 2. Pyrene-modified uridine with molecular structure display window of Modrast-P

Table 2. File format select menu of Modrast-P

1	Original PDB	以前の Modrast で用いられていた PDB フォーマット
2	AMBER PDB	AMBER や現在の PDB で用いられている PDB フォーマット
3	Modrast/Molda	Modrast-P や MOLDA より出力される mol ファイル

7.1 新規残基構造データの作成

MOLDA など分子モデリングソフトもしくは PDB などのデータベースから新規残基の構造データを作成する。

7.2 新規残基構造データの読込

Modrast-P を起動し、“File” “Open”から、ファイルフォーマットの選択となる。Table 2 に示す 3 種のフォーマットに対応しているため、6.1 において作成したファイルに適したものを選択し、ファイル名を指定する。

新規残基の構造データ読込が完了すると、Figure 2 に示す Modrast-P 分子構造表示ウィンドウに新規残基の構造が表示される。

7.3 MOPAC の実行

“MODELING” “Mopac”を選択すると、分子軌道計算 MOPAC の実行を選ぶメニューとなる。ここで、カルボニル基の有無など、MOPAC の実行に必要な情報を指定する。完了後、Modrast-P より MOPAC が実行され、計算結果から PREP 入力ファイルに必要な最適化構造と形式電荷の情報が Modrast-P に取得される。

7.4 PREP 入力ファイル作成

続いて、対話形式の PREP のデータ入力に移る。まずは、PREP 入力ファイルのタイトルと残基名、およびパスを入力する。

入力後、原子毎の情報を入力する段階へと移行する。ここでは最初に Atom Element を選択する指示が出るので、1:炭素・2:水素・3:窒素・4:酸素・5:その他の中から、指定原子の元素に相当する数字を入力する。

次にこの原子の Unique Name、Atom Type、Tree Structure Type の入力を行う。

Unique Name とは、各原子に対し任意の文字列 (4

文字まで)を与えるものであるが、混乱を避けるため PDB に沿った命名を行う事が望ましい。

Atom Type は、“C”ならば sp^3 混成軌道炭素や芳香族炭素であることを示すもので、ここでは Atom Element に対応した Atom Type のみが表示されるので、ユーザはその中から該当するものを選択し入力する。この Atom Type は、AMBER の parm94.dat で定義されたものであり、各原子の質量や原子間の結合距離および結合角等の決定に使用される。

Tree Structure Type は、主鎖・側鎖・枝分れ・末端の 4 つの構成状態から原子に見合った状態を選択肢の中から入力する。以上の 3 つの Type 入力を新規残基の全原子に対して行う。

全ての原子に対して Atom Element、Unique Name、Atom Type、Tree Structure Type を入力後、Define Loop において Tree Structure Type では定義しきれなかった原子間の結合を、画面上に表示されている原子の番号を入力することで定義をする。

最後に Define Loop に 2(No)を入力すれば、このプログラムが終了し、自動的に PREP 入力ファイル”prep.in”が作成される (Figure 3)

7.5 AMBER への登録

AMBER を起動し、モジュールの 1 つである xLEaP を起動する。Universe Editor より PREP 入力ファイルを読み込み、新規残基の登録を行う。

Unit Editor を起動し、Import Unit から先ほど登録した残基名を選択する。Editor 画面に新規残基が表示されるので、DNA 等と結合し、目的とする分子の作成を行う (Figure 4)

このように PREP 入力ファイルを作成および登録したピレン修飾ウリジンをを用いて、ピレン修飾 DNA および RNA についてシミュレーションを行ったところ、実験により観測された結果 [8-10] と符合する結果を得た。このことから、本プログラムを用いることで PREP 入力ファイルは正確に作成され、種々の新規残基に対応可能であることが示された。

8 結言

本プログラムにより、テキストエディターでの手作業により作成が困難であった PREP 入力ファイルの作成が容易なものとなった。特に、新規残基の各原子に関する情報入力においては、Modrast-P の分子構造表示機能の利用により、新規残基を視覚的に確認しながら、情報の入力が行えるため、入力している原子の判別が非常にしやすくなり、入力ミス等人為的なミスの軽減に大いに貢献するものと思われる。また、各原子の情報の入力に関しても、選択式となっているため簡便性に富み、効率的に PREP 入力ファイルが作成可能となった。今回はピレン修飾体を作成したが、核酸の修飾体だけでなく、アミノ酸の修飾体や補酵素など、AMBER の残基データベースに登録されていない残基を含む物質についてのシミュレーションが容易となった。

なお、本システムは次の URL にて公開予定である。
<http://www.chem.eng.himeji-tech.ac.jp/chem/labo04/>

参考文献

[1] Weiner, S. J., Kollman, P. A., Nguyan, D. T. and Case, D. A., *J. Comput. Chem.*, **7**, 230 (1986).

[2] <http://www.amber.ucsf.edu/amber/amber.html>

[3] Uno, T., Hayashi, H., Yamana, K. and Nakano, H., *J. Chem. Software*, **4**, 1 (1998).

[4] Uno, T., Hayashi, H., Yamana, K. and Nakano, H., *J. Chem. Software*, **5**, 39 (1999).

[5] <http://www.rcsb.org/pdb/>

[6] <http://www.molda.org/>

[7] Stewart, J. J., 内田希訳, *MOPAC マニュアル Version 6*, 日本化学プログラム交換機構 (1991).

[8] Yamana, K., Iwase R., Furutani S., Tsuchida H., Zako H., Yamaoka T. and Murakami A., *Nucleic Acids Research.*, **27**, 11 (1999).

[9] Yamana, K., Zako H., Asazuma K., Iwase R., Nakano H. and Murakami A., *Angew. Chem. Int. Ed.*, **40**, 6 (2001).

[10] Mahara A., Iwase R., Sakamoto T., Sakamoto T., Yamana, K., Yamaoka T. and Murakami A., *Angew. Chem. Int. Ed.*, **41**, 19 (2002).

Development of an Assisting System to Add New Residues to the Standard AMBER Residue Database

Kazuhiro SASA^{a*}, Takeshi UNO^b, Haruhisa HAYASHI^a, Kazushige YAMANA^a and Hidehiko NAKANO^a

^aDepartment of Materials Science and Chemistry, Graduate School of Engineering, Himeji Institute of Technology
2167 Shosha, Himeji, Hyogo, 671-2201, JAPAN

^bDepartment of Management and Information Sciences, Hiroshima Prefectural University
562 Nanatsuka, Shobara, Hiroshima, 727-0023, JAPAN

**e-mail: c4kaz@chem.eng.himeji-tech.ac.jp*

AMBER (Assisted Model Building with Energy Refinement) is a powerful tool for simulations of biomolecules and so on. To simulate DNA molecules with new residues that are not registered in the AMBER database, it is necessary to register them on this database by using one of the AMBER modules, PREP. However, the format of the PREP input file is very complicated, and it is hard to make by hand. Therefore, in this work, in order to make the PREP input file interactively with visualization, we add a function to our molecular structural display system, "Modrast-P". By using this function, the PREP input file can be made easily and efficiently.

Keywords: DNA, AMBER, PREP, Modrast-P, Molecular dynamics