Bioinformatics （Akiyama）　　Exercise #14

Student ID　　Name　　

問１．　以下の文章内空欄(1)～(8)に当てはまる語句を選び、記号で答えよ。

タンパク質立体構造データは、[(1)]、核磁気共鳴(NMR)、極低温電子顕微鏡などの手法で観測され、決定された各原子の三次元座標の値は国際的なデータベースである[(2)]に格納されている。タンパク質立体構造の観測手法は飛躍的に発達しており、収納エントリー総数も 2008 年 7 月時点で[(3)]を越えているものの、DNA 配列決定技術の圧倒的な進歩のために、現時点でも DNA 配列の収納エントリー総数の方が、立体構造の[(4)]倍前後であり、今後その差は圧倒的に広がると危惧されている。
このため、計算機による立体構造予測の技術が重要になっている。

立体構造予測の技術として、最も実用的に利用されているのは、主鎖は類縁の構造を使い、主に側鎖だけを調整する[(5)]の手法であり、代表的なソフトウェアとしては A. Sali らによる MODELLER がある。もしも主鎖の位置を正確に決められれば、側鎖をぶつかり合わないように配置する組み合わせは限られてくるので、制約充足問題として大規模な木探索などで解くことができる。この際、側鎖が取り得る候補構造を[(6)]と称し、そのライブラリとしては R. Dunbrack の SCWRL がある。

1990 年に M. Sippl らは、ある配列を、正しい構造と偽の（decoy）構造とに流し込んだとき、原子間の配置の安定性の違いから両者の区別を付けられることを発見した。これが、立体構造予測の一つの手法である[(7)]法の始まりである。すなわち、既知の代表的な（部分）構造を多数準備しておき、そのどれと対象配列が相性が良いかを判断することで、進化距離が遠くて配列上では明確な相同性が発見できないケースでも、立体構造の面では良い鋳型構造を見つけることが可能となった。

ほぼ同時に、D.Eisenberg らは、任意のタンパク質立体構造を与えられたときに、その構造に適する正解の配列を見つけ出すという「逆フォールディング問題」を提案し、その解決手法として、[(8)]プロファイル法を提案した。この手法では、望ましい配列に関するプロファイル情報をまず作成するので、後は立体構造に関係なく、通常の配列解析のような手法で膨大な配列データベース内から適切な配列を探索することができる。この手法を応用すれば、立体構造が未知のタンパク質間でも構造の類似性を指摘可能となる。

選択肢：

(a) GenBank　(b) PDB　(c) SWISS-PROT　(d) ５億件 (e) ５０万件　(f) ５万件
(g) X線結晶構造解析　(h) ハイブリダイゼーション　(i) 2000(倍)　(j) 200(倍)
(k) アブイニシオ予測　(l) ホモロジー・モデリング (m) スレッディング (n) 分子動力学
(o) 3D-1D　(p) 二次構造 (q) ロータマー

問１回答欄

(1)　　(2)　　(3)　　(4)

(5)　　(6)　　(7)　　(8)

Ans.

問２　簡単な二次元の正方格子モデルで、タンパク質の折りたたみを考えることにする。
下記は３×３の二次元空間であり、各残基は＊印の格子点にしか存在できない。
また連続する残基はエッジでつながれる縦か横の隣接する点に並ばねばならない。
今、長さ９残基のタンパク質配列があるとすると、図右側の Conformation 1 は、
ありえる一つの構造を表している。３×３の空間をはみ出してはいけない。

```
＊――＊――＊
|     |     |
＊――＊――＊
|     |     |
＊――＊――＊
```

2-D　Lattice

```
1――2――3
          |
6――5――4
|
7――8――9
```

Conformation 1

（１）残基番号４と残基番号９が隣接し、さらに９が空間の中心（他の残基に囲まれる位置）に位置するような構造を１つ書け。

（２）残基番号３と残基番号８が隣接する構造を１つ書け。