AA reference frame- x-, y- and z-axes

Stacey

Defining the AA reference frame in a quantitative way

1) Defining an ideal AA reference frame
   x- Cα→ Cβ
   y- cross product- (Cβ- Cα)×(N- Cα)
   z- complete a right-handed orthonormal coordinate system
   the length corresponds to the real bond length (e.g. C-C bond is about 1.5 nm)
2) Least square fitting
   Superimposing ideal AA reference frame and obtaining the AA reference frame for the AA-bp contacts
   Calculation is similar with the one in getting "middle frame" (base-pair)

DataAA reference frame

```
REMARK    Part of SNAP by Dr. Xiang-Jun Lu [October, 2006]
ATOM      1  N   AAA A   1      -0.498   1.364  -0.000
ATOM      2  CA  AAA A   1       0.000   0.000   0.000
ATOM      3  C   AAA A   1      -0.517  -0.772   1.200
ATOM      4  CB  AAA A   1       1.537  -0.000   0.000
END
```

Corresponding pdb data

```
MODEL       2
REMARK    Section #0002
REMARK    C111.A-D218.T:A5.ARG 9ant.mod.pdb
…
ATOM     42  N   ARG A   5     -11.308   2.285  -0.764  1.00 48.38           N
ATOM     43  CA  ARG A   5     -10.745   0.936  -0.712  1.00 50.59           C
ATOM     44  C   ARG A   5     -11.601  -0.056   0.061  1.00 51.43           C
ATOM     46  CB  ARG A   5      -9.353   0.937  -0.076  1.00 46.38           C
…
```

.par file

```
C111.A-D218.T:A5.ARG 9ant.mod.pdb +   -10.7995   25.8394
 -10.6263    1.2572    1.4595    8.3168  -24.2450    3.2969
 -10.7362    0.9216   -0.7160
   0.9104    0.0245    0.4130
  -0.0849    0.9881    0.1285
  -0.4050   -0.1521    0.9016
 -10.7447    0.9360   -0.7120
```

Least square fitting

(exactly follow the steps in “technical details”)

1. Standard and experimental matrix

Standard matrix

N -0.498 1.364 -0.000
CA 0.000 0.000 0.000
C -0.517 -0.772 1.200
CB 1.537 -0.000 0.000

Experimental matrix

N -11.308 2.285 -0.764
CA -10.745 0.936 -0.712
C -11.601 -0.056 0.061
CB -9.353 0.937 -0.076

```
> e_ave=c(-10.752, 1.0255,-0.373)
> s_ave=c(0.1305,0.148,0.3)
```

Where s_ave and e_ave are the geometric centers of the 4 toms in the standard and experimental bases respectively

2. Construct the covariance matrix C between s and e

(I used R to do the following calculation)

1)

```
# define standard and experimental matrix
t_s= matrix(c(-0.498,1.364,-0.000,0.000,0.000,0.000,-0.517, -0.772,1.200,1.537,  -
0.000,0.000), nrow=3)
> t_s
     [,1] [,2]   [,3]  [,4]
[1,] -0.498    0 -0.517 1.537
[2,]  1.364    0 -0.772 0.000
[3,]  0.000    0  1.200 0.000
```

[NOTES] here “t_s” refers to the transpose of standard matrix to simplify

```
 e= matrix(c( -11.308,2.285,-0.764,-10.745,0.936,-0.712,-11.601,-0.056,  0.061,-9.353,
0.937,-0.076 ),nrow=4, byrow=T)
> e
      [,1]   [,2]   [,3]
[1,] -11.308  2.285 -0.764
[2,] -10.745  0.936 -0.712
[3,] -11.601 -0.056  0.061
[4,]  -9.353  0.937 -0.076
```

2)

covariance matrix C

```
> c1= 1/3 * (t_s %*% e - (1/4)*t_s %*% c(1,1,1,1) %*% t(c(1,1,1,1)) %*% e )
> c2= 1/3 * (t_s %*% e - (1/4)*t_s %*% c(1,1,1,1) %*% t(c(1,1,1,1)) %*% e )
> c3= 1/3 * (t_s %*% e - (1/4)*t_s %*% c(1,1,1,1) %*% t(c(1,1,1,1)) %*% e )
> c4= 1/3 * (t_s %*% e - (1/4)*t_s %*% c(1,1,1,1) %*% t(c(1,1,1,1)) %*% e )
> c= c1+c2+c3+c4
> c
        [,1]      [,2]      [,3]
[1,]  3.821271 -0.272160  0.5689313
[2,] -0.137472  3.403835 -1.1580267
[3,] -1.358800 -1.730400  0.6940000
```

3. Construct matrix M

```
> m=matrix(c(c[1,1]+c[2,2]+c[3,3], c[2,3]-c[3,2], c[3,1]-c[1,3], c[1,2]-c[2,1],c[2,3]-c[3,2],
c[1,1]-c[2,2]-c[3,3],c[1,2]+c[2,1], c[3,1]+c[1,3], c[3,1]-c[1,3], c[1,2]+c[2,1], -c[1,1]+c[2,2]-
c[3,3], c[2,3]+c[3,2], c[1,2]-c[2,1], c[3,1]+c[1,3], c[2,3]+c[3,2], -c[1,1]-c[2,2]+c[3,3]),nrow=4,
byrow=T )
> m
        [,1]      [,2]      [,3]      [,4]
[1,]  7.9191060  0.5723733 -1.927731 -0.1346880
[2,]  0.5723733 -0.2765633 -0.409632 -0.7898687
[3,] -1.9277313 -0.4096320 -1.111437 -2.8884267
[4,] -0.1346880 -0.7898687 -2.888427 -6.5311060

> eigen(m)
$values
[1]  8.3723137 -0.1451454 -0.2983051 -7.9288633

$vectors
         [,1]       [,2]      [,3]       [,4]
[1,]  0.97470228 -0.09454432  0.1954690 -0.05299722
[2,]  0.07187939 -0.68256395 -0.7187259 -0.11123370
[3,] -0.20976833 -0.62392516  0.6343012 -0.40543382
[4,]  0.02803682  0.36862613 -0.2070916 -0.90578241
```

4. Rotation matrix R

qi corresponds to the largest unit eigenvector

```
> q0=0.97470228
> q1=0.07187939
> q2= -0.20976833
> q3= 0.02803682
```

```
> r=matrix(c(q0*q0+q1*q1-q2*q2-q3*q3, 2*(q1*q2-q0*q3), 2*(q1*q3+q0*q2),
2*(q1*q2+q0*q3),q0*q0-q1*q1+q2*q2-q3*q3, 2*(q2*q3-q0*q1), 2*(q1*q3-
q0*q2),2*(q2*q3+q0*q1),q0*q0-q1*q1-q2*q2+q3*q3), nrow=3,byrow=T)
> r
         [,1]       [,2]       [,3]
[1,] 0.91042237 -0.08481114 -0.4048928
[2,] 0.02449907  0.98809458 -0.1518845
[3,] 0.41295388  0.12835954  0.9016612
```

5. Origin

```
> o=e_ave- s_ave %*% t(r)
> o
         [,1]     [,2]      [,3]
[1,] -10.73679 0.9216302 -0.7163861
```

6. x-,y- and z-axes

```
> t(r)
          [,1]       [,2]      [,3]
[1,]  0.91042237  0.02449907 0.4129539
[2,] -0.08481114  0.98809458 0.1283595
[3,] -0.40489280 -0.15188448 0.9016612
```

7. Validation

.par data

```
C111.A-D218.T:A5.ARG 9ant.mod.pdb +   -10.7995   25.8394
  -10.6263    1.2572    1.4595    8.3168  -24.2450    3.2969
  -10.7362    0.9216   -0.7160
   0.9104    0.0245    0.4130
  -0.0849    0.9881    0.1285
  -0.4050   -0.1521    0.9016
 -10.7447    0.9360   -0.7120
```

Frame origin

.par file (line 2) -10.7362 0.9216 -0.7160

Calculation- -10.73679 0.9216302 -0.7163861

x-, y- and z- axes

.par file

```
0.9104    0.0245    0.4130
-0.0849    0.9881    0.1285
 -0.4050   -0.1521    0.9016
```

Calculation

```
0.91042237  0.02449907 0.4129539
-0.08481114  0.98809458 0.1283595
-0.40489280 -0.15188448 0.9016612
```