5 Some technical details

In our experience, following step-by-step working examples is the best way to understand any nontrivial algorithm. To make things right, every technical detail counts. Over the years, we have received many requests asking for exactly how 3DNA parameters are calculated. Here, we provide details of some key 3DNA components so that interested users could understand them better and possibly apply the same techniques to other related situations.

As an example, we will use the first dinucleotide step `GG/CC` from A-DNA `adh026` (with base sequence `GGGCGCCC`), included in 3DNA distribution, which shows Slide and Roll more clearly.

5.1 Least-squares fitting procedures

3DNA starts with a least-squares procedure to fit a standard base with an embedded reference frame to an observed base structure. It implements a closed-form solution of absolute orientation using unit quaternions first introduced by Horn (1987). This method can be applied when one or both of the structures are perfectly planar. This section is based on the following URL:

```
http://rutchem.rutgers.edu/~olson/jmb/ls_fit.html
```

- Standard reference frame (Olson *et al.*, 2001)

  Using base G as an example, its xyz coordinates in standard reference frame in PDB format are as follows (check BASEPARS directory for other cases):

  ```
  ATOM      1  C1'   G A   1      -2.477   5.399   0.000
  ATOM      2  N9    G A   1      -1.289   4.551   0.000
  ATOM      3  C8    G A   1       0.023   4.962   0.000
  ATOM      4  N7    G A   1       0.870   3.969   0.000
  ATOM      5  C5    G A   1       0.071   2.833   0.000
  ATOM      6  C6    G A   1       0.424   1.460   0.000
  ATOM      7  O6    G A   1       1.554   0.955   0.000
  ATOM      8  N1    G A   1      -0.700   0.641   0.000
  ATOM      9  C2    G A   1      -1.999   1.087   0.000
  ATOM     10  N2    G A   1      -2.949   0.139  -0.001
  ATOM     11  N3    G A   1      -2.342   2.364   0.001
  ATOM     12  C4    G A   1      -1.265   3.177   0.000
  ```

- Least-squares fitting procedure

  Least-squares fitting in 3DNA uses only (available) ring atoms: nine for purines (' N9 '; ' C8 '; ' N7 '; ' C5 '; ' C6 '; ' N1 '; ' C2 '; ' N3 '; '

C4 '), and six for pyrimidines (' N1 '; ' C2 '; ' N3 '; ' C4 '; ' C5 '; ' C6 ').

Using the first G residue in chain A of adh026, denoted thereafter as A_G1, the corresponding 9 ring atoms in standard (sX, sY, sZ) and experimental (eX, eY, eZ) structures are as follows:

```
          sX      sY      sZ           eX      eY      eZ
1  N9   -1.289   4.551   0.000      11.417  -2.904  -4.880
2  C8    0.023   4.962   0.000      10.759  -1.995  -5.662
3  N7    0.870   3.969   0.000      11.469  -0.913  -5.867
4  C5    0.071   2.833   0.000      12.638  -1.108  -5.156
5  C6    0.424   1.460   0.000      13.759  -0.273  -5.036
6  N1   -0.700   0.641   0.000      14.767  -0.848  -4.249
7  C2   -1.999   1.087   0.000      14.663  -2.116  -3.719
8  N3   -2.342   2.364   0.001      13.625  -2.934  -3.830
9  C4   -1.265   3.177   0.000      12.625  -2.328  -4.545
-----------------------------------------------------------
s_ave: -0.6897   2.7827  0.0001  e_ave:12.8580 -1.7132 -4.7716
```

Where s_ave and e_ave are the geometric centers of the nine ring atoms in the standard and experimental bases respectively.

We collect the two sets of coordinates in the 9 x 3 matrices S and E corresponding respectively to the standard and experimental bases. We then construct 3 x 3 covariance matrix (C) between S and E using the following formula:

```
        1                 1
 C = ------- [S' E -  --- S' i i' E]
     N - 1             N


   =
   -0.6849     0.8393   -0.7418
   -2.1919    -0.5683   -0.7888
    0.0001    -0.0002    0.0001
```

Here N, the number of atoms in each base, is 9, and i is an N x 1 column vector consisting of only ones. S' and i' are the transpose of matrix S and column vector i respectively.

From the nine elements of C, we subsequently generate the 4 x 4 real symmetric matrix M using the expression:

```
     _                                                              _
    | c11+c22+c33     c23-c32        c31-c13        c12-c21     |
```

```
M = |   c23-c32      c11-c22-c33      c12+c21       c31+c13     |
    |   c31-c13      c12+c21       -c11+c22-c33     c23+c32     |
    |   c12-c21      c31+c13         c23+c32      -c11-c22+c33  |


  =
  -1.2530    -0.7886     0.7419     3.0312
  -0.7886    -0.1167    -1.3526    -0.7417
   0.7419    -1.3526     0.1165    -0.7889
   3.0312    -0.7417    -0.7889     1.2533
```

The largest eigenvalue of matrix `M` is 3.5896, and its corresponding unit eigenvector (`qi`, `i = 0 -- 3`) is:

```
[ q0    q1     q2     q3 ] = [ 0.5460    -0.2921    0.0524    0.7835 ]
```

The rotation matrix `R` is deduced from the `qi` as:

```
    | q0q0+q1q1-q2q2-q3q3      2(q1q2-q0q3)         2(q1q3+q0q2)      |
R = |    2(q2q1+q0q3)       q0q0-q1q1+q2q2-q3q3     2(q2q3-q0q1)      |
    |    2(q3q1-q0q2)          2(q3q2+q0q1)      q0q0-q1q1-q2q2+q3q3  |


  =
  -0.2331    -0.8862    -0.4004
   0.8249    -0.3983     0.4012
  -0.5150    -0.2368     0.8238
```

Following coordinate transformation with matrix `R`, the translation vector for the least-squares fit between standard and experimental bases is defined as:

```
 o = e_ave - s_ave R' = [15.1632    -0.0362    -4.4678]
```

Here `o` and `R` are the origin and orientation, i.e., the reference frame, of the experimental base. Moreover, `R` is guaranteed to be orthogonal.

The least-squares fitted coordinates (`F`) of the standard base atoms on the experimental structure are then given by:

```
 F = S R' + o
   =
   11.4307    -2.9119    -4.8818
   10.7607    -1.9933    -5.6548
```

```
11.4432    -0.8991    -5.8558
12.6361    -1.1058    -5.1753
13.7706    -0.2678    -5.0319
14.7583    -0.8689    -4.2591
14.6659    -2.1180    -3.6958
13.6138    -2.9091    -3.8208
12.6427    -2.3449    -4.5688
```

Here `S` is the (`N  x  3`) matrix of the original coordinates of the standard base.

The difference matrix (`D`) between `F` and `E`, the (`N  x  3`) matrix of the original coordinates of the experimental base, and the root-mean-square (RMS) deviation between the two structures are found to be:

```
D = E - F
  =
  -0.0137     0.0079     0.0018
  -0.0017    -0.0017    -0.0072
   0.0258    -0.0139    -0.0112
   0.0019    -0.0022     0.0193
  -0.0116    -0.0052    -0.0041
   0.0087     0.0209     0.0101
  -0.0029     0.0020    -0.0232
   0.0112    -0.0249    -0.0092
  -0.0177     0.0169     0.0238


  RMS deviation = sqrt(sum(d^2)/N) = 0.0236
```

5.2   Base reference frames

Following the above least-squares procedure, we get the following for `A_G1`:

```
RMSD = 0.0236
  o = [15.1632    -0.0362    -4.4678]
  R = [
  -0.2331    -0.8862    -0.4004
   0.8249    -0.3983     0.4012
  -0.5150    -0.2368     0.8238]
     where the first column corresponds to x-axis, the 2nd to
        y-axis, and the 3rd to z-axis respectively.
```

Similarly for `B_C8` (i.e., C8 on chain B) which pairs with `A_G1`, we have the following:

```
RMSD = 0.0229
   o = [14.9124    0.2803   -4.7498]
   R = [
   -0.2339    0.9100    0.3422   ==>    -0.2339   -0.9100   -0.3422
    0.7496    0.3930   -0.5326   ==>     0.7496   -0.3930    0.5326
   -0.6191    0.1320   -0.7741   ==>    -0.6191   -0.1320    0.7741]
```

Since their dot product is negative (-0.9884), the z-axis of B_C8 is anti-parallel to that of A_G1. We reverse the y- and z-axes of B_C8 (by a 180-degree rotation around the x-axis) to make its z-axis parallel to that of A_G1, as shown on the right side of R following the arrows.

For A_G2, we have:

```
RMSD = 0.0171
   o = [14.8757    2.9250   -2.4635]
   R = [
   -0.6807   -0.6274   -0.3781
    0.3893   -0.7471    0.5388
   -0.6205    0.2195    0.7528]
```

For B_C7, which pairs with A_G2, we have:

```
RMSD = 0.0213
   o = [14.4982    3.0313   -2.3001]
   R = [
   -0.5797    0.6905    0.4326   ==>    -0.5797   -0.6905   -0.4326
    0.3207    0.6814   -0.6579   ==>     0.3207   -0.6814    0.6579
   -0.7491   -0.2426   -0.6165   ==>    -0.7491    0.2426    0.6165]
```

5.3   General rotation matrix

The general rotation matrix $\mathbf{R}_{\hat{\mathbf{u}}}(\varphi)$ (Eq. [1]) describes a rotation of magnitude φ about an arbitrary unit vector $\hat{\mathbf{u}} = u_1\hat{\mathbf{i}} + u_2\hat{\mathbf{j}} + u_3\hat{\mathbf{k}}$, where $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ are unit vectors along the axes of the local Cartesian frame:

$$\mathbf{R}_{\hat{\mathbf{u}}}(\varphi) = \begin{bmatrix} \cos\varphi + (1-\cos\varphi)u_1^2 & (1-\cos\varphi)u_1u_2 - u_3\sin\varphi & (1-\cos\varphi)u_1u_3 + u_2\sin\varphi \\ (1-\cos\varphi)u_1u_2 + u_3\sin\varphi & \cos\varphi + (1-\cos\varphi)u_2^2 & (1-\cos\varphi)u_2u_3 - u_1\sin\varphi \\ (1-\cos\varphi)u_1u_3 - u_2\sin\varphi & (1-\cos\varphi)u_2u_3 + u_1\sin\varphi & \cos\varphi + (1-\cos\varphi)u_3^2 \end{bmatrix},$$

$$(1)$$

5.4   Base-pair parameters

Given the above base reference frames for A_G1,B_C8 and A_G2,B_C7, we can calculate the base-pair parameters (Shear, Stretch, Stagger, Buckle, Propeller, and Opening). The procedure is exactly the same as for step parameters detailed below. Here we just give the parameters for the

two base-pairs, and users who are interested in knowing the details should work them through and should get exactly the same numbers.

	Shear	Stretch	Stagger	Buckle	Propeller	Opening
A_G1-B_C8	-0.4683	-0.1516	-0.0156	-5.4713	-6.7936	-2.8660
A_G2-B_C7	-0.1643	-0.2112	-0.3299	-4.6532	-9.8008	2.7357

The "middle frame" used in calculating the bp parameters becomes the bp reference frame. For base pair `A_G1-B_C8`, we have:

```
o1 = [15.0378    0.1221   -4.6088]
R1 = [
 -0.2323   -0.8985   -0.3724
  0.7889   -0.3980    0.4682
 -0.5689   -0.1851    0.8013]
```

For base pair `A_G2-B_C7`, we have:

```
o2 = [14.6869    2.9781   -2.3818]
R2 = [
 -0.6319   -0.6594   -0.4072
  0.3583   -0.7144    0.6010
 -0.6873    0.2339    0.6877]
```

These two reference frames are used in the next two sections to calculate step and helical parameters.

5.5 Step parameters

Given the 2 bp reference frames above, the following procedures are used to calculate step parameters (Shift, Slide, Rise, Tilt, Roll and Twist). It follows CEHS definition, as detailed in SCHNAaP.

1. Hinge axis is the cross product between z1 (3rd column of `R1`) and z2 (3rd column of `R2`):

```
[-0.3724    0.4682    0.8013] x [-0.4072    0.6010    0.6877]
  = [-0.1596   -0.0702   -0.0332]. When normalized, it becomes:
 [-0.8993  -0.3953  -0.1870]
```

Geometrically, hinge axis is the intersection line between the two base pair planes.

2. The RollTilt angle (Γ, i.e., net bending angle) is the magnitude between z1 and z2, which is given by their dot product: $\Gamma = 10.2221$ Å.

3. Now we rotate R2 by ($-0.5 * \Gamma = -5.1111$) degrees around the hinge axis (see Eq. [1]):

```
   R_hinge(-5.1111) = [
      0.9992   -0.0152    0.0359
      0.0181    0.9966   -0.0798
     -0.0345    0.0804    0.9962]
    R2' = R_hinge(-5.1111) * R2 = [
     -0.6616   -0.6396   -0.3914
      0.4005   -0.7426    0.5368
     -0.6340    0.1984    0.7475]
```

Similarly, we rotate R1 by $(+0.5 * \Gamma)$ degrees along the hinge axis:

```
   R1' = R_hinge(+5.1111) * R1 = [
     -0.1982   -0.8986   -0.3914
      0.7441   -0.3978    0.5368
     -0.6381   -0.1848    0.7475]
```

By definition, the z-axes of R1' and R2' (the third column) are the same, i.e., after symmetric rotations, we have perfectly aligned the z-axes of the two bps.

4. The x-, y- and z-axes of the "middle frame" are simply the average between those of R1' and R2', and by definition, they are orthogonal. The origin of the bp is the geometric average of o1 and o2. For the above case, we have the "middle frame" as follows:

```
   Rm = [
    -0.4490   -0.8033   -0.3914
     0.5977   -0.5955    0.5368
    -0.6642    0.0071    0.7475]
   om = [14.8624    1.5501   -3.4953]
```

5. The translational parameters (Shift, Slide, and Rise) are simply the projections of the vector linking from o1 to o2 onto the x-, y-, and z-axes of the "middle frame":

```
   [Shift Slide Rise] = (o2 - o1) * Rm
                      = [-0.3509    2.8561    2.2270] * Rm
                      = [0.3853   -1.4033    3.3349]
```

6. Twist is the angle from y1 to y2 (or from x1 to x2 of the "rotated" R1' and R2' matrices respectively). Its sign is defined with reference to the "middle frame" z-axis (the 3rd column of Rm), following right-handed rule for positive Twist angle. More specifically, the magnitude of the angle between y1 and y2 is: $33.5296°$. The signe is determined by:

```
   (y1' x y2') . z =
   [-0.2162  0.2965  0.4129] . [-0.3914  0.5368  0.7475]
                   = 0.5524
```

So Twist is positive.

7. The phase angle (ϕ) is the angle from the hinge-axis to the "middle frame" y-axis (2nd column of Rm). By definition, hinge-axis lies in the xy-plane of the "middle-frame" since it is perpendicular to the z-axis of the "middle-frame". The phase angle also has a sign associate with it, determined in the same way as Twist shown above. In this case, $\phi = +16.9598°$.

8. Roll is defined as:

$$\Gamma * \cos(\phi) = 10.2221 * \cos(16.9598 * \pi/180) = 9.7776°.$$

Similarly, Tilt is defined as:

$$\Gamma * \sin(\phi) = 10.2221 * \sin(16.9598 * \pi/180) = 2.9818°.$$

9. Overall, the six step parameters are:

```
    Shift      Slide      Rise       Tilt       Roll      Twist
   0.3853    -1.4033     3.3349     2.9818     9.7776    33.5296
```

5.6   Local helical parameters

The geometric approach described below gives exactly the same numerical values as those from the RNA (Babcock *et al.*, 1994). The pivot point issue does not apply here: it is only used in defining the base pair reference frame. Given the reference frames of the two base pairs, the procedure used in 3DNA to calculate the local helical parameters (x-displacement, y-displacement, helical rise, inclination, tip, and helical rise) is analogous to the one detailed above for step parameters, by using a tip-inclination combination.

Please note that to define a local helical axis, we need **two** base-pair reference frames. 3DNA finds the single-helical axis (which is actually dx × dy) that brings 1 to coincide with 2 by a Helical Twist angle. The position which this helix passes through is defined by Chasles' theorem as detailed in Figures 12 & 13 of Babcock *et al.* (1994). The calculation of x-displacement, y-displacement, tip and inclination is then exactly as described in SCHNAaP (Lu *et al.*, 1997a).

To make the above point clear, let's use A1-A2-A3 triplet as an example. First, A1-A2 define a local helical axis and a set of local base-pair helical parameters are calculated. In 3DNA, these parameters are defined in a symmetric manner that bp A1:T1 and bp A2:T2 have exactly the same values. Similarly, step A2-A3 defines another set of local base-pair helical parameters. Thus bp A2:T2 has two sets of helical parameters associated with it depending on its context, i.e., either with bp A1:T1 or with bp A3:T3. Moreover, the local Helical Rise and Helical Twist are directly related to a dinucleotide step. These are the reasons that "Local base-pair helical parameters" as given in 3DNA refer to base-pair steps. As a matter of fact, the schematic diagrams illustrating the local helical parameters (x-displacement, y-displacement, inclination and tip) as given in 3DNA website and this user's manual were based on two perfectly overlapped base pair blocks.

Using the first GG/CC dinucleotide step in adh026 as an example, the detailed precedure is as follows:

1. The local helical axis is defined by the cross product of (x2 - x1) and (y2 - y1):

```
[-0.3997   -0.4307   -0.1183] x [0.2391   -0.3164   0.4190] =
[-0.2179   0.1392   0.2294], which when normalized, gives:
h = [-0.6303   0.4026   0.6638]
```

2. Local helical frame of base-pair 1:

   - TipInclination angle (Ψ, in magnitude) is the angle between unit vectors h and z1, and is $17.2279°$.

   - Hinge axis is defined by a cross product from h to z1, normalized to give:

     ```
     [0.0399 0.8707 -0.4902].
     ```

     It lies in the xy-plane of local helical frame given below.

   - Rotate R1 through the above hinge axis by negative TipInclination angle (i.e., $-17.2279°$) will align the resultant z-axis with h, which gives us the local helical reference frame:

     ```
     H1 = R_hinge(-17.2279) * R1 = [
         -0.1880   -0.7532   -0.6303
          0.7504   -0.5242    0.4026
         -0.6337   -0.3973    0.6638]
     ```

3. Similarly, the local helical frame for base pair 2:

   - TipInclination angle (Ψ, in magnitude) is the angle between unit vectors h and z2, and is $17.2279°$. Note that by (symmetric) definition, it is the same as for base pair 1.

   - Hinge axis is defined by a cross product from h to z2, normalized to give:

     ```
     [-0.4121 0.5511 -0.7256]
     ```

   - Rotate R2 through the above hinge axis by negative TipInclination angle (i.e., $-17.2279°$) will align the resultant z-axis with h, which gives us the local helical reference frame:

     ```
     H2 = R_hinge(-17.2279) * R2 = [
         -0.5861   -0.5091   -0.6303
          0.3139   -0.8599    0.4026
         -0.7470    0.0381    0.6638]
     ```

4. The "middle helical frame" is the average of H1 and H2, and by definition, it is orthogonal.

   ```
   Hm = [
       -0.4058   -0.6618   -0.6303
        0.5580   -0.7256    0.4026
       -0.7238   -0.1883    0.6638]
   ```

5. Helical twist is the angle from y(H1) to y(H2) (or from x(H1) to x(H2)) of the local helical reference frames defined above, with reference to the local helical axis h for sign determination: $35.0103°$.

6. Helical rise is the projection of the vector linking `o1` to `o2` onto the local helical axis `h`: 2.8493 Å.

7. Use **base pair 1** to calculate tip and inclination (base pair 2 gives the same result):

   - Phase angle ψ is defined from hinge axis (`[0.0399 0.8707 -0.4902]`) to positive tip axis (2nd column of `H1`, `[-0.7532 -0.5242 -0.3973]`) with reference the local helical axis `h` (`[-0.6303 0.4026 0.6638]`) for sign control: 106.9598°.

   - Tip is defined as: $\Psi * \cos(\psi) = 17.2279 * \cos(106.9598 * \pi/180) = -5.0254°$

   - Inclination is defined as: $\Psi * \sin(\psi) = 17.2279 * \sin(106.9598 * \pi/180) = 16.4787°$

8. Get the origin of local helical frame of base-pair 1. It is based on Chasles' theorem as used in `RNA` (Babcock *et al.*, 1994, see Figure 13, page 141 for a detailed illustration). The procedure implemented in 3DNA is as follows:

   - Vector `AB` is defined by:

     ```
     AB =  [o2 - o1] - Helical_Rise * h =
      [-0.3509 2.8561 2.2270] - 2.8493 * [-0.6303 0.4026 0.6638]
       = [1.4451    1.7089    0.3357]
     ```

   - Vector `AD` is defined by an angle of

     ```
     90 - 0.5 * Helical_Twist = 72.4948
     ```

     from vector `AB` with reference to `h`:

     ```
     AD = R_h(72.4948) * AB' = [-0.5182    1.6306    -1.4812]'
     ```

     `R_h(72.4948)` is the rotation matrix along `h` (Eq. [1]) by an angle of 72.4948°. `AB'` changes row vector `AB` to a column vector to be compatible with the `3-by-3` rotation matrix. When normalized, the `AD` vector is: `[-0.2290 0.7205 -0.6545]`

   - The magnitude of `AD` is defined by:

     ```
     AD_magnitude =
     0.5 * AB_magnitude / sin(0.5 * Helical_Twist * pi / 180)
             = 0.5 * 2.2631 / 0.3008 = 3.7618
     ```

   - We then get "D", the position vector where the local helical axis passes through. It is the origin of local helical frame of base pair 1.

     ```
     o1_h = o1 + AD_magnitude * AD =
      [15.0378 0.1221 -4.6088] + 3.7618 * [-0.2290 0.7205 -0.6545]
     = [14.1764    2.8326    -7.0709]
     ```

     The origin of the local helical frame of base pair 2 is given by:

     ```
     o2_h = o1_h + Helical_Rise * h = [12.3805  3.9798  -5.1797]
     ```

The origin of the "middle helical frame" is the average of `o1_h` and `o2\_h`: `[13.2785 3.4062 -6.1253]`

9. X-displacement and y-displacement are defined as the projections of the vector from `o1_h` to `o1` onto the x- and y-axis of local helical frame of base pair 1 (`H1`):

```
[o1 - o1_h] * H1(:, 1:2) =
([15.0378 0.1221 -4.6088] - [14.1764 2.8326 -7.0709]) * H1(:, 1:2)
= [-3.7562 -0.2063]
```

So x-displacement = -3.7562 Å and y-displacement = -0.2063 Å.

Using base pair 2 gives *exactly the same values* due to symmetric definition.

10. Overall, the six local helical parameters are:

X-disp	Y-disp	Rise	Incl.	Tip	Twist
-3.7562	-0.2063	2.8493	16.4787	-5.0254	35.0103

5.7 Rebuilding based on local step parameters

The local step parameters defined in 3DNA are rigorous and thus reversible. Given a set of step parameters, the relative position and orientation of the two base pairs can be exactly reproduced by a rebuilding procedure as detailed in `SCHNAaP`/`SCHNArP` (Lu *et al.*, 1997a; Lu *et al.*, 1997b). Here we provide step-by-step working example so users can understand the algorithm better.

Using the six step parameters of the first GG/CC step in `adh026` as calculated above:

Shift	Slide	Rise	Tilt	Roll	Twist
0.3853	-1.4033	3.3349	2.9818	9.7776	33.5296

1. We first Twist base-pairs 2 and 1 about the "middle frame" z-axis (which at this stage is coincident with the z-axes of both base-pair reference frames) through $+\Omega/2$ and $-\Omega/2$ respectively:

$$\text{T2} = \mathbf{R}_z(+\Omega/2), \qquad \text{T1} = \mathbf{R}_z(-\Omega/2) \tag{2}$$

Note that to be consistent with previous literatures, the symbol used for Twist here is still Ω (not ω), which is 33.5296°. The rotation matrix $\mathbf{R}_z$ corresponds to the Eq. [1] with unit vector `[0 0 1]`. Similarly, $\mathbf{R}_x$ corresponds to `[1 0 0]` and $\mathbf{R}_y$ corresponds to `[0 1 0]`. Here, the "middle frame" is the "global" reference frame and corresponds to an identity matrix.

2. The RollTilt angle (Γ) is given by:

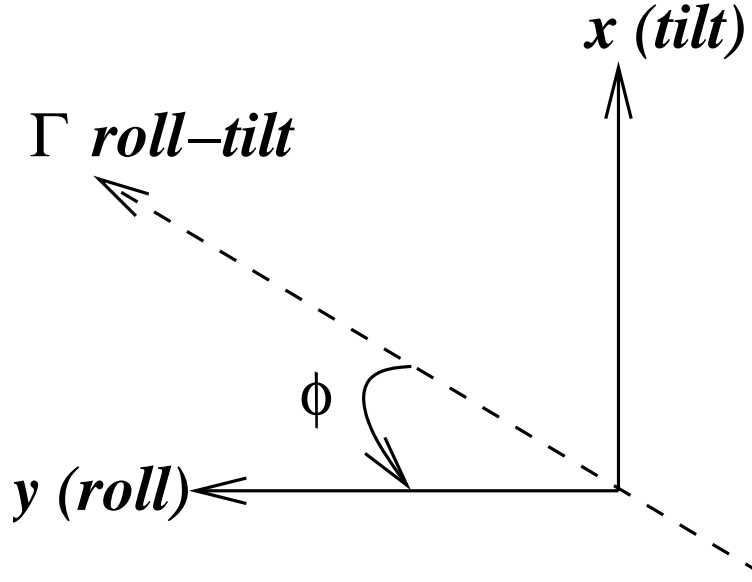$$\Gamma = \sqrt{\rho^2 + \tau^2} = \sqrt{9.7776^2 + 2.9818^2} = 10.2221° \tag{3}$$

Figure 7: Roll-tilt combination

3. The Roll-Tilt axis is determined by the values of Roll and Tilt, and it lies in the x-y plane of the "middle frame". For the above case, it is `[0.2917 0.9565 0]`, when normalized.

   The angle from the Roll-Tilt axis (`[0.2917 0.9565 0]`) to the "middle frame" y-axis (Roll) (`[0 1 0]`) is called ϕ. Its sign is determined with reference to the "middle frame" z-axis (`[0 0 1]`), with values in range from $-180°$ to $180°$. In this case, it is $16.9598°$.

4. We now rotate T2 and T1 through $+\Gamma/2$ and $-\Gamma/2$ respectively about the Roll-Tilt axis to give the final orientation of the two base-pairs with respect to the "middle frame". Analytically, the rotation about the Roll-Tilt axis can be expressed in terms of three rotations with reference to "middle frame" axes: First a rotation of angle ϕ around z-axis to align Roll-Tilt axis with y-axis, then a rotation of angle $\pm\Gamma/2$ around y-axis, and finally a rotation of angle $-\phi$ around z-axis to bring the Roll-Tilt axis back. Overall, the required rotations can be written as:

$$
\begin{aligned}
\mathrm{T2}^{\mathrm{m}} &= [\mathbf{R}_z(-\phi)\,\mathbf{R}_y(+\Gamma/2)\,\mathbf{R}_z(+\phi)]\,\mathrm{T2} \\
&= \mathbf{R}_z(-\phi)\,\mathbf{R}_y(+\Gamma/2)\,\mathbf{R}_z(\phi+\Omega/2) & (4) \\
\mathrm{T1}^{\mathrm{m}} &= [\mathbf{R}_z(-\phi)\,\mathbf{R}_y(-\Gamma/2)\,\mathbf{R}_z(+\phi)]\,\mathrm{T1} \\
&= \mathbf{R}_z(-\phi)\,\mathbf{R}_y(-\Gamma/2)\,\mathbf{R}_z(\phi-\Omega/2) & (5)
\end{aligned}
$$

5. To get the correct positions of the origins of the reference frame of base-pair 2 and 1, we apply the translation vectors $(+D_x/2, +D_y/2, +D_z/2)$ and $(-D_x/2, -D_y/2, -D_z/2)$ respectively. Both translations are defined with reference to the "middle frame". Here D_x, D_y, D_z represent Shift, Slide and Rise.

6. The set of equations described above for the synthesis of a dinucleotide step defines the positions and orientations of the two base-pairs in a dinucleotide step with respect to the

"middle frame". However, by means of a set of simple matrix transformations, we can change the reference frame from the "middle frame" to the reference frame of the base-pair 1. The orientation and position of base-pair 2 and the "middle frame" with respect to base-pair 1 are now given by the following equations:

$$\text{Rm} = [T1^m]^{-1} = \mathbf{R}_z(\Omega/2 - \phi)\, \mathbf{R}_y(\Gamma/2)\, \mathbf{R}_z(\phi) \tag{6}$$

$$\text{R2} = [T1^m]^{-1}\, T2^m = \mathbf{R}_z(\Omega/2 - \phi)\, \mathbf{R}_y(\Gamma)\, \mathbf{R}_z(\Omega/2 + \phi) \tag{7}$$

$$\text{o2} = [D_x\ D_y\ D_z]\, \text{Rm}' \tag{8}$$

Where R_m' means the transpose of the "middle frame" expressed with reference to base pair 1.

Using the above set of step parameters, we have the following:

```
Rm = [
    0.9537    -0.2873     0.0891
    0.2885     0.9575    -0.0003
   -0.0852     0.0260     0.9960]
R2 = [
    0.8204    -0.5436     0.1775
    0.5524     0.8336    -0.0006
   -0.1476     0.0985     0.9841]
o2 = [0.3853  -1.4033    3.3349] * Rm' = [1.0677   -1.2336    3.2524]
```

Expressed in terms of the experimental coordinate reference frame of base pair 1, we get the reference frame for base pair 2 which are the same as shown in section 5.4: "Base-pair parameters"

```
R1_exp = [
   -0.2323    -0.8985    -0.3724
    0.7889    -0.3980     0.4682
   -0.5689    -0.1851     0.8013]


R2_exp = R1_exp * R2 = [
   -0.6319    -0.6594    -0.4072
    0.3583    -0.7144     0.6010
   -0.6873     0.2339     0.6877]


o1_exp = [15.0378     0.1221    -4.6088]


o2_exp = o1_exp + o2 * R1_exp' = [14.6869     2.9781    -2.3818]
```

Xiang-Jun Lu (3dna.lu@gmail.com)

5.8 Rebuilding based on local helical parameters

The local helical parameters defined in 3DNA are rigorous and thus reversible. Given a set of helical parameters, the relative position and orientation of the two base pairs can be exactly reproduced by a rebuilding procedure. Here we provide step-by-step working example so users can understand the algorithm better.

Using the six helical parameters of the first GG/CC step in `adh026` as calculated above:

```
  X-disp     Y-disp      Rise      Incl.        Tip      Twist
 -3.7562    -0.2063     2.8493    16.4787    -5.0254    35.0103
```

There are four frames for a dinucleotide step consisting of base-pairs 1 and 2: the base-pair frames $\mathbf{R}_1$ and $\mathbf{R}_2$ and their corresponding helical frames $\mathbf{R}_{1h}$ and $\mathbf{R}_{2h}$. $\mathbf{R}_{1h}$ and $\mathbf{R}_1$ are related by x-displacement, y-displacement, inclination and tip, so are $\mathbf{R}_{2h}$ with $\mathbf{R}_2$. $\mathbf{R}_{1h}$ and $\mathbf{R}_{2h}$ are related by *helical* rise and *helical* twist.
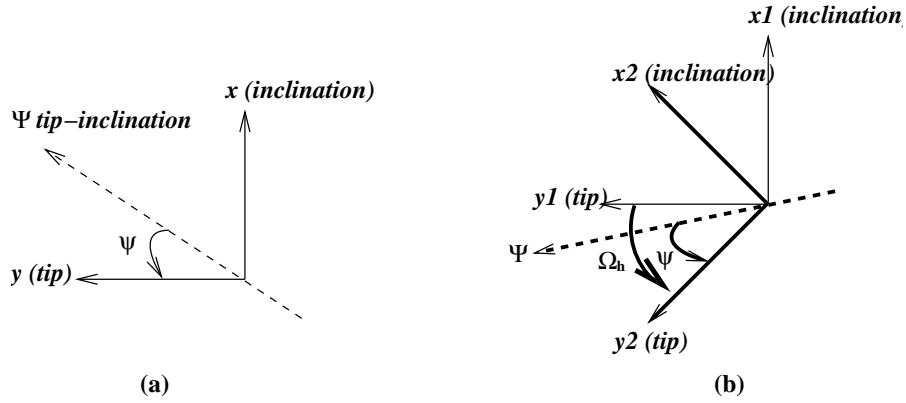

Figure 8: Combination of Tip and Inclination, and calculation of helical Twist.

1. $\mathbf{R}_{1h}$ as reference ($1h$ in superscript)

   Starting with $\mathbf{R}_{1h}$ as the reference for the dinucleotide step concerned, then $\mathbf{R}_{1h}^{1h}$ is the identity matrix $\mathbf{I}$:

   $$\mathbf{R}_{1h}^{1h} = \mathbf{I} \tag{9}$$

   $\mathbf{R}_{2h}^{1h}$ is got by the helical twist:

   $$\mathbf{R}_{2h}^{1h} = \mathbf{R}_z(\Omega_h) \tag{10}$$

   $\mathbf{R}_1^{1h}$ is got by a combined rotation of magnitude Ψ $(= \sqrt{\eta^2 + \theta^2})$ along tip-inclination axis (Fig. 8 (a)). Thus Ψ is the exact angle between base-pair normal and the local helical helix. Let the angle from tip-inclination axis to tip axis be ψ, the rotation can be expressed as a rotation about z-axis by angle ψ, followed by a rotation about y-axis by angle Ψ, and then a rotation about z-axis by angle $-\psi$ to bring the axis back:

$$\mathbf{R}_1^{1h} = \mathbf{R}_z(-\psi)\,\mathbf{R}_y(\Psi)\,\mathbf{R}_z(\psi) \tag{11}$$

As shown in Fig. 8 (b), $\mathbf{R}_2^{1h}$ is got by a helical twist, and then a rotation about the tip-inclination axis of $\mathbf{R}_{2h}^{1h}$ (which is related to the tip y-axis of $\mathbf{R}_{1h}$ by the helical twist Ω_h):

$$\mathbf{R}_2^{1h} = [\mathbf{R}_z(\Omega_h - \psi)\,\mathbf{R}_y(\Psi)\,\mathbf{R}_z(\psi - \Omega_h)]\,\mathbf{R}_z(\Omega_h) = \mathbf{R}_z(\Omega_h - \psi)\,\mathbf{R}_y(\Psi)\,\mathbf{R}_z(\psi) \tag{12}$$

2. $\mathbf{R}_1$ as reference

   With Eqs. 9 to 12, we can do some simple matrix transformations to make $\mathbf{R}_1$ as the reference:

$$
\begin{aligned}
\mathbf{R}_1 &= \mathbf{I} & (13) \\
\mathbf{R}_{1h} &= [\mathbf{R}_1^{1h}]^{-1} = \mathbf{R}_z(-\psi)\,\mathbf{R}_y(-\Psi)\,\mathbf{R}_z(\psi) & (14) \\
\mathbf{R}_{2h} &= [\mathbf{R}_1^{1h}]^{-1}\,\mathbf{R}_{2h}^{1h} = \mathbf{R}_z(-\psi)\,\mathbf{R}_y(-\Psi)\,\mathbf{R}_z(\psi + \Omega_h) & (15) \\
\mathbf{R}_2 &= [\mathbf{R}_1^{1h}]^{-1}\,\mathbf{R}_2^{1h} \\
&= \mathbf{R}_z(-\psi)\,\mathbf{R}_y(-\Psi)\,\mathbf{R}_z(\psi)\,\mathbf{R}_z(\Omega_h - \psi)\,\mathbf{R}_y(\Psi)\,\mathbf{R}_z(\psi) \\
&= \mathbf{R}_z(-\psi)\,\mathbf{R}_y(-\Psi)\,\mathbf{R}_z(\Omega_h)\,\mathbf{R}_y(\Psi)\,\mathbf{R}_z(\psi) & (16)
\end{aligned}
$$

   Equ. 16 describes the orientation of base-pair 2 with reference to 1.

3. Position vector of base-pair 2 with reference to 1

   - From base-pair 1 origin to base-pair 1 helical origin: $-[dx\ dy]\,[x_{1h}\ y_{1h}]'$
   - From base-pair 1 helical origin to 2 helical origin: $dz \cdot z_{1h}' = dz \cdot z_{2h}'$
   - From base-pair 2 helical origin to base-pair 2 origin: $+[dx\ dy]\,[x_{2h}\ y_{2h}]'$

   Combining the above three items, we have the position (base-pair origin) of 2 relative to 1 as follows:

$$\mathbf{o_2} = [dx\ dy]\,\{[x_{2h}\ y_{2h}]' - [x_{1h}\ y_{1h}]'\} + dz \cdot z_{1h}' \tag{17}$$

   Note that $x_{1h}, y_{1h}, z_{1h}, x_{2h}, y_{2h}, z_{2h}$ are 3×1 column vectors. $'$ means the transpose of a vector to change it to a row vector.

4. Using the example where inclination = $16.4787°$ and tip = $-5.0254°$, we have $\Psi = \sqrt{\eta^2 + \theta^2} = 17.2279°$. The tip-inclination axis lies in the x-y plane of local helical frame of base pair 1: [16.4787 -5.0254 0]; which when normalized gives: [0.9565 -0.2917 0]. The angle from the tip-inclination axis ([0.9565 -0.2917 0]) to the base pair 1 helical y-axis (tip) ([0 1 0]) with reference to the helical z-axis ([0 0 1]) is $\psi = 106.9598°$.

   Using equations 15 to 17, we having the following:

```
R1h = [
    0.9962   -0.0125    0.0864
   -0.0125    0.9590    0.2833
   -0.0864   -0.2833    0.9551]


R2h = [
    0.8087   -0.5818    0.0864
    0.5399    0.7926    0.2833
   -0.2333   -0.1825    0.9551]


R2 = [
    0.8204   -0.5436    0.1775
    0.5524    0.8336   -0.0006
   -0.1476    0.0985    0.9841]


o2 = [1.0677   -1.2336    3.2524]
```

Compared with the numbers based on step parameters, it is clear that Eqs. 16 and 7, Eqs. 17 and 8 are equivalent.

5.9 Relation between local helical and step parameters

To refer the orientation and position of one base-pair relative to the other, 6 parameters (3 rotations and 3 translations) are required. One set of such parameters is (Shift, Slide, Rise, Tilt, Roll and Twist), and the other set is (X-displacement, Y-displacement, Helical Rise, Inclination, Tip and Helical Twist).

Obviously these two sets should be completely reversible/dependent: from any one set you can get the other, rigorously. You can verify this point using `step_hel`, a utility program in 3DNA. Graphically this is best illustrated by the Calladine-Drew A to B transition model by introducing uniform Roll and Slide values at each dinucleotide step. The key point is that by introducing Roll, you also get Inclination, and with Slide, you get X-displacement.

The `rebuild` program in 3DNA can construct a DNA structure using either set of these parameters. Examples of such input files (e.g., `bp_step.par` and `bp_helical.par`) can be generated by `analyze` (`Examples/Analyze_Rebuild` directory.)

We have two sets of simple equations:

$$\theta/\eta = -\tau/\rho$$

$$2\cos\Omega_h = \cos\Omega(1 + \cos\Gamma) - (1 - \cos\Gamma)$$