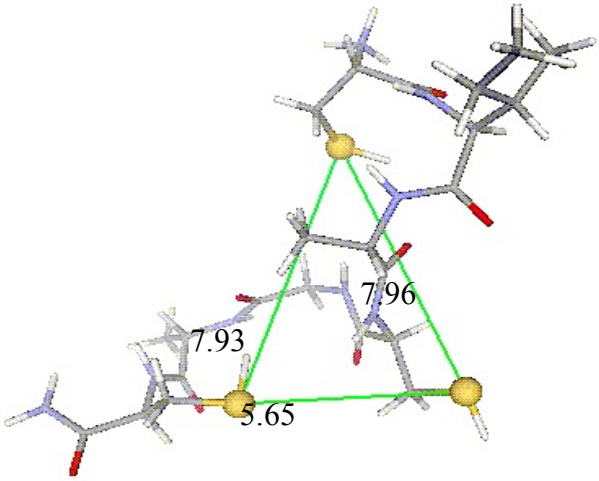
Threading, [4Fe-4S] Nesting Ceiling

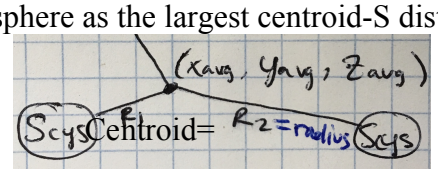
Identifies frames where the cysteine S-S distances are less than 8 Å and the backbone of the peptide is not transecting the cysteine S-S-S plane and disrupting the nest. This provides a maximum number of possible frames for [4Fe-4S] nesting, including all non-threading frames where the cysteine S-S distances are less than 8 Å.



CIACGAC optimized: frame #1585

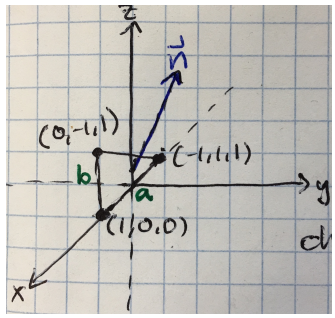
Algorithm: For each frame…

1. Read in the cysteine S coordinates, continue only if all cysteine S-S distances are less than 8 Å.
    a. Cysteine S-S distances in FdM7 1DUR are 6.54, 6.03 and 6.32 Å for S-S_{NX}, S-S_{XC} and S-S_{NC} respectively for an average distance of 6.30 Å. Since each frame is only a snapshot in time taken every 1 ps, 8 Å was chosen as the maximum distance allowed.
2. Define a spherical model of the nest.
    a. Calculate the centroid of the nest as the average of the cysteine S coordinates.
    b. Calculate the radius of the sphere as the largest centroid-S distance.

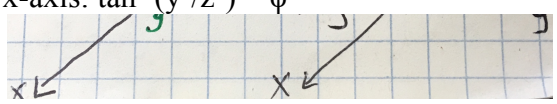


3. Find and store all the backbone atoms within the sphere. Store connectivity data to identify the length of each backbone fragment within the sphere.
4. Rotate the peptide so the cysteine S-S-S plane is aligned with the x-y plane so the normal vector to the S-S-S plane is aligned with the z-axis.
    a. Calculate and normalize the normal vector of the cysteine S-S-S plane

        a = coordinates of cysteine S #1 — coordinates of cysteine S #2
        b = coordinates of cysteine S #1 — coordinates of cysteine S #3
        || **normal** || = **a** x **b** / magnitude of the normal vector



b. Rotate normal vector to align with z-axis (x=0 and y=0).
   i. Rotate around z-axis: $\tan^{-1}(x/y) = \theta$
   ii. Rotate around x-axis: $\tan^{-1}(y'/z') = \phi$



c. Transform the coordinates of each atom within the sphere.
   i. Shift coordinates so the centroid is at the origin.
   ii. Rotate coordinates θ around the z-axis

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

   iii. Rotate coordinates ϕ around the x-axis

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

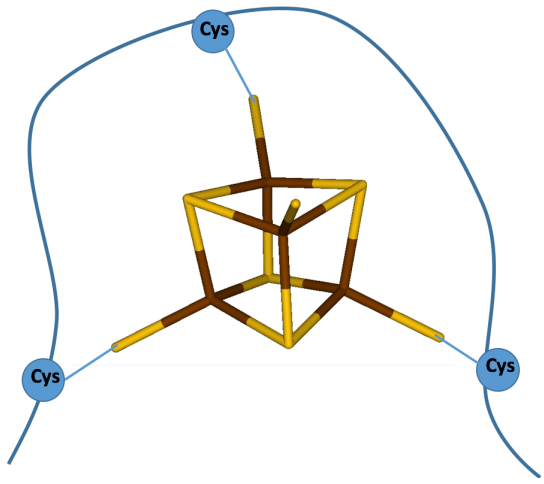
5. Check if backbone atoms transect the z-axis in the rotated peptide. If + and – z coordinates exist for any backbone fragment within the sphere, then the frame is threading.

Input Files:
- input.txt
- frame coordinate file (.txyz)

Output Files:
- CIACGACxrd-threading-nest.txt (list of frames non-threading and with cysteine S-S distances less than 8 Å)

[4Fe-4S] Nesting

Identifies frames where the peptide is available to bind a [4Fe-4S] cluster, where all the cysteine S-S distances are less than 8 Å, a spherical model of the cluster fits within the cysteine S atoms, there are no non-cysteine atoms within the sphere, and there are no atoms within a 1.5 Å expanded sphere to exclude weak interactions.



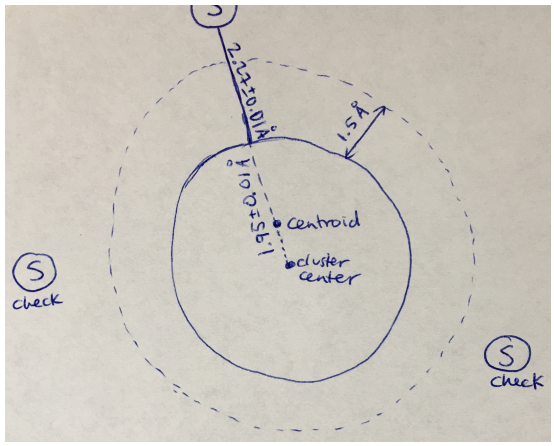
5D8V

Cluster fitting:

1. Ellipsoid model: fit coordinates of 5D8V [4Fe-4S] cluster with an ellipsoid model. The radii of the model are a = b = c = 1.96 Å. Therefore, the cluster can be more accurately fit with a spherical model based on crystallographic data.
2. Spherical model: for the [4Fe-4S] cluster in 5D8V, the maximum radius using Fe-S distances within the cluster is 1.95 ± 0.01 Å.

Algorithm: For each frame…

1. Read in the cysteine S coordinates, continue only if all cysteine S-S distances are less than 8 Å.
    a. Cysteine S-S distances in FdM7 1DUR are 6.54, 6.03 and 6.32 Å for S-S_{NX}, S-S_{XC} and S-S_{NC} respectively for an average distance of 6.30 Å. Since each frame is only a snapshot in time taken every 1 ps, 8 Å was chosen as the maximum distance allowed and no minimum distance was defined.
2. Pick a cysteine S and define the center of the nest based on the position that primary S.
    a. Calculate the centroid of the nest as the average of the cysteine S coordinates.
    b. Define the center of the cluster by moving along the primary cysteine S-centroid vector a distance of the cysteine S-Fe bond length plus the radius of the sphere.
        i. The average cysteine S-Fe bond length in 5D8V is 2.27 ± 0.01 Å and the radius of the sphere (described above) is 1.95 ± 0.01 Å.
        ii. Ratio = (sphere radius + S_{cys}-Fe bond length) / (S_{cys}-centroid distance)
            = (1.95 ± 0.01 Å + 2.27 ± 0.01 Å) / (S_{cys}-centroid distance)
            = (4.22 ± 0.02 Å) / (S_{cys}-centroid distance)
        iii. Cluster center = ratio(centroid) + (1 – ratio)(S_{cys})



3. Check remaining cysteine S distances from the cluster center. If they are both within 4.22 ± 0.02 Å (sphere radius + S_{cys}-Fe bond length) from the cluster center, then continue.
4. Repeat steps 2 and 3 with each of the remaining cysteine S atoms acting as primary. If all three cysteine S atoms pass step 3, then continue.
5. Check sphere for any atom besides the cluster binding cysteine S, C, or H atoms. If no atoms found, then frame is [4Fe-4S] cluster nesting.
    a. Redefine the sphere origin as the average cluster center from step 2biii.
    b. Expand the radius of the sphere by 1.5 Å to allow hydrogen bonds but exclude ionic bonds to the cluster.

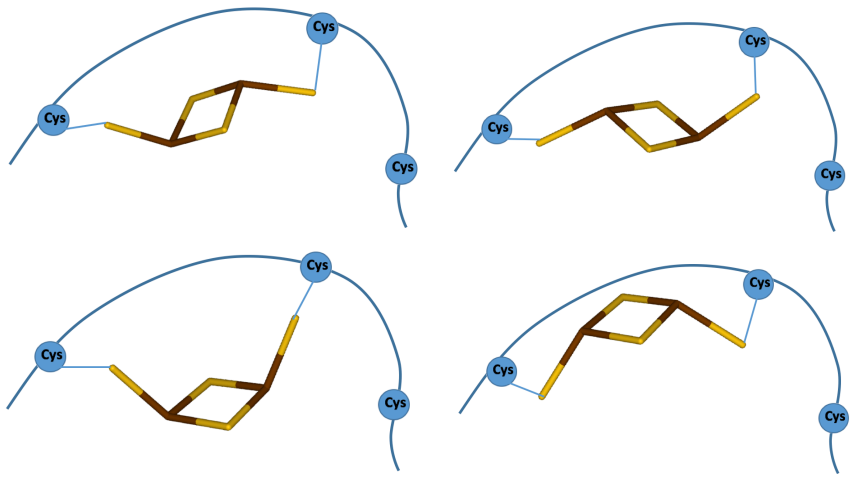
Input Files:
- input.txt
- frame coordinate file (.txyz)

Output Files:
- CIACGACxrd-4FeS-nest.txt (list of frames [4Fe-4S] cluster nesting)

[2Fe-2S] Nesting

Identifies frames where the peptide is available to bind at least one [2Fe-2S] cluster, where at least one cysteine S-S distance is within 5-7 Å, a quad sphere model or an ellipsoid model of the cluster fits within those cysteine S atoms, there are no non-cysteine atoms within the defined cluster space, and there are no atoms within a 1.5 Å expanded model to exclude weak interactions.

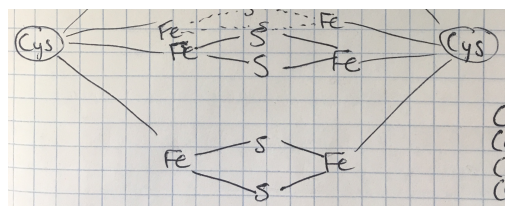
- For Cx_nCx_mC there are N-terminal cysteine (N), central cysteine (X) and C-terminal cysteine (C).
- Single cluster formed between NX, XC, NC
- Two clusters formed between NX+XC, NX+NC, and XC+NC
- Three clusters formed between NX+XC+NC



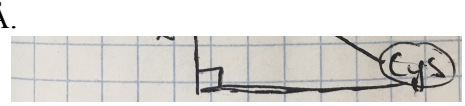
Coordination isomers for the NX single cluster model.

Cluster Fitting; quad sphere model (less restrictive)

Using a spherical model of the cluster, rotate the model around the cysteine S-S vector placing it in four possible locations.

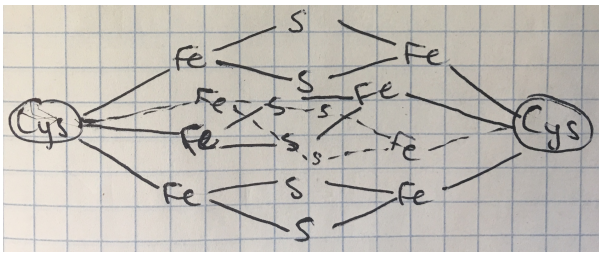


1. Fit the [2Fe-2S] cluster in 3WCQ with a spherical model, the maximum radius using Fe-Fe distances within the cluster is 1.76 ± 0.01 Å.
    a. The trans binding motif for a [2Fe-2S] cluster will sit in the center of the four spheres modelling the cluster's possible positions. However, expansion of the radius of the spheres by 1.5 Å to allow hydrogen bonds but exclude ionic bonds to the cluster, preclude the necessity of a fifth sphere centered at the centroid. The combination of the four spheres included in the model cover the area around the centroid with an overlap diameter of 2.82 Å.
2. Using the bond lengths and angles of the [2Fe-2S] cluster in 3WCQ, the displacement of the spheres from the center is the average ½Fe-Fe distance between the two possible cis binding motifs with cysteine S, 1.85 ± 0.01 Å.



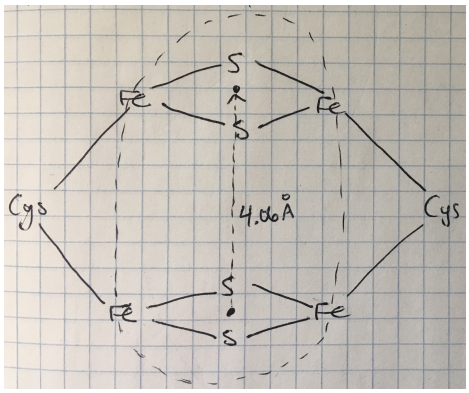
Cluster Fitting; ellipsoidal model (more restrictive)

Place the [2Fe-2S] cluster from 3WCQ in extremes of the cis binding motif and fit using an ellipsoidal model.

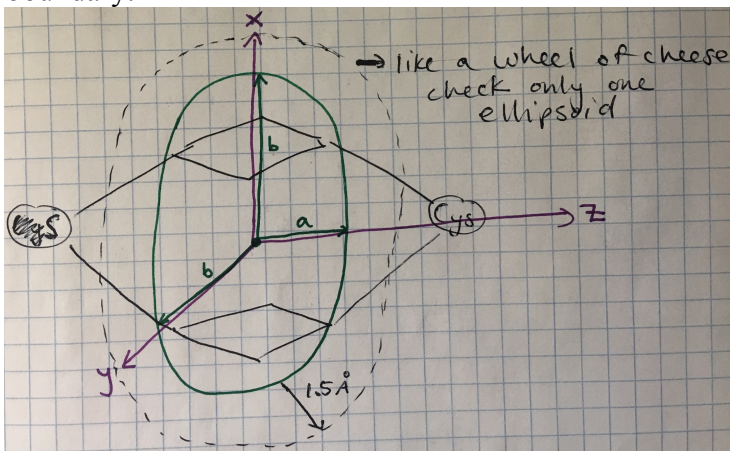


1. Using the bond lengths and angles of the [2Fe-2S] cluster in 3WCQ, the maximum displacement of the cluster from the center is the maximum ½Fe-Fe distance between the two possible cis binding motifs with cysteine S, 2.03 Å.

2. Using the coordinates of the [2Fe-2S] cluster in 3WCQ, two clusters placed 180˚ apart at a distance of 4.06 Å to form a "cube".



3. Fit the "cube" with an ellipsoid model with radii of $r_x = r_y = 2.8120$ Å and $r_z = 1.8506$ Å. The formula of the ellipsoid in coefficient form is $0.1264x^2 + 0.1264y^2 + 0.2920z^2 = 1$
   a. The trans binding motifs for a [2Fe-2S] cluster will sit in the center of the ellipsoid model and are fully incorporated by the ellipsoid with a maximum of 0.2 Å leeway from the ellipsoid boundary.



Algorithm: For each frame…

1. Read in the cysteine S coordinates, continue only if at least one cysteine S-S distance is less than 7 Å.
    a. Cysteine S-S distances in 3WCQ for the cis binding motif are 5.86 and 6.58 Å, while for the trans binding motif, the distances are 6.66 and 5.15 Å. Since each frame is only a snapshot in time taken every 1 ps, 7 Å was chosen as the maximum distance allowed to include both binding motifs. Similarly, 5 Å was chosen as the minimum distance allowed.
2. For each cysteine S pairing NX, XC and NC, continue only if the S-S distance for that cysteine pair is greater than 5 Å and less than 7 Å.
3. Define the centers of each sphere in the quad sphere model and the center of the ellipsoid in the ellipsoidal model.

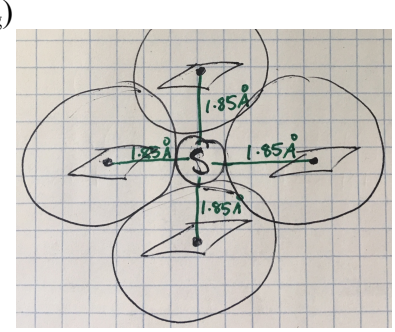
    Quad Spheres Model:

    a. Calculate the centroid of the model as the average coordinates of the cysteine S atoms in the chosen pair.

        Centroid = (x_{avg}, y_{avg}, z_{avg})

    b. The centers of each sphere are offset from the center of the model by 1.85 Å in the positive and negative directions along both the x and y axes.

        $Center_1$ = (x_{avg}+1.85 Å, y_{avg}, z_{avg})
        $Center_2$ = (x_{avg}—1.85 Å, y_{avg}, z_{avg})
        $Center_3$ = (x_{avg}, y_{avg}+1.85 Å, z_{avg})
        $Center_4$ = (x_{avg}, y_{avg}—1.85 Å, z_{avg})



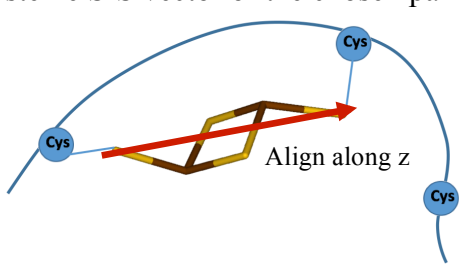
Looking down cysteine S-S bond:

    Ellipsoid Model:

    a. Define the center of the ellipsoid as the average coordinates of the cysteine S atoms in the chosen pair.

        Centroid = (x_{avg}, y_{avg}, z_{avg})

4. Rotate the peptide so the cysteine S-S vector of the chosen pair is aligned with the z-axis.

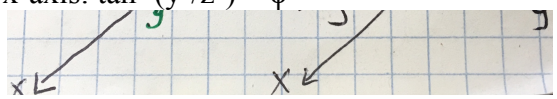


    a. Calculate and normalize the cysteine S-S vector.

        vector = coordinates of cysteine S — average coordinates of S atoms in pair
        || **normal** || = **vector** / magnitude of the vector

    b. Rotate the normalized cysteine S-S vector to align with the z-axis (x=0 and y=0).

i. Rotate around z-axis: $\tan^{-1}(x/y) = \theta$
ii. Rotate around x-axis: $\tan^{-1}(y'/z') = \phi$



c. Transform the coordinates of the entire peptide.
   i. Shift coordinates so the centroid is at the origin.
   ii. Rotate coordinates θ around the z-axis

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

   iii. Rotate coordinates ϕ around the x-axis

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

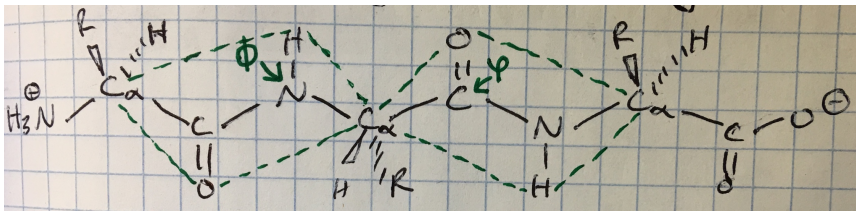
5. Check model (spheres or ellipsoid) for any atom besides the cluster binding cysteine S, C, or H atoms. If no atoms found, then frame is [2Fe-2S] cluster nesting.
   d. Expand the radius of the spheres or the radii of the ellipsoid by 1.5 Å to allow hydrogen bonds but exclude ionic bonds to the cluster. The radii of the spheres are now 3.26 Å and the radii of the ellipsoid are now $r_x = r_y = 4.31$ Å and $r_z = 3.35$ Å.
6. Repeat steps 2-5 for each cysteine S pairing NX, XC and NC.

Input Files:
- input.txt
- frame coordinate file (.txyz)

Output Files:
- CIACGACxrd-2FeS-spheres-nest.txt or CIACGACopt-2FeS-ellipsoid-nest.txt (list of frames [2Fe-2S] cluster nesting)
- CIACGACxrd-2FeS-spheres-matrix.txt or CIACGACopt-2FeS-ellipsoid-matrix.txt (matrix of frames by binding motif NX, XC, NC, NX+XC, NX+NC, XC+NC and NX+XC+NC for [2Fe-2S] cluster nesting)

Torsion Angles

Calculates the torsion angles of the peptide, excluding the terminal residues, and creates a Ramachandran histogram plot.



Per Trajectory

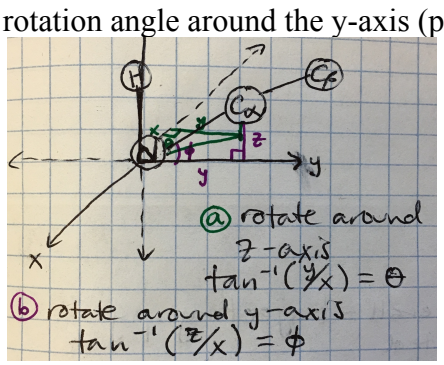
Creates a Ramachandran plot for the torsion angles of all the frames in a trajectory.

Per Peptide

Creates a Ramachandran plot for the torsion angles of all the frames in all the trajectories for a given peptide.

Algorithm (per trajectory): For each frame…

1. Moving along the peptide from the N to C-terminal, store the coordinates for the backbone atoms for each pair of residues.
2. Calculate ϕ (H-N-C_α-CO dihedral) for the residue pair.
   a. Translate the coordinates of the backbone H, N, C_α, and CO so that the N atom is at the origin.
   b. Rotate the coordinates so that the C_α is on the x-axis.
      i. Calculate rotation angle around the z-axis (green): $\tan^{-1}(y/x) = \theta$
      ii. Calculate rotation angle around the y-axis (purple): $\tan^{-1}(z'/x') = \omega$



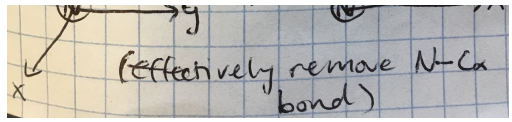
      iii. Rotate coordinates θ around the z-axis

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

      iv. Rotate coordinates ω around the y-axis

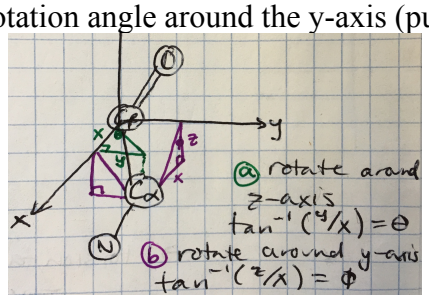
$$R_y(\omega) = \begin{bmatrix} \cos\omega & 0 & \sin\omega \\ 0 & 1 & 0 \\ -\sin\omega & 0 & \cos\omega \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

   c. Convert coordinates to 2D system, effectively removing the N- C_α bond. Change y coordinates to x and change z coordinates to y.



d. Measure ϕ as H'-N'-CO' angle.

3. Calculate φ (N-C_α-C=O dihedral) for the residue pair.
   a. Translate the coordinates of the backbone N, C_α, C, and O so that the carbonyl C atom is at the origin.
   b. Rotate the coordinates so that the C_α is on the x-axis.
      i. Calculate rotation angle around the z-axis (green): $\tan^{-1}(y/x) = \theta$
      ii. Calculate rotation angle around the y-axis (purple): $\tan^{-1}(z'/x') = \omega$



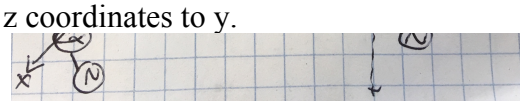
      iii. Rotate coordinates θ around the z-axis

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

      iv. Rotate coordinates ω around the y-axis

$$R_y(\omega) = \begin{bmatrix} \cos\omega & 0 & \sin\omega \\ 0 & 1 & 0 \\ -\sin\omega & 0 & \cos\omega \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

   c. Convert coordinates to 2D system, effectively removing the C_α-C bond. Change y coordinates to x and change z coordinates to y.



   d. Measure φ as N'-C'=O' angle.

4. Repeat steps 2 and 3 for each pair of residues in the peptide fragment, excluding the N and C-terminal residues.
5. Loop through all calculated torsion angles and create histogram for the trajectory.
   a. Use 50 bins per axis with 7.2° per bin.

b. Count the number of data points in each bin.
c. Take the center of the bin as the coordinate (ϕ_{center}, φ_{center}, n).

Algorithm (per peptide):

1. Read in the torsion histogram coordinates for each trajectory of a given peptide.
2. Sum the z-coordinates for each bin to get the total number of torsion angles in a given bin.

Input Files:

- input.txt
- frame coordinate file (.txyz)

Output Files:

- CIACGACxrd-torsion-histogram.txt (coordinates for torsion angles histogram of trajectory (ϕ_{center}, φ_{center}, n))
- CIACGACxrd-dihedrals.txt (list of torsion angles for each frame in trajectory)
- CIACGAC-torsion-histogram.txt (coordinates for overall torsion angles histogram of peptide (ϕ_{center}, φ_{center}, n_{total}))

Nesting Frequency and Rate

Calculates the cumulative percentage of nesting frames and the frequency of nesting throughout a trajectory for the [4Fe-4S] nesting ceiling, [4Fe-4S] and [2Fe-2S] nesting definitions.

Nesting Frequency Algorithm: for each frame…

1. Check if frame is [4Fe-4S] or [2Fe-2S] (ellipsoid model) nesting.
2. If nesting, count number of consecutive nesting frames.

Nesting Rate Algorithm: for each frame…

1. Check if frame is non-threading, [4Fe-4S] or [2Fe-2S] (ellipsoid model) nesting.
2. If nesting or non-threading, increase total count.
3. Calculate the rate of nesting or non-threading up to the current frame.

   rate = (total nesting or non-threading frames / frame number) * 100

Input Files:

- input.txt
- CIACGACxrd-threading-nest.txt
- CIACGACxrd-4FeS-nest.txt
- CIACGACxrd-2FeS-ellipsoid-nest.txt

Output Files:

- CIACGACxrd-nestFrequency_time.txt (frequency of consecutive nesting frames by frame number for [4Fe-4S] and ellipsoid model [2Fe-2S] nesting)
- CIACGACxrd-nestRate_time.txt (rate of nesting percentage by frame number for the [4Fe-4S] nesting ceiling, [4Fe-4S] and ellipsoid model [2Fe-2S] nesting)

S-S Distances and Histogram Integration

Calculates the cysteine S-S distances for each frame of the trajectory, creates a histogram of the S-S distances and calculates the total relative time the cysteine pairs are available for nesting by integrating the histograms from 0-8 Å.

S-S Distances Algorithm: for each frame…

1. Find each cysteine S potentially involved in cluster nesting.
2. Calculate the distances between each cysteine S pair.
    a. For Cx_nCx_mC there are N-terminal cysteine (N), central cysteine (X) and C-terminal cysteine (C) for cysteine S pairings of NX, XC, NC.
    b. S-S' distance = $\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}$

Input Files:

- input.txt
- frame coordinate file (.txyz)

Output Files:

- CIACGACxrd-Sdistances.txt (S-S distances for each cysteine S pairing NX, XC, NC)

S-S Distances Histogram and Integration Algorithm: for each frame…

1. Loop through each 0.5 Å bin from 0 to 30 Å.
2. Count the number of data points in each bin for each cysteine S pairing NX, XC, NC.
3. Integrate each cysteine S pairing's histogram curve from 0-8 Å.

$h =$ bin size (0.5 Å)
$n =$ total number of bins (60)
$f(x) =$ percentage of frames in trajectory per bin

    a. Right-hand Rectangle Rule: underestimate because histogram most likely monotonically increasing between 0 and 8 Å, 2nd order error $\mathcal{O}(h^2)$.

$$I_{Rectangle} = h\sum_{i=1}^{n} f(x_i)$$

    b. Simpson's Rule: composite rule derived from integrating a 3rd order Lagrange interpolating polynomial, 4th order error $\mathcal{O}(h^4)$.

$$I_{Simpson} = \frac{h}{3}\left[f(x_0)+f(x_n)+2\sum_{i=2}^{n/2} f(x_{2i-2})+4\sum_{i=1}^{n/2} f(x_{2i-1})\right]$$

Input Files:

- input.txt
- CIACGACxrd-Sdistances.txt

Output Files:

- CIACGACxrd-Sdistance-histogram.txt (coordinates for S-S distance histogram of each cysteine S pairing NX, XC, NC)

- CIACGACxrd-Sdistance-histogram-integration.txt (Right-hand and Simpson's Rule integration of histogram curves, temporary file added to analysis summary then deleted)

Crossover

Identifies pairs of frames between two trajectories (same or different) where the conformation of the peptide fragment is the "same" based on the torsion angles of the backbone and the cysteine S-S distances.

- This analysis assumes the side chains of the peptide fragment move more rapidly than the backbone atoms. Thus, if the backbones are in the "same" conformation then the sidechains move enough to also be in the "same" conformation.

Step 1: create input files for step 2 and step 3.

For each frame in both trajectories, list the sum of the dihedral angles of the backbone and the sum of the cysteine S-S distances.

Input Files:

- Input-step1.txt
- CIACGACopt-Sdistances.txt
- CIACGACopt-dihedrals.txt
- CIACGACxrd-Sdistances.txt
- CIACGACxrd-dihedrals.txt

Output Files:

- CIACGACopt-sums.txt (sums of backbone dihedral angles and S-S distances per frame)
- CIACGACxrd-sums.txt (sums of backbone dihedral angles and S-S distances per frame)

Step 2: filter frame pairs by loose comparison.

For each frame pairing, compare the sum of the backbone dihedral angles and the sum of the cysteine S-S distances. For each surviving frame pairing, an empty file is created with those indices.

Input Files:

- Input-step2.txt
- CIACGACopt-sums.txt
- CIACGACxrd-sums.txt

Output Files:

- 8421-337.optxrd (indices of filtered frame pairs, empty file moved into directory CIACGACoptxrd-filtered)
- CIACGACoptxrd-filtered.txt (list of filtered frame pairs)

Step 3: check each surviving frame pairing for "sameness".

For each surviving frame pair from step 2, compare the absolute value of each backbone dihedral angle (excluding the terminal residues) and each cysteine S-S distance.

Input Files:

- Input-step3.txt
- CIACGACoptxrd-filtered.txt

- CIACGACopt-Sdistances.txt
- CIACGACopt-dihedrals.txt
- CIACGACxrd-Sdistances.txt
- CIACGACxrd-dihedrals.txt

Output Files:

- CIACGACoptxrd-correlated.txt (list of correlated frame pairs)