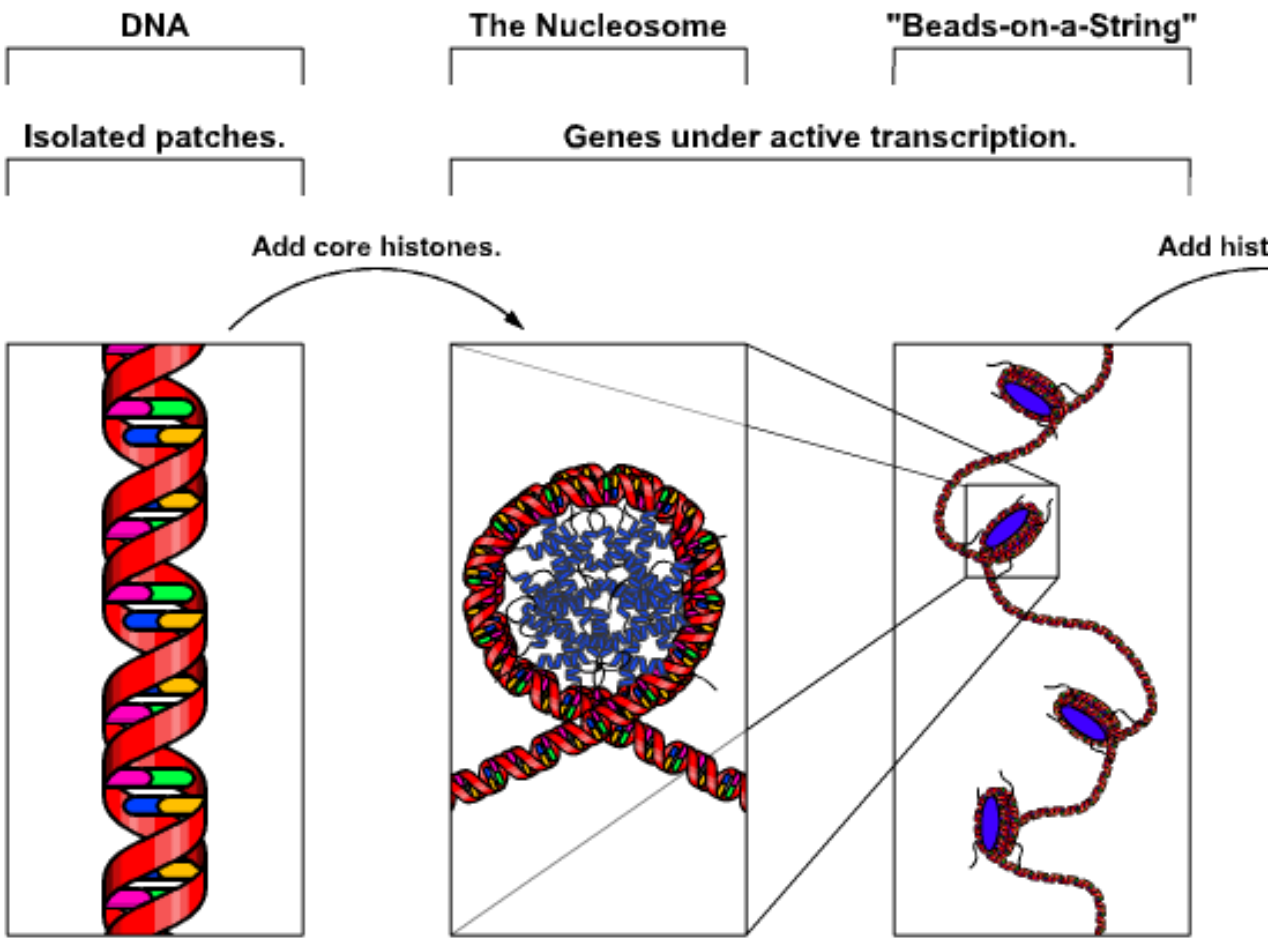
Using Mathematical Tools to Analyze Chromatin Structures

Ondrej Maxian
DIMACS REU
Olson Research Group
July 13, 2017

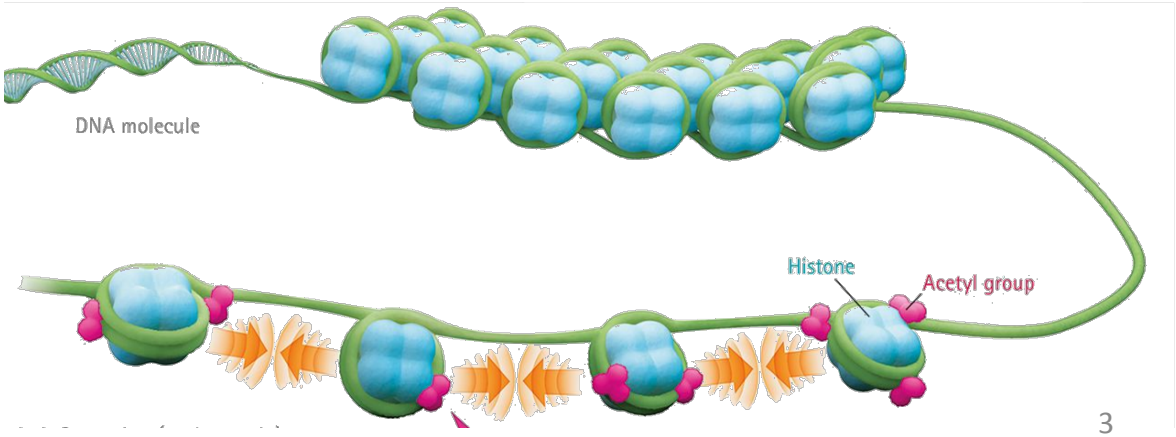
Chromatin



Wikipedia

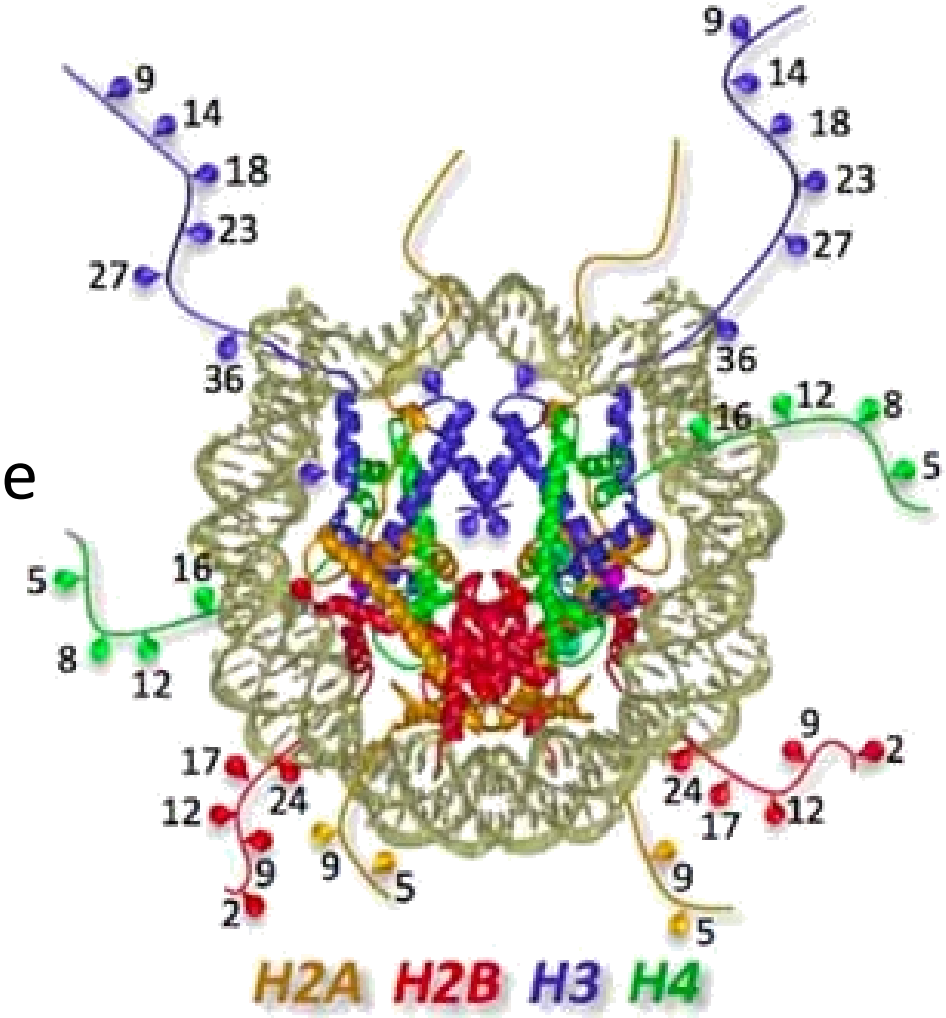
Why are structures important?

- "Beads on a string" structures relate to expression of genes
- Key questions
  - How are nucleosomes packed in loose chromatin fibers (in solution)?
  - What about crystal structures?



Core nucleosome

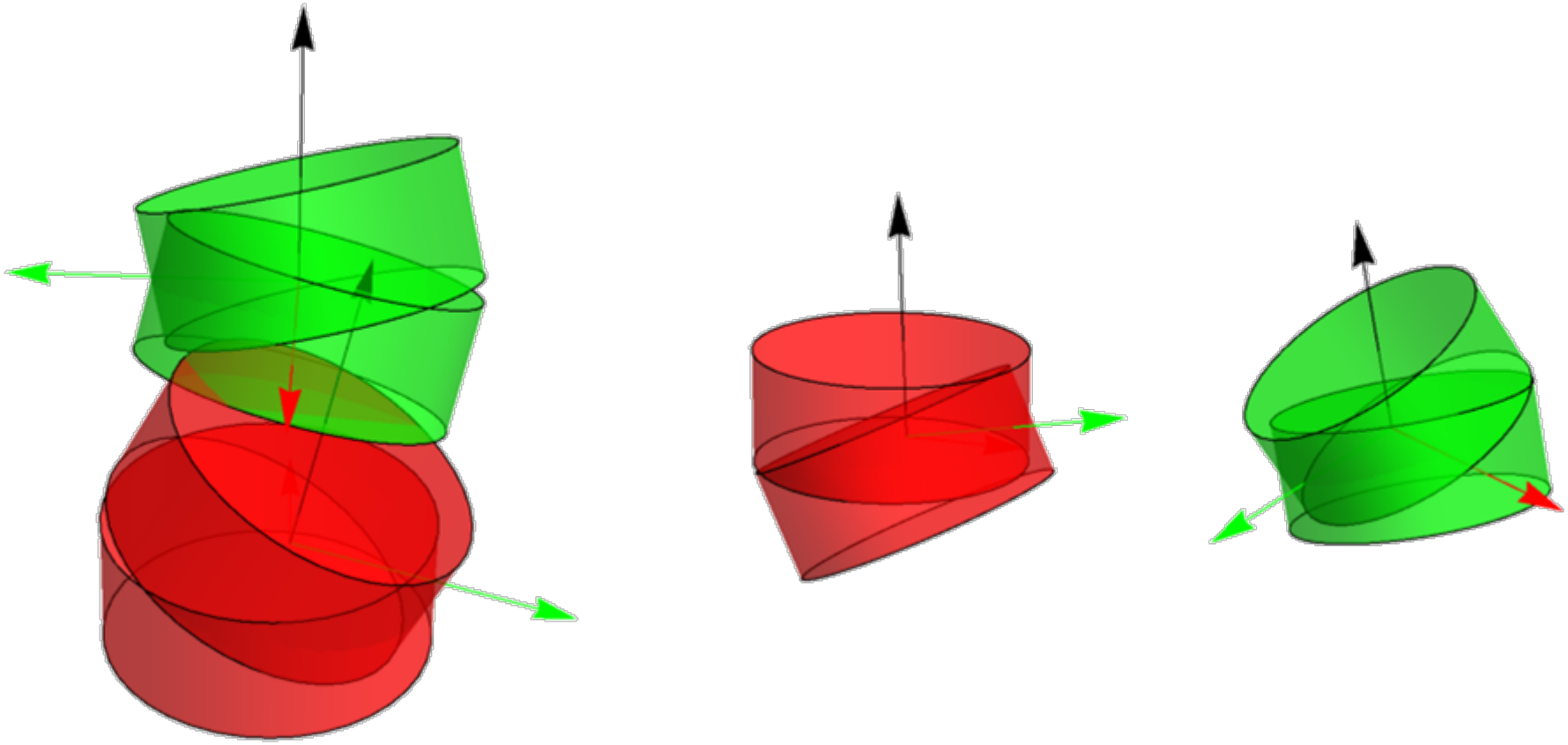
- Core histones and tails H2A,H2B,H3,H4
- Surrounded by DNA
- Locations of phosphate, centers of base pairs can be used to determine coordinate frame



Wikipedia

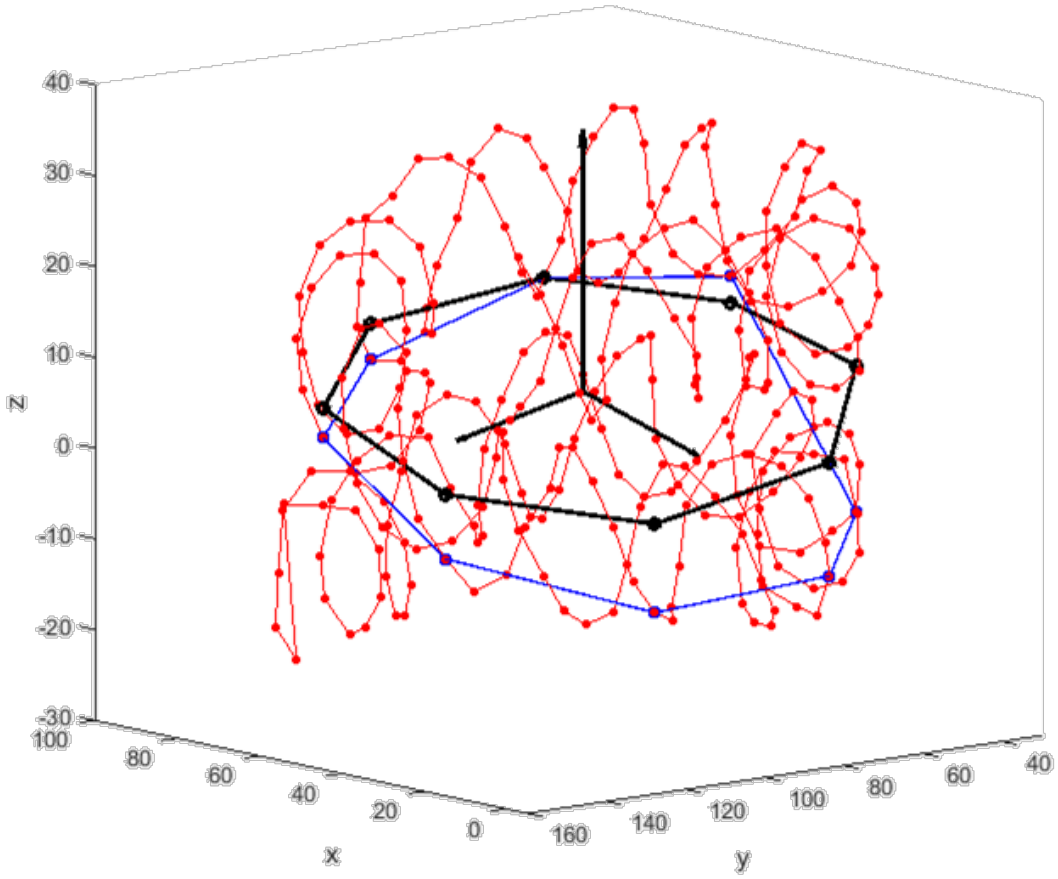
Tools to Analyze

- Interaction head on or side to side



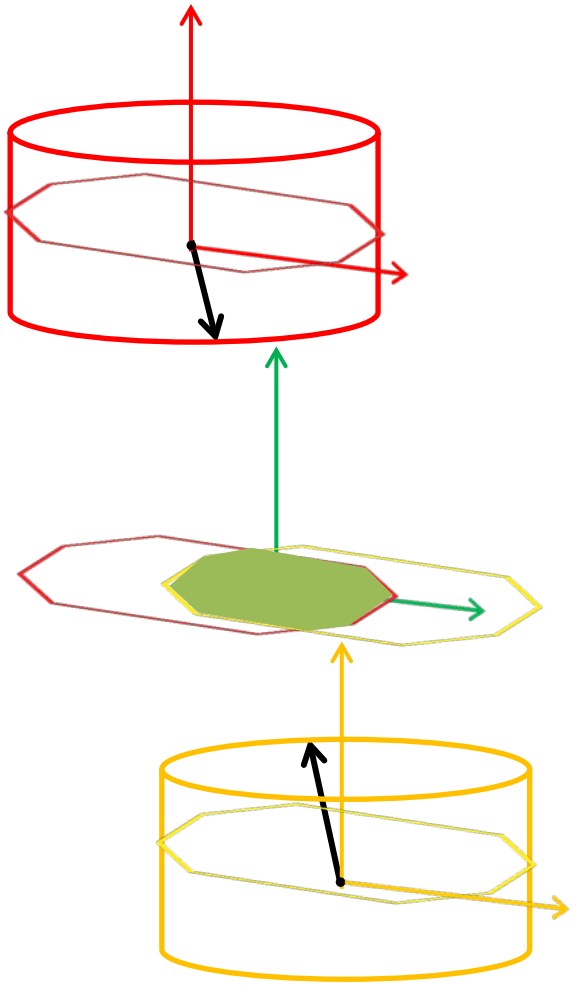
Computing overlap - steps

1. Use phosphate or some atoms to find reference frame (and polygon in ref. frame)
   – The same for all nucleosomes



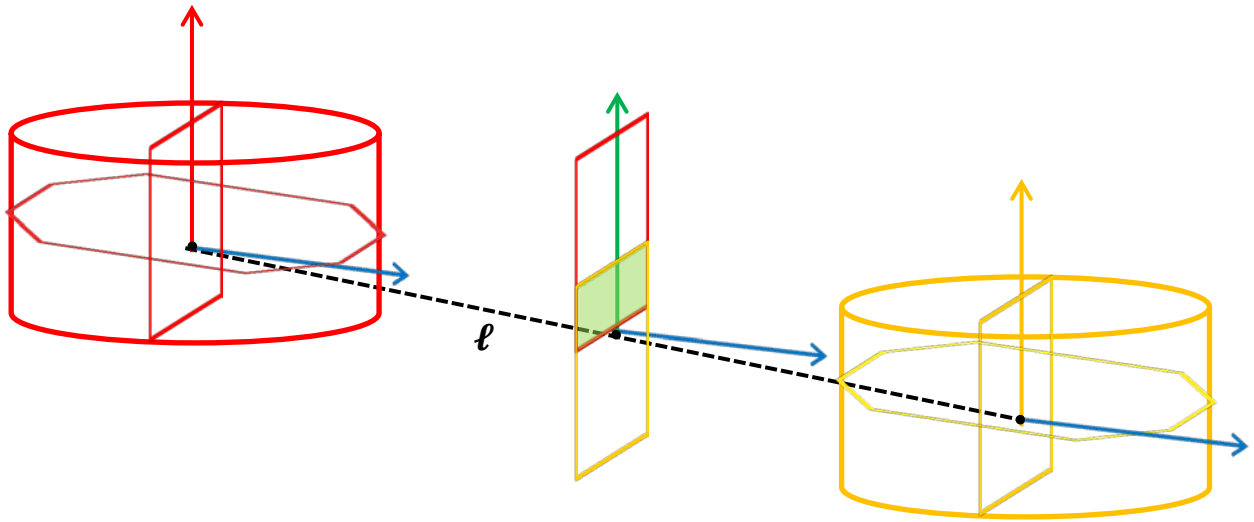
Computing overlap – steps

2. Given 2 nucleosomes, compute mid-frame
3. Project polygons onto mid-frame
4. Find overlap region & area
5. Draw unit vector to other origin to see which histone it points to



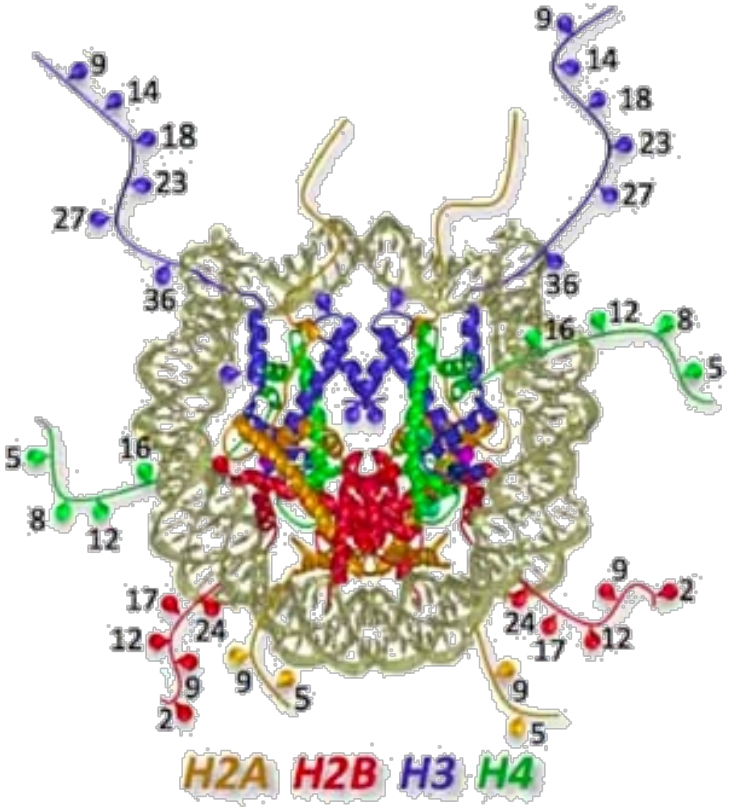
Computing side overlap - steps

1. Draw line ℓ from one origin to another
2. Remove part normal to nucleosome to get 3 vector basis (other from cross product)
3. Lay down polygon (rectangle in this case)
4. Get mid-frame
5. Project and get area
6. Blue vector points to DNA bp coordinate and histone tail



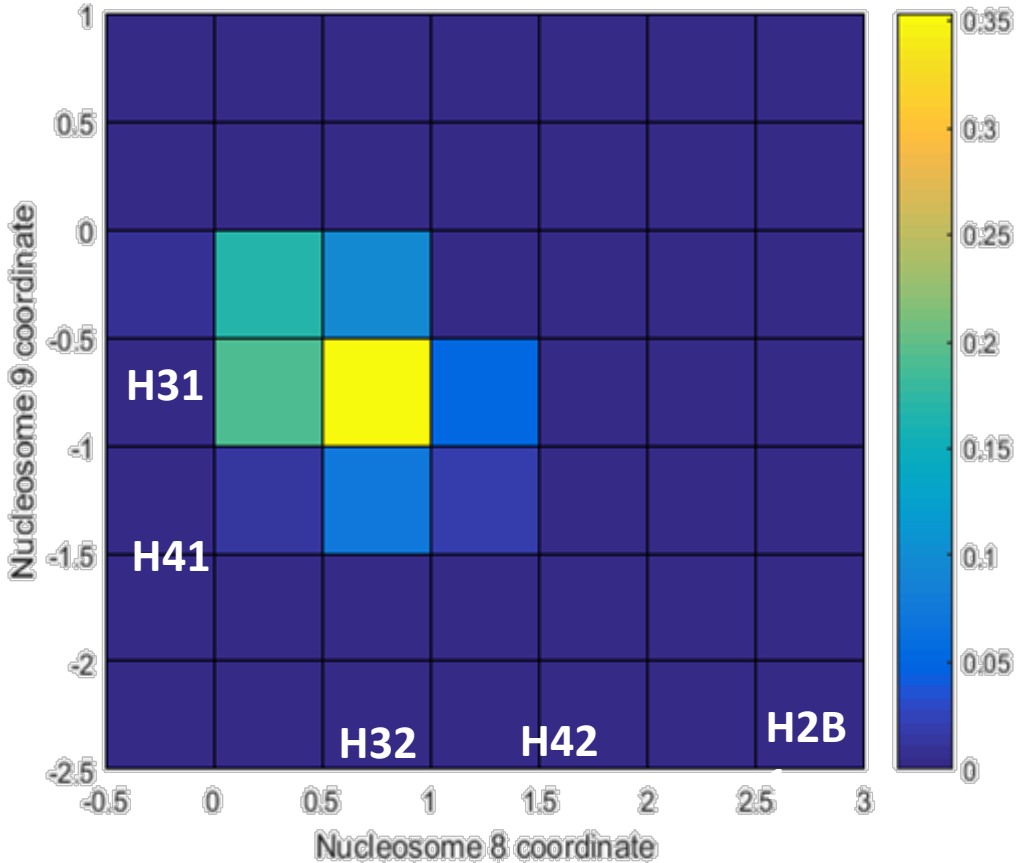
Analysis – Simulated Structures

- MCMC simulations of structures with and without histone tails
- Found H2A/H2B tails irrelevant



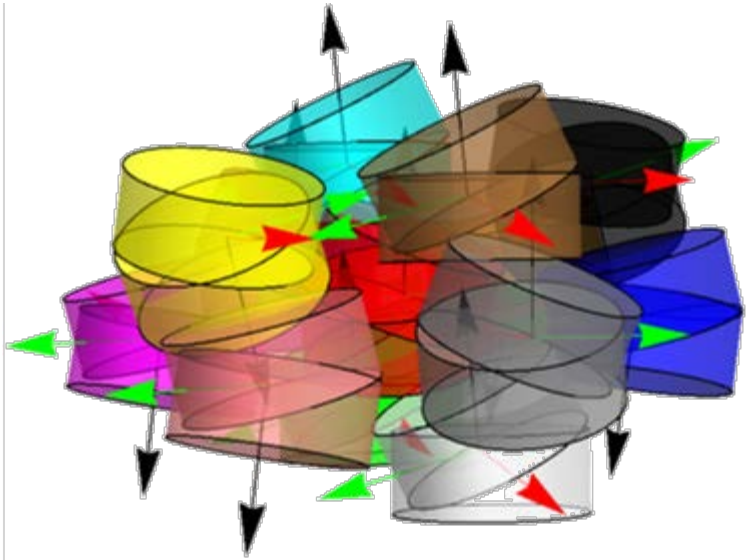
Analysis – Simulated Structures

- Easy to see why H3/H4 tails were important looking at contact matrix!
- H2A/H2B have no contacts near their tails

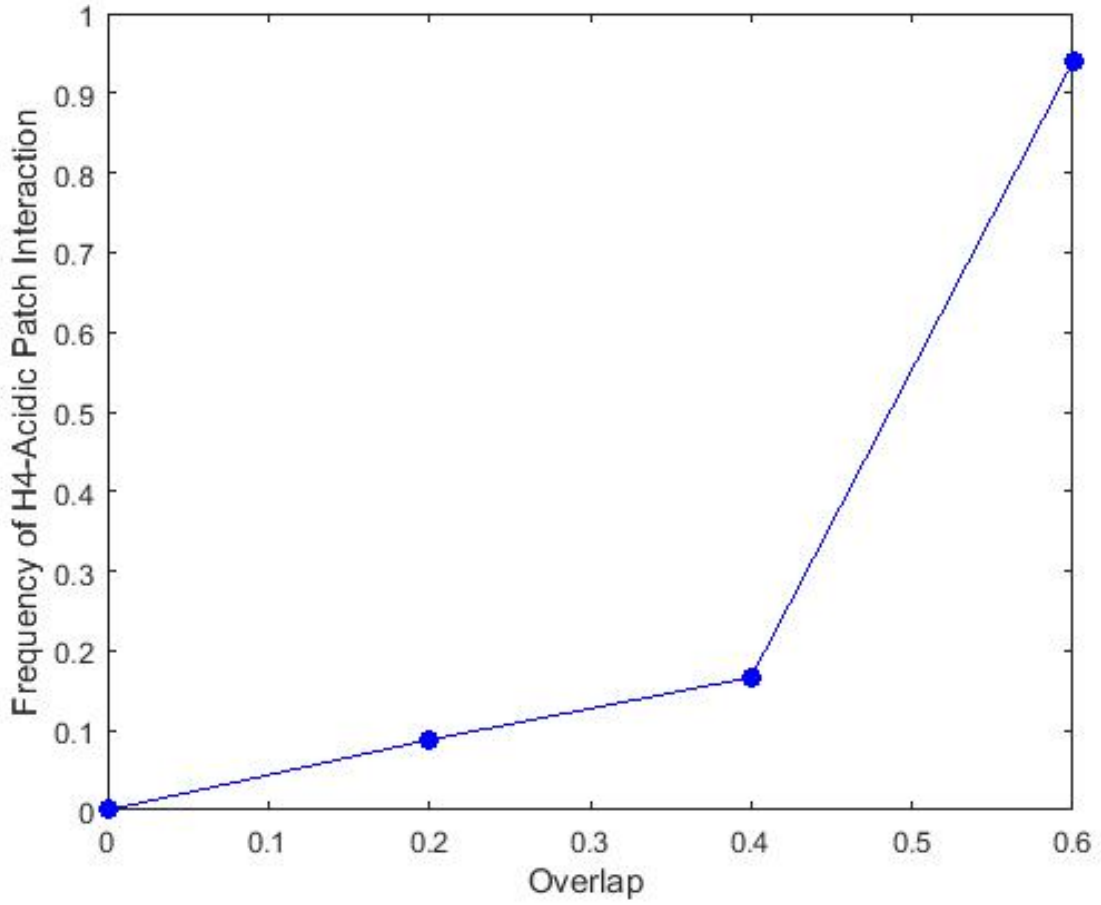


Analysis – Crystal Structures

- Nucleosomes packed together
- Experiments show acidic (- charge) patch on histone core H2A/H2B dimer
- Attracted to H4 (+ charge)?
- Confirm this by studying nucleosomes with large overlap



Larger Overlap = H4-AP interaction!



Conclusions

- Developed algorithms to analyze nucleosome overlap head-on and side to side
- Used to show effects of histone tails in simulations
- And confirm biological results in crystals

Thank you!

- Dr. Olson, Stefjord Todolli, other members of Olson group
- Dr. Gallos, Parker, DIMACS
- NSF grant CCF-1559855 (DIMACS)
- USPHS grant GM 34809 (Olson group)

(1) Computing the reference frame

- Given locations of P atoms, P_i
- Let r_i be vector difference of point i from mean

$$\boldsymbol{r_i} = \begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix} = \begin{pmatrix} x_i - \bar{x} \\ y_i - \bar{y} \\ z_i - \bar{z} \end{pmatrix}$$

- Consider the symmetric matrix

$$R = < \boldsymbol{r_i} \boldsymbol{r_i}^{\mathrm{T}} >$$

(1) Computing the reference frame

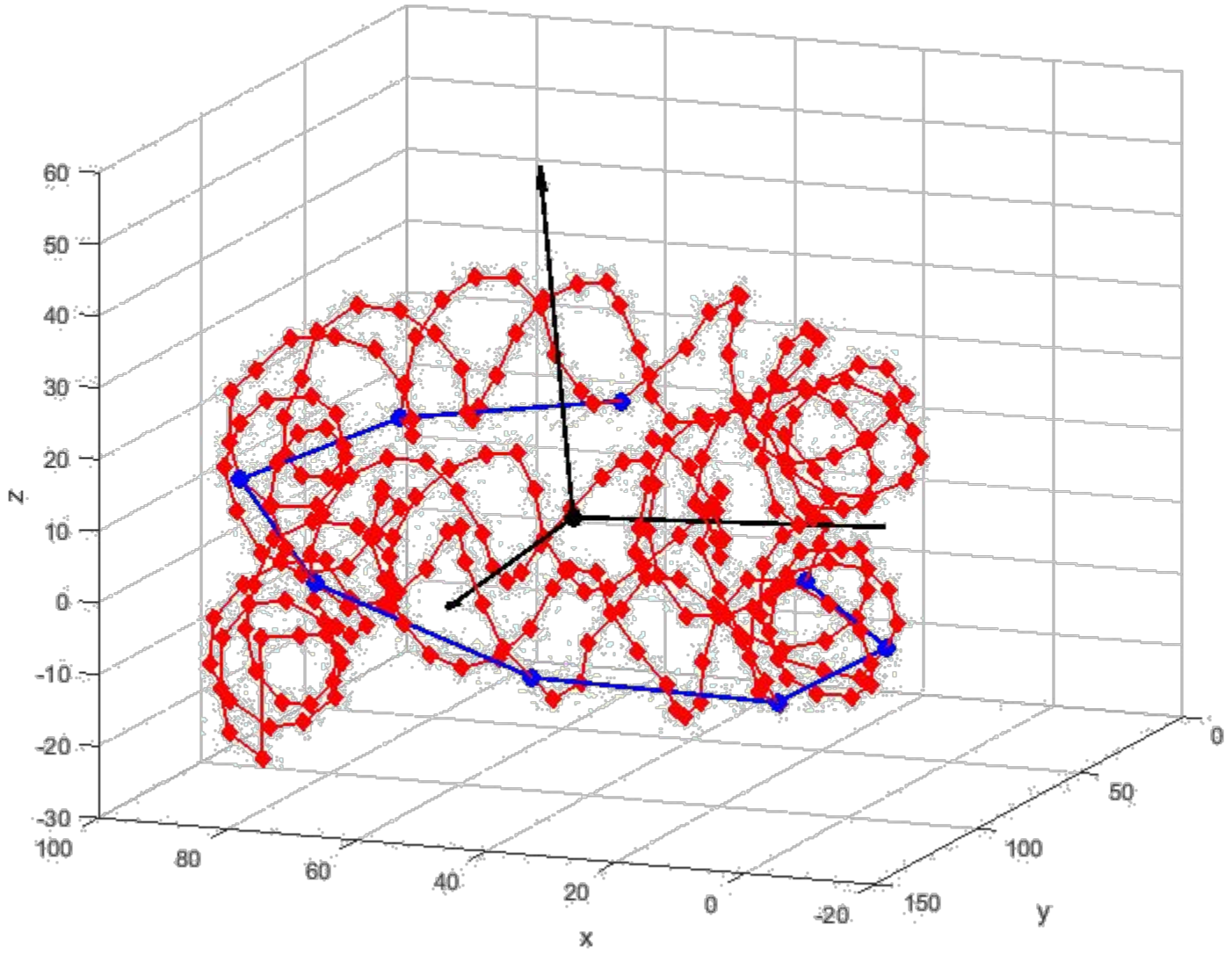
$$R = \begin{pmatrix} < u^2 > & < uv > & < uw > \\ < uv > & < v^2 > & < vw > \\ < uw > & < vw > & < w^2 > \end{pmatrix}$$

- Since *R* is symmetric, it has eigenvalue decomposition

$$R = TDT^*$$

- *D* is a diagonal matrix of eigenvalues, the columns of *V* are the eigenvectors (principal axes)
- Largest eigenvalue corresponds to largest deviation
  - ***n*** for tall cylinder

Also shown is polygon formed from every 10th phosphate on either side of dimer for 1 strand

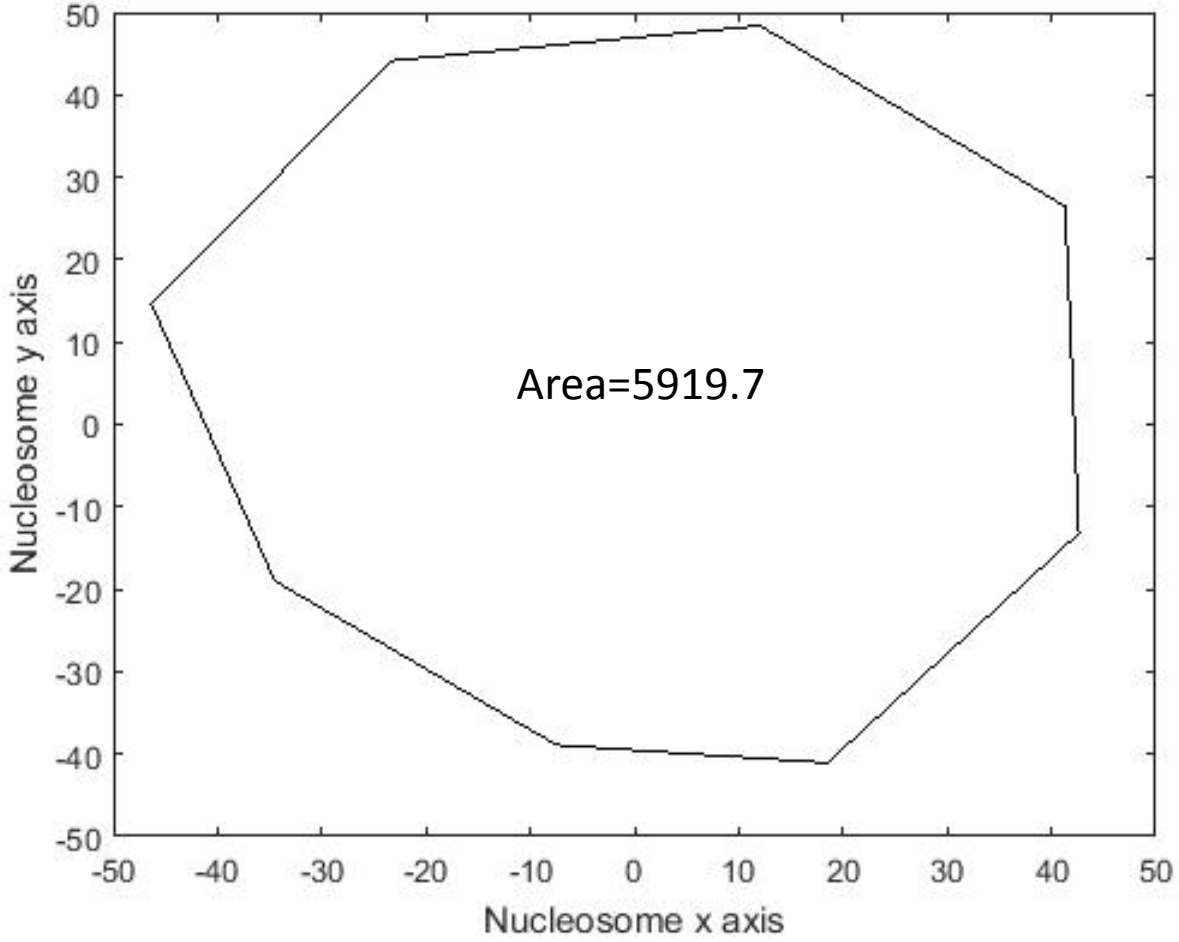


Important Observation

- The polygon coordinates wrt to the reference frame are constant

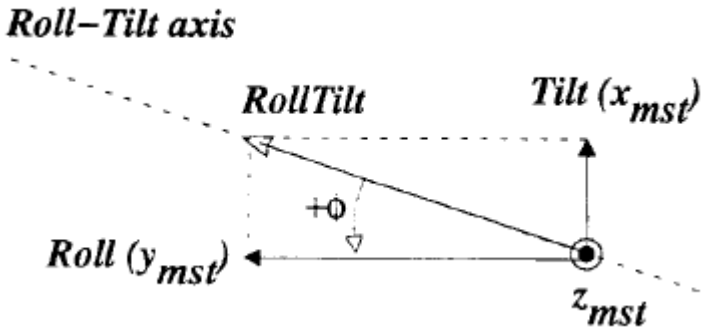
$$P_{nuc} = T^* P_{std}$$

- Remove normal coordinate (project points onto *n*=0 plane) to get polygon in basis


Area=5919.7
Nucleosome y axis
Nucleosome x axis
50
40
30
20
10
0
-10
-20
-30
-40
-50
-50
-40
-30
-20
-10
0
10
20
30
40
50

(2) Computing the midframe

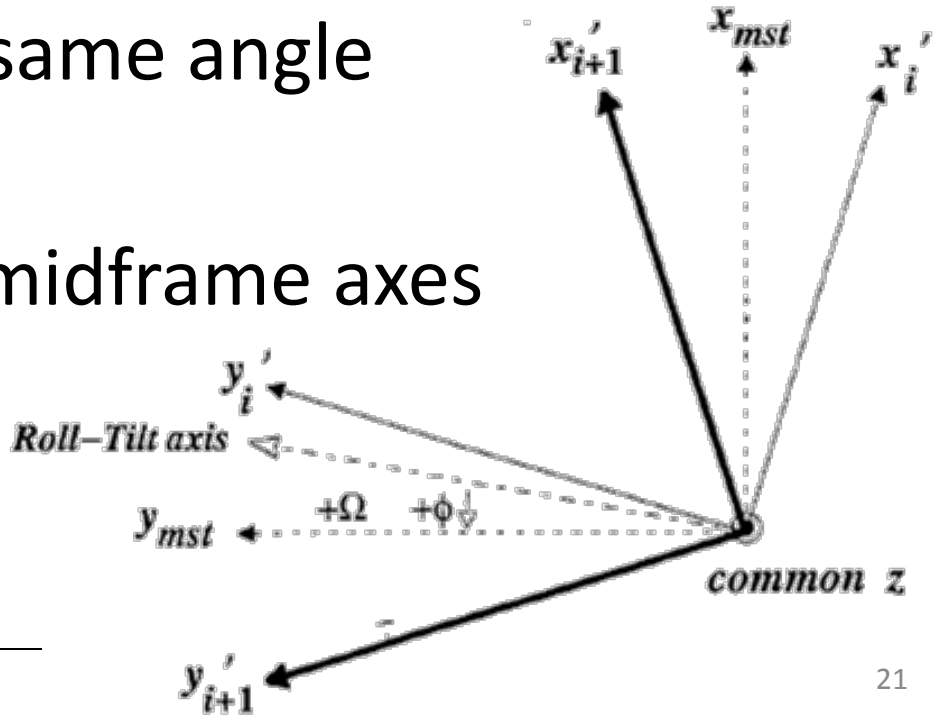
- Suppose 2 principal axes are given – how can we find the midframe?
  1. Define the “roll-tilt” axis



Lu et al. *J. Mol. Biol.* (1997) 273, 668-680

(2) Computing the midframe

1. Define RT axis
2. Rotate base pairs about RT axis so that they share z axis
3. Rotate x, y axes by same angle so they match
4. Matching axes are midframe axes



Lu et al. *J. Mol. Biol.* (1997) 273, 668-680

(2) Computing the midframe

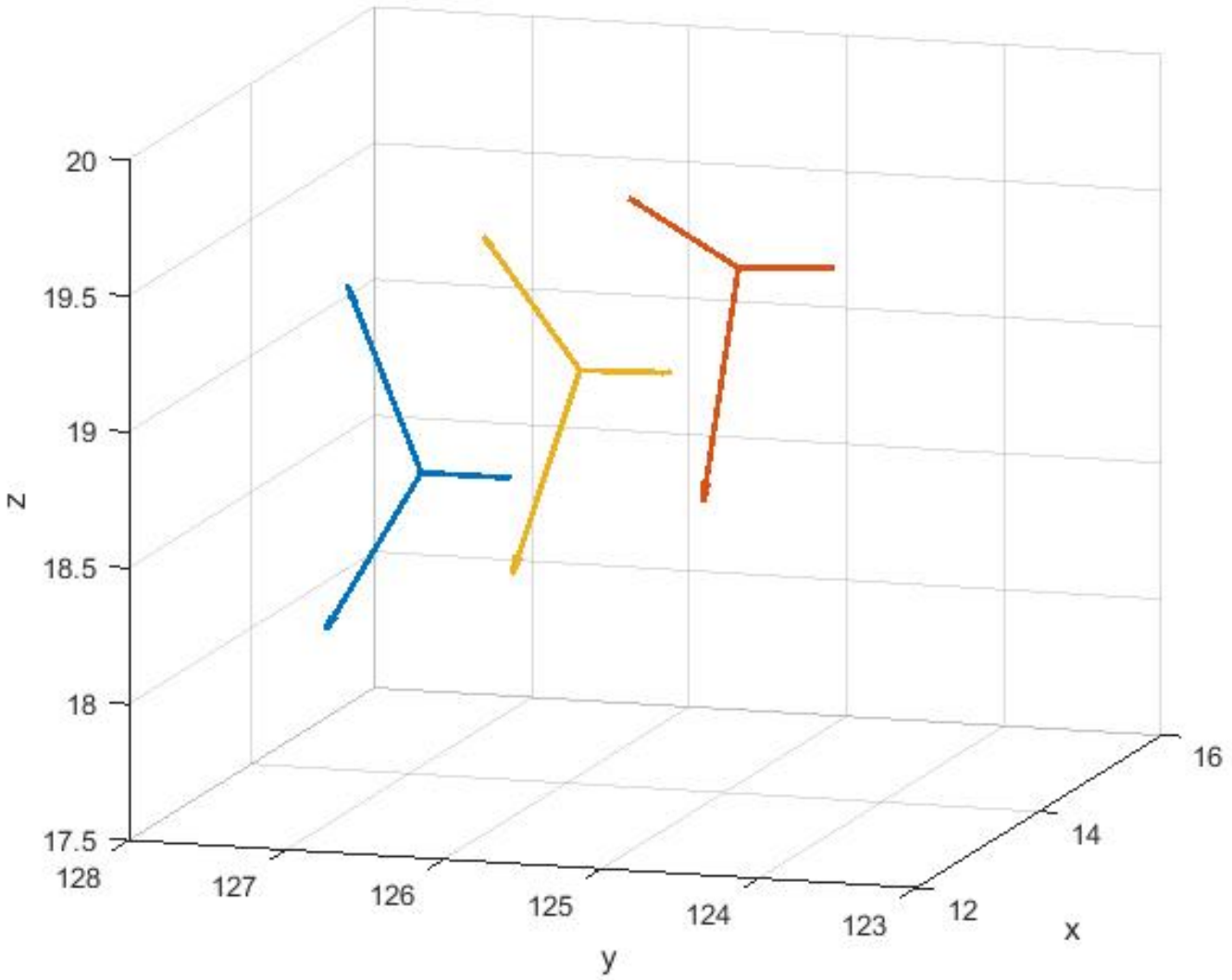
- This formulation recovers angular parameters tilt, roll, twist
- Define the midframe matrix

$$T_m = (\boldsymbol{x_{mst}} \quad \boldsymbol{y_{mst}} \quad \boldsymbol{z_{mst}})$$

- Translational parameters shift, slide, rise given by $T_m^* \boldsymbol{r}$, where $\boldsymbol{r} = o_2 - o_1$
- Computations match w3DNA for given coordinate frame

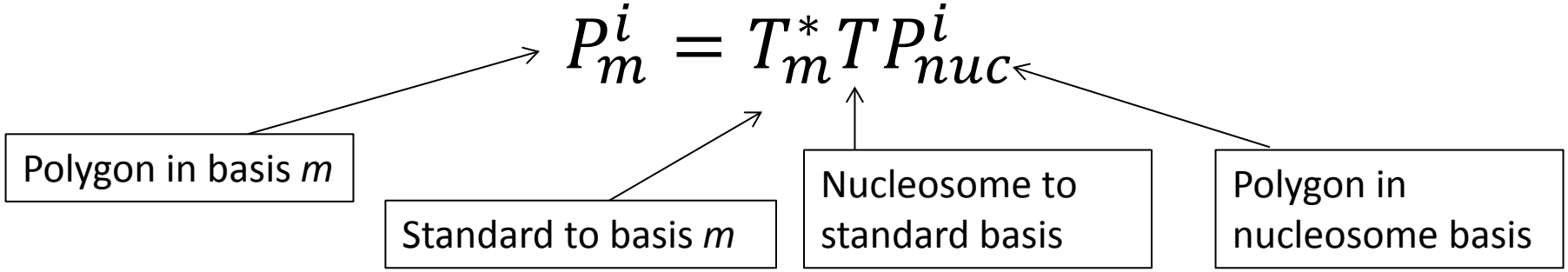
Lu et al. *J. Mol. Biol.* (1997) 273, 668-680

(2) Computing the midframe

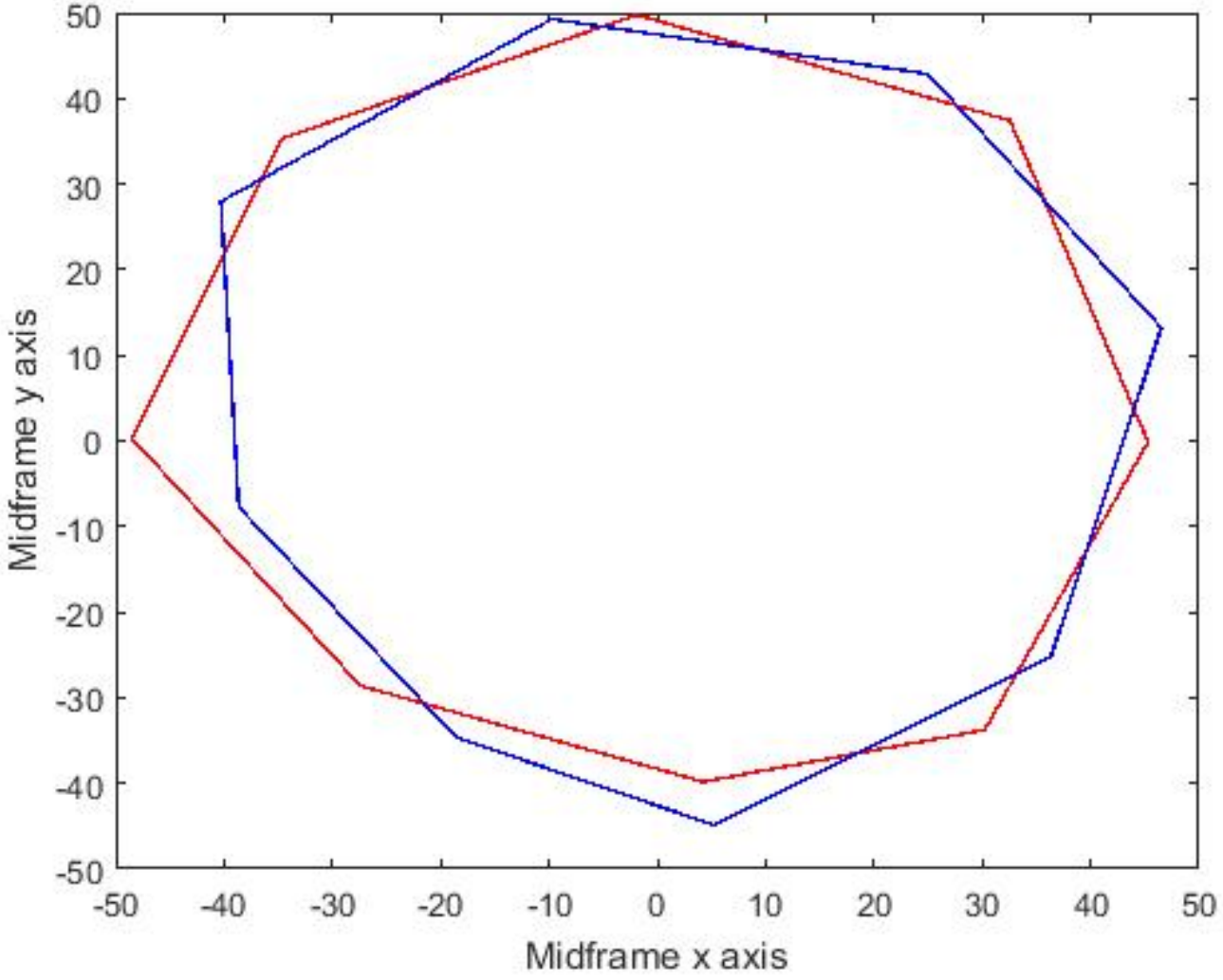


(3) Project polygons onto same plane

- The coordinates of polygon *i* wrt the midframe are essentially given by

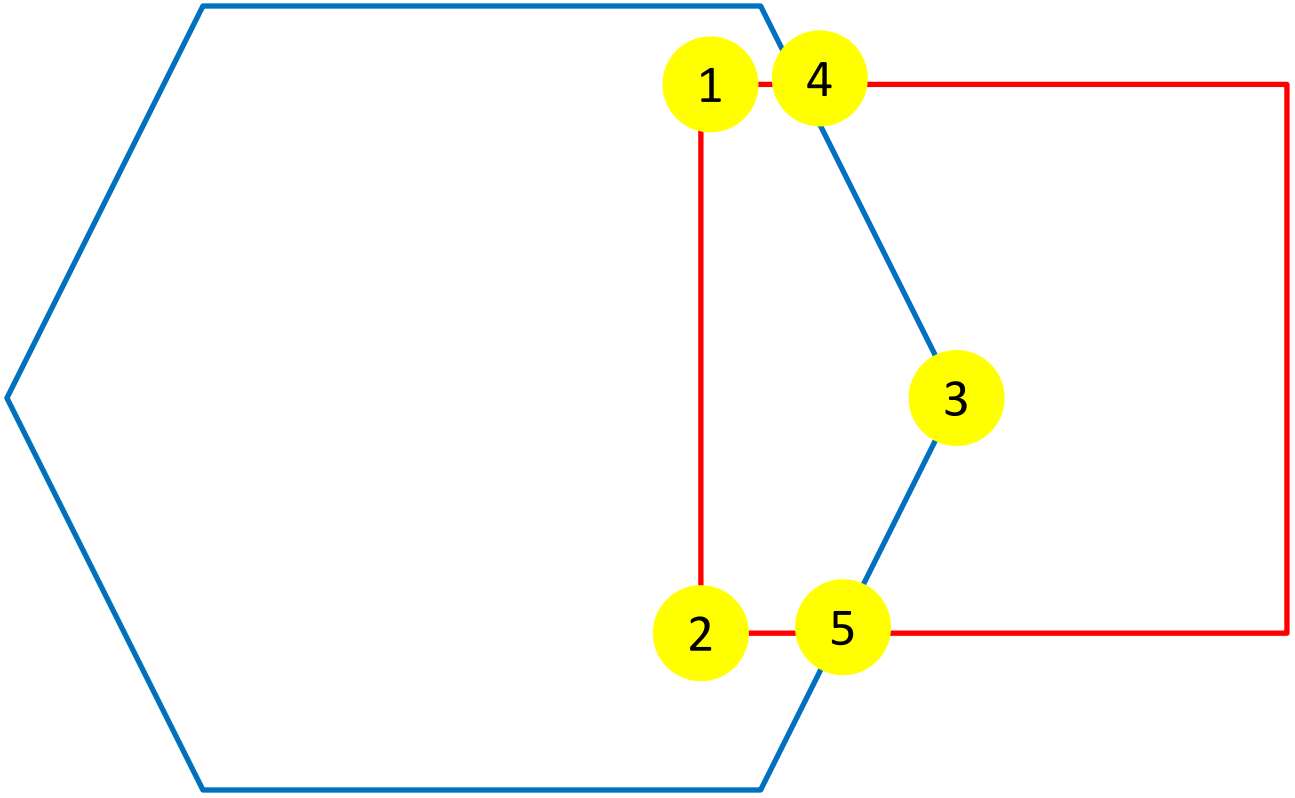


- (some differences for translation)
- Project the polygons onto xy plane by removing normal coordinate


Midframe y axis
Midframe x axis
50
40
30
20
10
0
-10
-20
-30
-40
-50
-50
-40
-30
-20
-10
0
10
20
30
40
50

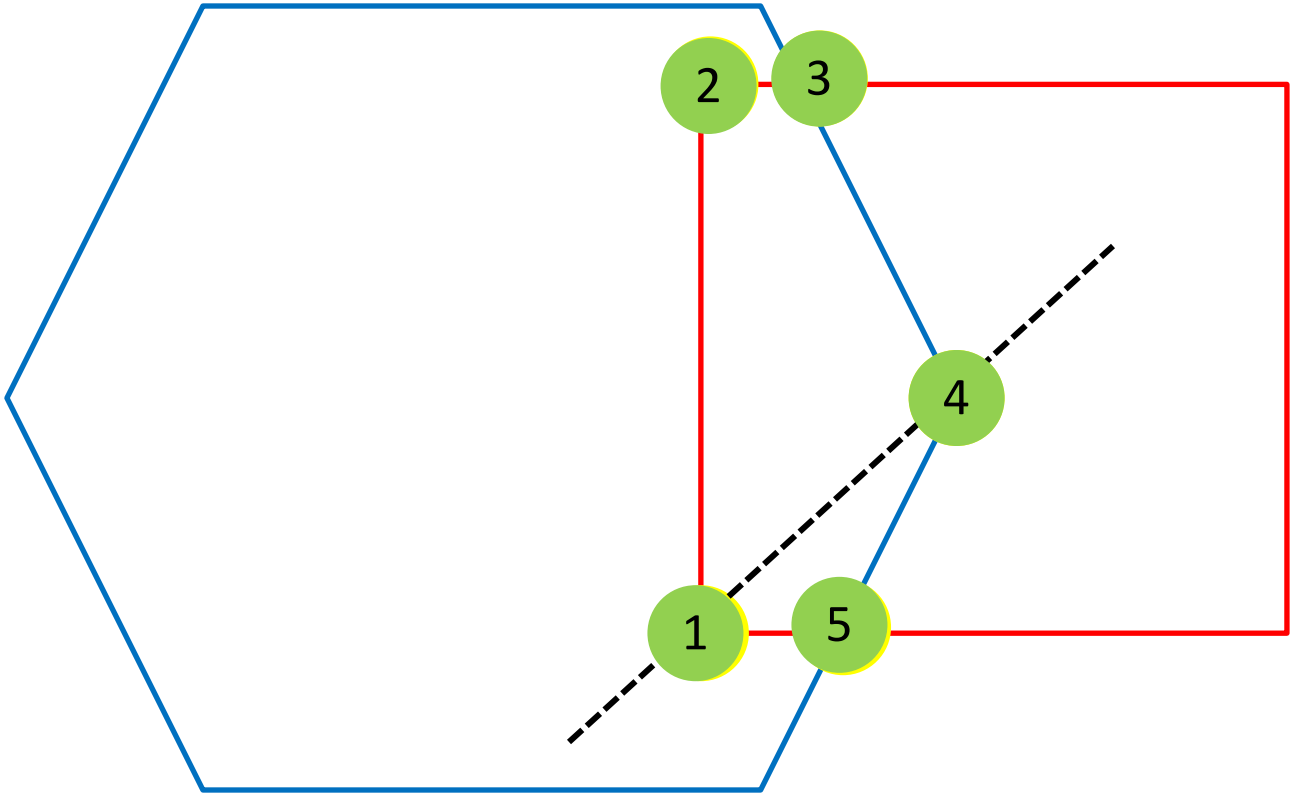
(4) Find overlapping area

1. Identify points from one inside the other
2. Identify intersecting points



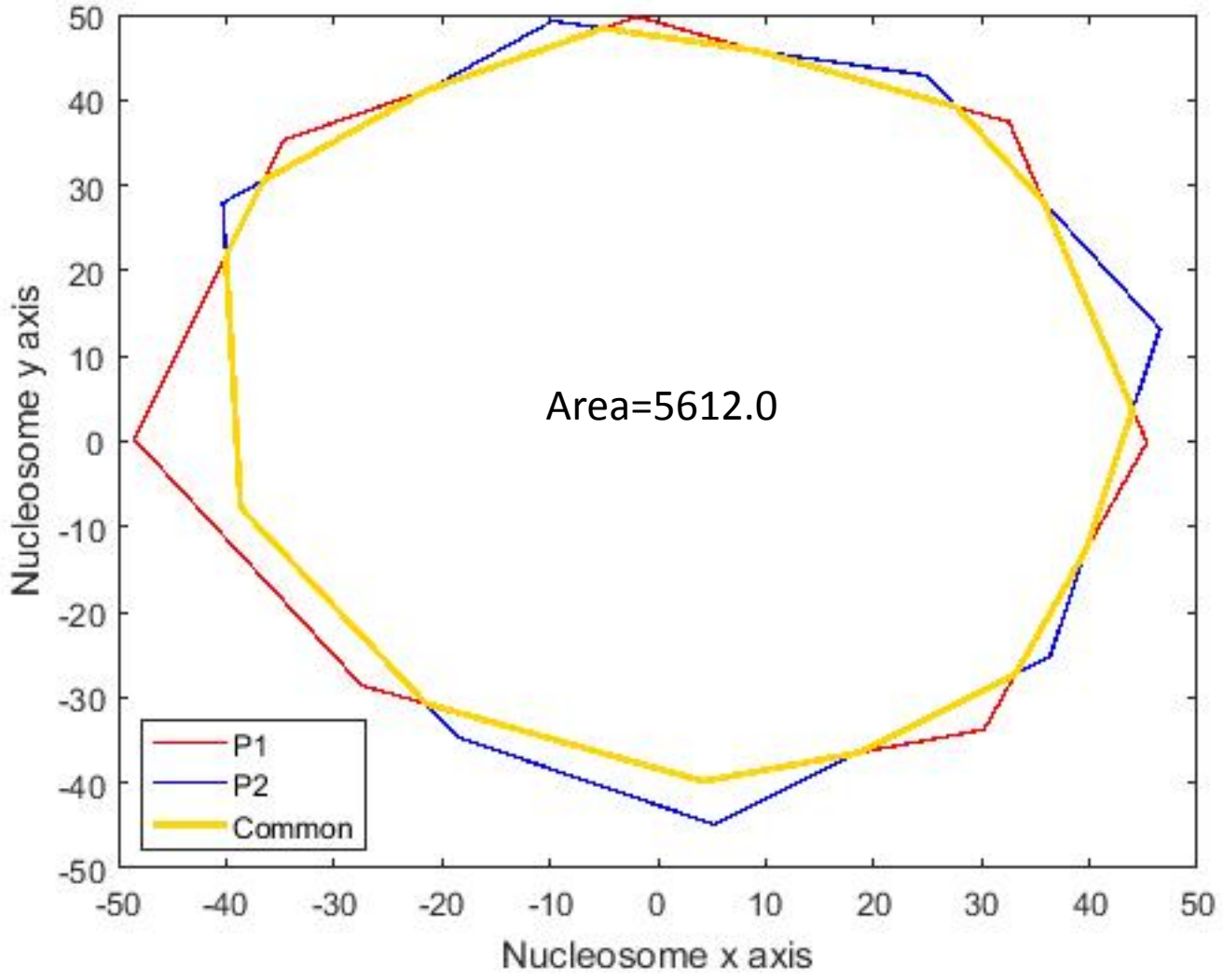
(4) Find overlapping area

3. Arrange points clockwise
4. Compute area



(4) Find overlapping area

- Method will always work
  - P_{nuc}^{i} is convex
    - Careful! Only in x,y space
    - Must 0 out normal *before* translating
  - Any linear map applied to it preserves convexity


Area=5612.0
P1
P2
Common
Nucleosome y axis
Nucleosome x axis
50
40
30
20
10
0
-10
-20
-30
-40
-50
-50
-40
-30
-20
-10
0
10
20
30
40
50

Overlap Coefficient for Simulations

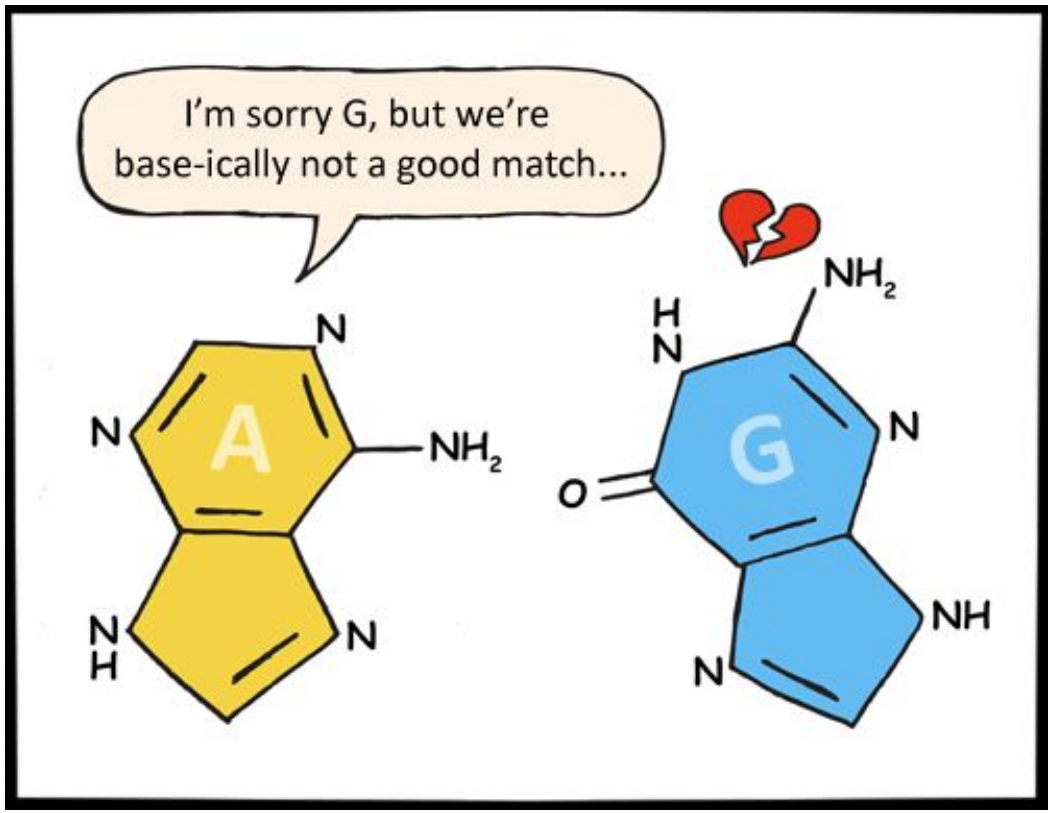
$$OC = \frac{A_{overlap}}{A_{orig}}?$$

In this case,

$$OC = \frac{5612.0}{5919.7} = 0.948$$

Thank you

- Questions?



http://www.clontech.com/US/Support/Clontech_Cartoons/February_2012