1. *Structure math:*

a. Absolute value $e^{ix}= \sqrt{e^{ix}e^{ix}*}$

Because : $e^{ix}=\cos x +i\sin x$; $e^{ix}*= \cos x -i\sin x$

So: $e^{ix}\, e^{ix*}= (\cos x +i\sin x)(\cos x -i\sin x)$

$=\cos^2 x - (i^2)\sin^2 x= \cos^2 + \sin^2 x=1$

Absolute value $e^{ix}= \sqrt{e^{ix}e^{ix}*}=1$

So the absolute value of e^{ix} is one

b. $e^{iz}= e^{i\,(a+ib)}= e^{ia+i*ib}= e^{-b+ia}=e^{-b}e^{ia}=e^{-b}(\cos a + i\sin a)$

$e^{i\,z*}= e^{-b}(\cos a - i\sin a)$

$|e^{iz}| = \sqrt{e^{iz}\, e^{i\,z*}}= \sqrt{e^{-b}(\cos a + i\sin a)\, e^{-b}(\cos a - i\sin a)}$

$= \sqrt{(e^{-b})^2[\cos^2 x - (i^2)\sin^2 x]}= \sqrt{(e^{-b})^2(\cos^2 + \sin^2 x)}$

$= \sqrt{(e^{-b})^2}= e^{-b}$

2. *Crystallography*

a. orthorhombic

b. It is twofold axis about **a**, **b**, **c**; the twofold screw axis along the **X** and **Y** direction.

c. The equivalent positions in space group are (x,y,z); (-x,-y,z); (-x+1/2, y+ 1/2, -z); (x+1/2, -y+1/2, -z)

There will therefore be a strong peak in the Patterson function at

1): (u,v,w) =(x,y,z) - (-x,-y,z) =(2x, 2y, 0)

2) (u,v,w) = (x,y,z) - (-x+1/2, y+ 1/2, -z) =(2x-1/2, -1/2, 2z)

3) (u,v,w) = (x,y,z) - (x+1/2, -y+1/2, -z) =(-1/2, 2y-1/2, 2z)

4) (u,v,w) =(-x,-y,z) - (-x+1/2, y+ 1/2, -z) =(-1/2, -2y-1/2, 2z)

5) (u,v,w) =(-x,-y,z) - (x+1/2, -y+1/2, -z) = (-2x-1/2, -1/2, 2z)

6) (u,v,w) = (-x+1/2, y+ 1/2, -z) - (x+1/2, -y+1/2, -z)= (-2x, 2y, 0)

Since -1/2=-1/2+1=1/2 ; 0=0+1=1

So u=1/2; v=1/2; w=1 are all the Harker section

3. *Multiple methods*

1): Solution structure of Human normal adult hemoblobin

PDB Accession Code:2H35

2): 1.25A resolution crystal structure of human hemoglobin in the deoxy form.

PDB Accession Code: 2DFQ

Comparison

Hemoglobin has a very physiologically significance to human healthy, and has been very well studied by multiple methods such as X-ray crystallography and NMR. The paper published by Yingqi Xu. etc, demonstrated the solution NMR structure of human adult hemoglobin. And the other paper stated the X-ray structure of human hemoglobin. Both of the methods can provide large structural information about the protein. They all determined the protein structure at atomic level and determine the distance of chemical bonds. Almost all the atoms have a 1.0 occupancy in its PDB file, which means their structure are been well identified.

Beyond the similarities, significant differences between the two methods in determining protein structures are exist. X-ray provides static structural information, but NMR offers a dynamic structure. The motions of atoms are incorporated into the atomic model by a B-value. There is no B value demonstrated on PDB file of NMR structure of human hemoglobin, because NMR structures are dynamic. This is not the special case, and at least certain percent of NMR structure lacking of B-value in its PDB file. The B value of X-ray structure are collected in PDB file, which is about 40 on average, it is not a very static structure although.

NMR tends to be effective for smaller molecule, and X-ray works well for bigger molecules. Now determining the protein structure of a protein as large as hemogblbin is on the trial stage, and it cannot provide as many information as X-ray can. The PDB file of the X-ray structure contained more information, because hemoglobin is a big

molecule. The NMR PDB file lacks many of the information such as the R_{free} value and the completeness. Since the X-ray structures are determines at 1.5Å resolution. There is no resolution information provided by NMR structure.

Reference

Xu, Yingqi. Zheng, Yu. Fan, Jing-Song, Yang, Daiwen. "A new strategy for structure determination of large proteins in solution without deuteration." *Nature Methods.* (Nov, 2006) 3.11: 931-937. Print

Park SY, Yokoyama T, Shibayama N, Shiro Y, Tame JR . "1.25 A resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms." *J Mol Biol*. (Jul,2006)360.3:690-701. Print

4. *Structure refinement*

a.

This one amino acid's side chain error might affect the original structure amplitudes, but it will not affect the final structure amplitudes after structure refinement.

Usually, a small number of structure amplitudes errors will not affect the overall structure amplitudes if the rest structure amplitudes are correct. You can always refine small errors by refinement methods.

b.

There are several refinement methods. For example, usually, take 5-10% of your structure amplitudes out of the list that are being refined on. Use the other 90-95% of the structure amplitudes in modifying the model to fit the observed data. Everytime you make a change in your structure based on the 90-95% of you data. If there are only small percent of errors, it always can be excluded eventually. So a small number of structure amplitudes errors will not affect the overall structure amplitudes if the rest structure amplitudes are correct.

5. *Crystallographic structure determination:*

Molecular replacement has proven effective for solving macromolecular crystal structure based upon the knowledge of homologous structure. The overall structure of the search model must be very close to that of the unknown structure for the technique to work. The Molecular replacement is straight and reduces the time and effort require for structure determination because there is no need to prepare heavy atom derivatives and collect their data. Model building is also simplified, since little or no chain tracing is required. The risk of model bias .i.e. the inability to correct errors in one's model is greater with molecular replacement than with experimental phasing. For molecular replacement you need figure out the orientation of the model and the location of oriented model.

For MAD and MIR. Accurate measurement of Friedel pair difference can be used to extract starting phase if the AD effect is large enough. The lack of isomorphism problem is much milder for anomalous data than molecular replacement. MAD only need one sample, and molecular replacement need two samples. The unit cell is by definition identical, and the molecule is in the same place within that unit cell.

6. *Nucleic acid struture*

The 12 different parameters of two adjacent base pairs are:

(1): vertical rise: the distance between two base pairs

(2): base pairs slide: base pairs not always across right the middle of helix at the same way

(3): base pair shift

(4): base pair twist

(5): base pair roll

(6): base pair tilt

(7): base pair stagger

(8): base pair stretch

(9): base pair shear

(10): base pair buckle

(11): base pair propeller twist

(12): base pair opening

7. *Muscle diffraction*

1). The long-spacing repeat is 232 nm.

2). The statistics implying the layer lines parallel to the equator can give the information of long-spacing repeat. Assuming the long-spacing repeat is N. The long-spacing repeat must be the common multiple of that statistics. It means that the layer lines only appear at the position of N/n (where N is the long-spacing repeat and n is an integer). So the 38.7nm(n=6), 23.2nm(n=10), 19.3nm(n=12), 14.5nm(n=16), 12.9nm(n=18), 7.25nm(n=32) and 5.95nm(n=39).

The 232 nm repeat can be thought of as the fundamental repeat

distances of the thick filament with that of the thin filament. The observed diffraction pattern arises from the periodic array of the thick and the thin filaments within the sarcomeres. When the muscle contracting, the length of sarcomere became short, that means the long-spacing repeat decrease. So the distances among the diffraction layer lines become short.

8. *Computational structural*

The computational tools at various time scales:

- Quantum mechanics methods work at faster time scales that faster than picoseconds.
- All-atom molecular dynamics simulation operating at range from picoseconds to microsecond range.
- NMR works at microseconds to millisecond range.
- Electrophysiology experiment works from ten microseconds up to seconds time scale.
- FRET works at milliseconds to thousand seconds scale.
- AFM/optical tweezers operating range from milliseconds to thousand seconds.
- EM/X-ray are telling us about static structure. They work at infinite time scale.

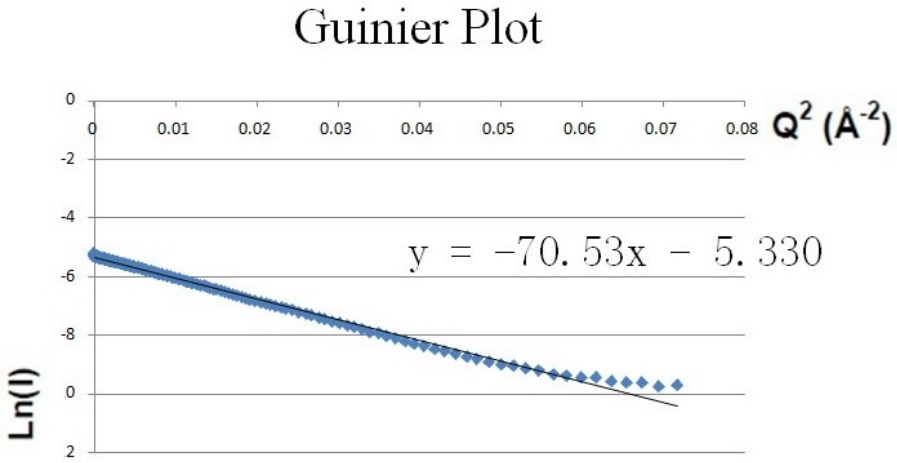
2). It is better to pick up pretty short time interval between calculations, and the appropriate time scales for reasonable calculations at the quantum level is 10^{-15} seconds. We can approach to a problem at several different

time scales. We get experimental data at quantum level. Then we can calibrate the results based on what's happen at the higher length scale and time scales such as at atomic level. Then determine whether they are consistent with each other. Then we treat atoms at particular position and have particular velocity. We can made atomistic model, which works at microsecond level. It may offer some statistic data which is useful about the mechanical models which reflects the motions about atoms at seconds time scale.

Or we can figure out whether those results of atomistic models make sense through calibration processes that happened at even high time scales (milliseconds) and length scale. And we also we choose bigger time interval which is milliseconds scale. Then we want to identify whether these are consistent with kinetic model we derived from other spectroscopic techniques which works at time scales of hours.

9. *X-ray Scattering*

1). According to the data I got. I made the Guinier plot as below:



According to the equation:

$$Ln[I(Q)] = Ln[I_0] - \frac{Q^2 R_g^2}{3},$$

$-R_g^2/3 = -70.53$; So $R_g = 14.55$

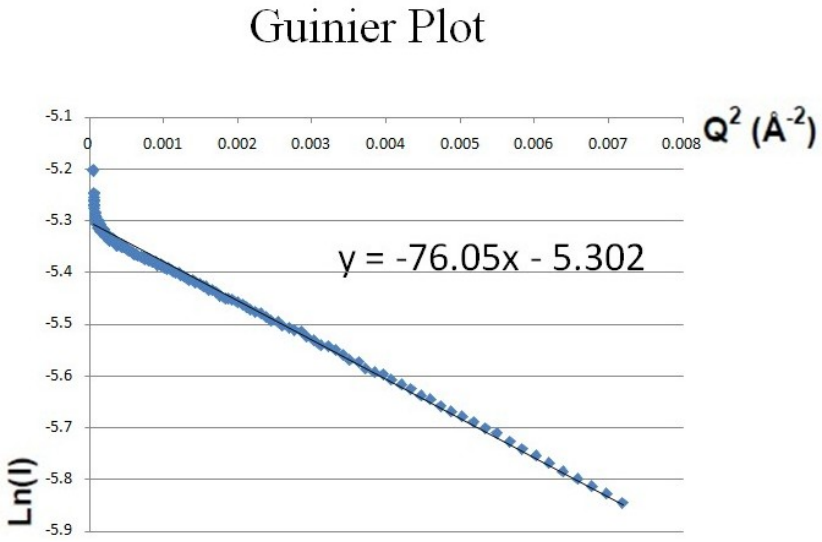
At that circumstance， $R_g \times Q_{max} = 3.89 > 1.3$

Assume when $R_g = 14.55$， and $R_g \times Q = 1.3$； So $Q = 0.89$

So we use the first 171 data to recalculate the R_g again. I get $R_g \times Q_{max} = 1.33 > 1.3$

I tried a couple more times. When I used first 169 data I got $R_g \times Q_{max} = 1.28 < 1.3$, which is good.

The new plot is below:



At that time $R_g = 15.1$

2). If it is a globular particle, $R_g = R \sqrt{3/5}$, where Rg is radius of gyration, and R is the radius of particle.

$R = R_g/0.775 = 19.48$

If it is a ellipsoid, $R_g^2 = (a^2 + b^2 + c^2)/5$, where Rg is radius of gyration, a,b,c are the three semi-axis lengths of the particle.

Since it is a prolate ellipsoid, so a=b<c.

So $R_g^2 = (2a^2 + c^2)/5$