**Colin Sloniker**
**Bio 555**
**7-31-12**

# Take-home Final

## 1) Structure Math

1. Show : $|e^{ix}| = 1$

   $|e^{ix}| = \sqrt{(e^{ix} \times e^{ix^*})}$
   $|e^{ix}| = \sqrt{\{[\cos(x) + i\sin(x)] * [\cos(x) - i\sin(x)]\}}$
   $|e^{ix}| = \sqrt{[\cos^2(x) + \sin^2(x)]}$
   $|e^{ix}| = \sqrt{(1)}$
   $|e^{ix}| = 1$

2. Show : $|e^{iz}| = e^{-b}$

   When: $z = a+ib$
   $|e^{i(a+ib)}| = e^{-b}$
   $|e^{ia+i(ib)}| = e^{-b}$
   $|e^{ia-b}| = e^{-b}$
   $|e^{ia} \times e^{-b}| = e^{-b}$
   $|e^{ia}| = 1$
   therefore $|1 \times e^{-b}| = e^{-b}$
   $e^{-b} = e^{-b}$

## 2) Crystallography

a) Orthorhombic
b) 2-fold axis around the z, 2-fold screw in the x & y direction
c) $(u,v,w) = (x,y,z) - (-x+½, y+½, -z)$
   $(u,v,w) = (x-(-x+½), (y-(y+½), z+z))$
   $(u,v,w) = (2x-½, -½, 2z)$

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

   $(u,v,w) = (x,y,z) - (x+½, -y+½, -z)$
   $(u,v,w) = (x-(x+½), (y-(-y+½), z+z))$
   $(u,v,w) = (-½, 2y-½, 2z)$

## 3) Multiple Methods (PDB ascension codes : 1OA5 & 1K6U)

Bovine pancreatic trypsin inhibitor (BPTI) as the name suggests inhibits the activity of trypsin. This protein is also one of the most extensively studied proteins in terms of structure. As a result, a broad range of genetic modifications and environmental variations have been exhibited on this protein with the resulting structural changes being recorded. In this essay we compare two such experiments, a high pressure NMR structure determination, and an X-ray crystallographic determination of a cyclic form of BPTI.

The immediate differences observed between the two determined structures are the result of the environmental conditions, and form of BPTI being used. In the paper in which

BPTi has been cyclized into cBPTI the peptide bond has stabilizing effects on residues Gly56 and Gly57 which normally have duel conformations (1). The major structure changed in the BPTI structure under high pressure is noted to be slight expansion of the protein. It was also seen that there was a change in the dihedral angles of the proteins back bone. (2) In comparing the two structures it seems increasing pressure causes more perturbations from the molecules normal structure than that of cyclization.

Despite the differing conditions and one being a cyclized version of BPTI, the two proteins fold remarkably similar to one and other, and natural BPTI. The high pressure model sees most of the effects of the higher pressure in the length of hydrogen bonds, reduction in packing defects and deformation of the active site. Despite this, the protein maintains a form similar, yet markedly different to that of BPTI under normal condition (2). In a similar manner cBPTI, despite hindrances due to the bond between the two termini, is able to fold in the expected manner and maintain its enzymatic properties. (1)


1. Botos, I., Z. Wu, W. Lu, and A. Wlodawer. "Crystal Structure of a Cyclic Form of Bovine Pancreatic Trypsin Inhibitor." *FEBS Letters* 509.1 (2001): 90-94. *Science Direct*. Web. 31 July 2012.

2. Williamson, M., K. Akasaka, and M. Refaee. "The Solution Structure of Bovine Pancreatic Trypsin Inhibitor at High Pressure." *Protein Science*12.9 (2003): 1971-979. *PMC*. Web. 31 July 2012.


**4) Structure Refinement**

a. The misplacing of an amino acid side chain would result in a large number of structure amplitudes being slightly wrong. This is due to the structure factor, and therefore the structure amplitude being a summation of all phases produced by atoms in a unit cell.
b. During refinement, discrepancies could be found when comparing the mathematically derived expected diffraction pattern to the experimentally obtained pattern using the least squares refinement process ($\sum w(Y_o-Y_c)^2$).

**5) Crystallographic Structure Determination.**

For structure determination of a protein, multiple methods are available. In the provided case having a 58% identical protein would allow for the use of molecular replacement. Molecular replacement is a practical method and requires no classical laboratory techniques and is driven solely on computational power. Thus it requires no special equipment. This allows for a researcher to perform the identification in their home office, rather than having to travel or collaborate with researchers who have access to equipment needed for other methods of structure determination. This lack of needing specialized equipment also reduces the cost of the method. Flaws with molecular replacement are that there may be false positives and negatives based on the search model used. Also, molecular replacement requires a high resolution and generally works best with smaller molecules. Another option is structure determination is the use of experimental phasing. The plus side of using experimental phasing is that it can result in good quality phases in a potentially short amount of time. Experimental phasing also

works with larger molecules. The drawbacks of this method of structure determination are that it can require the use of potentially toxic heavy metals, and need multiple crystals. Time and cost are another factor. Physical laboratory work must be performed, and expensive equipment is required specifically in the case of MAD, a synchrotron is needed to tune X-rays. Both methods of determination present benefits and drawbacks, it is up to the researcher to know these traits and determine which method will function best for the protein they are attempting to identify.

**6) Nucleic Acid Structure**

1. Vertical Rise
2. Slide
3. Shift
4. Twist
5. Roll
6. Tilt
7. Stagger
8. Stretch
9. Buckle
10. Propeller Twist
11. Shear
12. Opening

**7) Muscle Diffraction**

The long spacing repeat for the insect muscle is 232nm. This is due to actin being a double helix and having a full repeat every 77.4nm, and myosin being a four strain helix which fully repeats every 116nm. The two have their full repeats coincide every 232nm. The diffraction image for a contracting muscle would have a change in the intensities of the dots at both the 14.5nm and 7.25nm layer lines would be seen because the intensities are representative of the activity levels of the myosin filaments.

**8) Computational Structural Biology**

Computational structural determination tools used across the various timescales provide a level of information directly corresponding to the timescale needed to perform the calculations for the model. At the lowest time scale the computational tools mostly are concerned with the sub-atomic level. As the time scale progresses in length the level grows beyond the sub atomic level to a more atomic level, and looks at how atoms interact. Further along the time scale the resolution decreases to the point where the computational techniques evaluate the interactions between molecules. The information collected on the shorter time scales allows for researches to evaluate their experimental data and determine if they are proceeding in the correct direction, i.e. does the experimental data collected so far match, or agree with the computed values.

**9) X-ray Scattering**

Using a graph plotting ln(I) verses $Q^2$, the slope for a range of values was obtained. This slope corresponds to $R_g$. The range of values used to obtain $R_g$ was determined using the standard $R_g \times Q_{max} < 1.3$. Using this as the search criteria a range of values were found resulting in an $R_g$ equal to 46.11Å, with a $Q_{max}$ of .0221 $Å^{-1}$. The range was determined by trying to include the largest group of data, yet maintaining the search criteria of $R_g \times Q_{max} < 1.3$. Using this value and assuming a globular protein, the radius was calculated:

$46.11 = \sqrt{(3/5)}\ R$
$46.11 = 3/5\ R^2$
$2126.1321 = 3/5\ R^2$
$3543.5535 = R^2$
$59.53\ Å = R$

Assuming a prolated ellipsoid:

$R_g = \sqrt{(2+\gamma^2/5)}\ a$
$46.11Å = \sqrt{(2+\gamma^2/5)}a$
$2126.1321 = (2+\gamma^2/5)a^2$
$2126.1321/a^2 = 2+\gamma^2/5$
$2124.1321/a^2 = \gamma^2/5$
$2124.1321 = (\gamma^2/5)a^2$
$10620.6605 = (\gamma^{2\,*}\ a^2)$
$\sqrt{10620.6605} = \sqrt{(\gamma^{2\,*}\ a^2)}$
$103.05 = \gamma * a$
If γ is the ratio of the axes, then $\gamma = b/a$ therefore
$103.05 = (b/a) * a$
$103.05 = b$