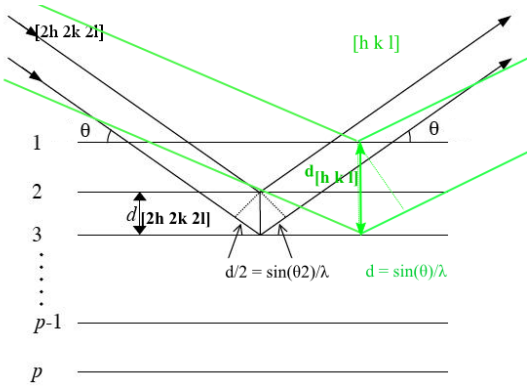
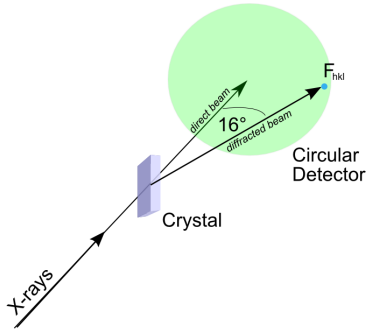
BC530 2014 X-ray Crystallography Unit Homework

Bragg's Law and measured Bragg reflections

The Bragg planes for a reflection [h k l] and another reflection [2h 2k 2l] are parallel. Therefore the same diagram showing the conditions for satisfying the Bragg diffraction condition can be used to show the values of λ, d , and θ required for both reflections.





1. Sketch on the diagram above what is different about diffraction of the [h k l] and [2h 2k 2l] elections.
   The distance d2 between [2h 2k 2l] planes is half the distance d between [h k l] planes. That means the angle $\theta 2$ for diffraction of the [2h 2k 2l] reflection is larger than the angle θ for the [h k l] reflection. From the diagram above
   $$\theta 2 = asin(\frac{1}{2}sin(\theta))$$

2. Does this mean that the amplitudes ($|F| = amplitude = \sqrt{intensity}$) measured experimentally for the two reflections are the same? I.e. is it true that $|F_{hkl}| = |F_{2h2k2l}|$? Why or why not?
   They are not the same. Consider one atom lying on plane *p1* in the figure, and a second atom below it lying on plane *p2*. They are separated by distance d/2. Scattering contributions from this pair of atoms will be in phase for reflection [2h 2k 2l] but will be exactly out of phase for reflection [h k l]. In general the vector sum of contributions from atoms everywhere in the crystal will be different for the two reflections, and hence the measured amplitudes will be different.

3. Where do the [h k l] and [2h 2k 2l] "spots" appear on the recorded diffraction image (right figure)?
   The higher resolution reflection [2h 2k 2l] will be further from the center. A line drawn from the center will pass through both spots.

Evaluate Structures from a Recent Report

Here is the Supplementary Table 1 (to which I have added a few lines) from a recent structure determination of the c-di-AMP riboswitch, an RNA mediator of a bacterial signaling pathway related to DNA damage [*Gao & Serganov, Nature Chemical Biology 10, 787–792 (2014) doi:10.1038/nchembio.1607*].

Supplementary Table 1 Data collection, phasing and refinement statistics

	T. pseudethanolicus $[Ir(NH_3)_6]^{3+}$-soaked	*T. pseudethanolicus* Native	*T. lienii* Native
PDB Entry Code	**4QKA**	**4QK8**	**4QK9**
X-ray source			
synchrotron	NSLS X25	APS 24-ID	NSLS X25
wavelength	1.105Å	0.9792Å	1.105Å
Data collection			
Space group	$P3_121$	$P3_121$	$P2_13$
Cell dimensions			
a, *b*, *c* (Å)	116.0, 116.0, 114.1	114.9, 114.9, 114.7	110.3, 110.3, 110.3
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 90.0
Resolution (Å)	30.00-3.20 (3.31-3.20) *	30.00-3.05 (3.16-3.05)	30.00-3.00 (3.11-3.00)
R_{sym} or R_{merge}	0.12 (0.71)	0.07 (0.52)	0.11 (0.55)
$I/\sigma I$	37.3 (3.0)	34.7 (3.5)	40.4 (6.0)
Completeness (%)	100.0 (99.7)	99.6 (99.9)	99.8 (100.0)
Redundancy	17.7 (12.1)	6.0 (6.1)	11.5 (12.2)
Refinement			
Resolution (Å)	30.00-3.20	30.00-3.05	30.00-3.00
No. reflections	28,200	16,997	9,173
R_{work}/ R_{free}	17.1 / 19.7	18.1 / 19.5	18.2 / 22.4
No. atoms			
RNA	2,578	2,626	2,516
Ligand/ion	130	98	99
Water	2	4	0
B-factors			
RNA	99.4	90.9	57.2
Ligand/ion	84.9	63.0	43.3
Water	56.7	56.1	----
R.m.s deviations			
Bond lengths (Å)	1.049	1.051	0.987
Bond angles (°)	0.005	0.005	0.005

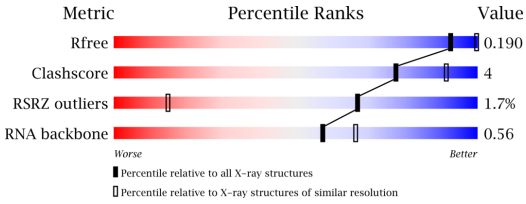
*Highest resolution shell is shown in parenthesis.

1. What is the resolution of the structure? Why are there two lines in the table marked "Resolution"?
   The resolution of both native structures is 3.05Å. The resolution of the Irridium-hexamine derivative is 3.2Å. One line shows the resolution limit of the data measurements (maximum 2θ).The other line shows the resolution of the data used to refine the models. In this case they are the same, but often not all the data measured are used in refinement. For example the reflections measured in the highest resolution part of the data may be so weak that they are swamped by the background noise, and hence are not used in refinement. That clearly is not the case here, because the signal-to-noise ratio I/σ_I is still gratifyingly large (> 3) in the highest resolution range of the data measured.

2. When you collect data at a synchrotron you can choose what wavelength X-rays are used. Why do you suppose they chose one wavelength for the data collected at NSLS and a different wavelength for the data collected at APS?
   Irridium is in the 6th row of the periodic table and hence has a very complicated set of anomalous scattering edges. You can see the full set by clicking on the Ir symbol in the periodic table here: `http://skuld.bmsc.washington.edu/scatter/AS_periodic.html`, or you can generate a plot by typing Ir in the web tool here and also selecting the "Elemental absorption edges" tick box: `http://skuld.bmsc.washington.edu/scatter/AS_form.html`. They chose a wavelength of 1.105Å in order to maximize the anomalous scattering signal from the Irridium L_{III} edge.

3. What indications do you have that this structure determination and refinement was done well, or poorly?

From the table: 99% complete data (excellent), $R_{merge} <$ 0.10 for the native data sets (we didn't discuss this in class, but it is another example of using an R factor to measure quality of agreement: lower R is better), $R_{free}$is only slightly higher than R_{work} (very good). It all looks very good. You can also easily look at the summary charts for these PDB entries on the PDB web site. I show one here for 4QK8.



4. I said in class that the Ramachandran chart is one of the few quality checks you can expect to see in a good crystal structure report, but there is no indication of Ramachandran chart quality in this table. Should we fault them for this lack? You can find equivalent information on the PDB web site for these three structures. Is the quality good or bad?
   http://www.pdb.org/pdb/explore/explore.do?structureId=4QK8
   Trick question. Ramachandran plots are for proteins, but this is an RNA structure. The "RNA backbone" slider in the validation chart above is roughly equivalent, but the energy landscape for RNA conformations is not so easily captured in a simple chart as it is for proteins.

5. Just considering the information in the Table 1, how do you think they solved these structure? That is, how might they have gotten initial phases to calculate electron density maps into which they could build and refine structural models?
   They could have used the crystals soaked in Irridium-hexamine as a derivative for SIR phasing. Irridium has lots of electrons, and the unit cell parameters for the native and derivative crystals are the same to roughly 1%, close enough to treat as being isomorphous. However in this case they clearly went to the trouble of collecting data at the Ir L_{III} edge, so it seems clear they wanted to use anomalous scattering to calculate SAD (Single wavelength anomalous diffraction) phases. Note that the crystals of the *T. lienii* homolog are in a different space group, so they could not have been used as an isomorphous derivative despite the similar values for cell edges a, b, c. After solving the first structure by SAD they could use it as a molecular replacement model to solve the homologous structure.