Chem 406 - Lecture 7

Structure Determination by X'Ray Crystallography

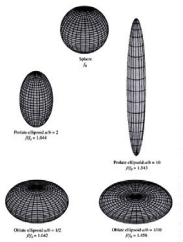
Introduction

- Earlier in the semester we developed a crude model for macromolecules using sedimentation analysis.
  - Prolate and oblate ellipsoids of revolution

TABLE 5.1 FRICTIONAL COEFFICIENT RATIOS

Shape	f/f_0	R_e
Prolate ellipsoid	$\frac{P^{-1/3}(P^2-1)^{1/2}}{\ln[P+(P^2-1)^{1/2}]}$	$(ab^2)^{1/3}$
Oblate ellipsoid	$\frac{(P^2-1)^{1/2}}{P^{1/3}\tan^{-1}(P^2-1)^{1/2}}$	$(a^2b)^{1/3}$
Long rod	$\frac{(2/3)^{1/3}P^{2/3}}{\ln 2P-0.30}$	$\left(\frac{3a^2a}{2}\right)^{1/3}$

In these equations, $P = a/b$, where a is the semimajor axis (or the half-length for a rod) and b is the minor axis (or radius of a rod). R_e is the radius of a sphere equal in volume to the ellipsoid or rod, so $f_0 = 6\pi\eta R_e$.



Introduction

- Since the late 1950's we have had models with atomic resolution.

Protein Data Bank

HTTP Status 500 -

type Exception report

message

description The server encountered an internal error () that prevented it from fulfilling this request.

exception

```
org.apache.tiles.DefinitionsFactoryException: XML error reading definitions.
        org.apache.tiles.digester.DigesterDefinitionsReader.read(DigesterDefinitionsReader.java:156)
        org.apache.tiles.definition.UrlDefinitionsFactory.readDefinitions(UrlDefinitionsFactory.java:220)
        com.opensymphony.webwork.views.tiles.TilesResult.getComponentDefinition(TilesResult.java:137)
        com.opensymphony.webwork.views.tiles.TilesResult.doExecute(TilesResult.java:91)
        com.opensymphony.webwork.dispatcher.WebWorkResultSupport.execute(WebWorkResultSupport.java:143)
        com.opensymphony.xwork.DefaultActionInvocation.executeResult(DefaultActionInvocation.java:311)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:206)
        org.rcsb.mypdb.AuthenticationInterceptor.intercept(AuthenticationInterceptor.java:28)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        com.opensymphony.xwork.interceptor.TimerInterceptor.invokeUnderTiming(TimerInterceptor.java:131)
        com.opensymphony.xwork.interceptor.TimerInterceptor.intercept(TimerInterceptor.java:118)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        org.rcsb.util.interceptor.StructureObjInterceptor.intercept(StructureObjInterceptor.java:61)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        org.rcsb.util.interceptor.StructureIdInterceptor.intercept(StructureIdInterceptor.java:62)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        org.rcsb.util.interceptor.NavigationInterceptor.intercept(NavigationInterceptor.java:53)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        com.opensymphony.xwork.interceptor.DefaultWorkFlowInterceptor.doIntercept(DefaultWorkFlowInterceptor.ja
        com.opensymphony.xwork.interceptor.MethodFilterInterceptor.intercept(MethodFilterInterceptor.java:86)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        com.opensymphony.xwork.validator.ValidationInterceptor.doIntercept(ValidationInterceptor.java:115)
        com.opensymphony.xwork.interceptor.MethodFilterInterceptor.intercept(MethodFilterInterceptor.java:86)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
```

Protein Data Bank

HTTP Status 500 -

type Exception report

message

description The server encountered an internal error () that prevented it from fulfilling this request.

exception

```
org.apache.tiles.NoSuchDefinitionException: Error while resolving definition inheritance: child 'pdb.statistics
        org.apache.tiles.ComponentDefinition.resolveInheritance(ComponentDefinition.java:598)
        org.apache.tiles.definition.ComponentDefinitionsImpl.resolveInheritances(ComponentDefinitionsImpl.java:
        org.apache.tiles.definition.ComponentDefinitionsImpl.addDefinitions(ComponentDefinitionsImpl.java:70)
        org.apache.tiles.definition.UrlDefinitionsFactory.readDefinitions(UrlDefinitionsFactory.java:221)
        com.opensymphony.webwork.views.tiles.TilesResult.getComponentDefinition(TilesResult.java:137)
        com.opensymphony.webwork.views.tiles.TilesResult.doExecute(TilesResult.java:91)
        com.opensymphony.webwork.dispatcher.WebWorkResultSupport.execute(WebWorkResultSupport.java:143)
        com.opensymphony.xwork.DefaultActionInvocation.executeResult(DefaultActionInvocation.java:311)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:206)
        org.rcsb.mypdb.AuthenticationInterceptor.intercept(AuthenticationInterceptor.java:28)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        com.opensymphony.xwork.interceptor.TimerInterceptor.invokeUnderTiming(TimerInterceptor.java:131)
        com.opensymphony.xwork.interceptor.TimerInterceptor.intercept(TimerInterceptor.java:118)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        org.rcsb.util.interceptor.StructureObjInterceptor.intercept(StructureObjInterceptor.java:61)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        org.rcsb.util.interceptor.StructureIdInterceptor.intercept(StructureIdInterceptor.java:62)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        org.rcsb.util.interceptor.NavigationInterceptor.intercept(NavigationInterceptor.java:53)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        com.opensymphony.xwork.interceptor.DefaultWorkflowInterceptor.doIntercept(DefaultWorkflowInterceptor.ja
        com.opensymphony.xwork.interceptor.MethodFilterInterceptor.intercept(MethodFilterInterceptor.java:86)
        com.opensymphony.xwork.DefaultActionInvocation.invoke(DefaultActionInvocation.java:188)
        com.opensymphony.xwork.validator.ValidationInterceptor.doIntercept(ValidationInterceptor.java:115)
```

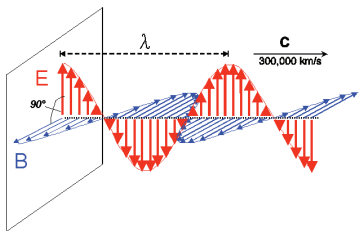
5

Protein Data Bank

```
HEADER    HYDROLASE                               26-MAR-04   1SU4
TITLE     CRYSTAL STRUCTURE OF CALCIUM ATPASE WITH TWO BOUND CALCIUM
TITLE    2 IONS
COMPND    MOL_ID: 1;
COMPND    2 MOLECULE: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM
COMPND    3 ATPASE 1;
COMPND    4 CHAIN: A;
COMPND    5 SYNONYM: CALCIUM-TRANSPROTING ATPASE SARCOPLASMIC
COMPND    6 RETICULUM TYPE, FAST TWITCH SKELETAL MUSCLE, SERCA1A;
COMPND    7 EC: 3.6.3.8
SOURCE    MOL_ID: 1;
SOURCE    2 ORGANISM_SCIENTIFIC: ORYCTOLAGUS CUNICULUS;
SOURCE    3 ORGANISM_COMMON: RABBIT;
SOURCE    4 TISSUE: MUSCLE
KEYWDS    ION PUMP, CALCIUM, MEMBRANE PROTEIN, P-TYPE ATPASE, ACTIVE
KEYWDS   2 TRANSPORT
EXPDTA    X-RAY DIFFRACTION
AUTHOR    C.TOYOSHIMA,M.NAKASAKO,H.NOMURA,H.OGAWA
REVDAT   1   04-MAY-04 1SU4    0
SPRSDE     04-MAY-04 1SU4      1EUL
JRNL        AUTH   C.TOYOSHIMA,M.NAKASAKO,H.NOMURA,H.OGAWA
JRNL        TITL   CRYSTAL STRUCTURE OF THE CALCIUM PUMP OF
JRNL        TITL 2 SARCOPLASMIC RETICULUM AT 2.6 A RESOLUTION.
JRNL        REF    NATURE                         V. 405    647 2000
JRNL        REFN   ASTM NATUAS  UK ISSN 0028-0836
REMARK   1
REMARK   1 REFERENCE 1
REMARK   1  AUTH   C.TOYOSHIMA,H.NOMURA
REMARK   1  TITL   STRUCTURAL CHANGES IN THE CALCIUM PUMP
REMARK   1  TITL 2 ACCOMPANYING THE DISSOCIATION OF CALCIUM.
REMARK   1  REF    NATURE                         V. 418    605 2002
REMARK   1  REFN   ASTM NATUAS  UK ISSN 0028-0836
REMARK   1 REFERENCE 2
REMARK   1  AUTH   C.TOYOSHIMA,H.NOMURA,Y.SUGITA
```

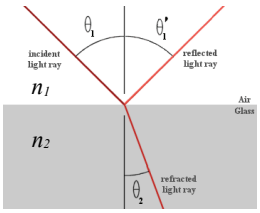
6

Electromagnetic Radiation



7

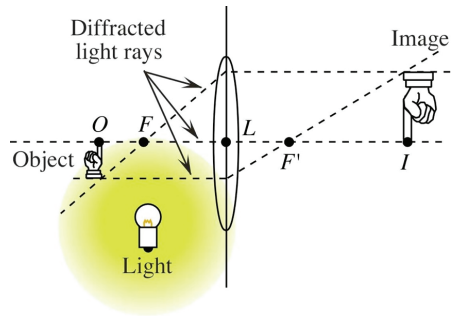
Snell's Law



$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$$

8

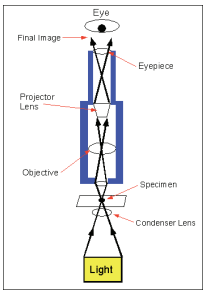
Lenses



9

Microscopy

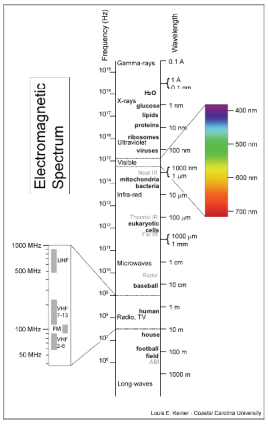
Light microscopy



10

Electromagnetic Spectrum

To use electromagnetic radiation to create images of objects requires electromagnetic radiation having wavelengths that are equal to or small than the objects being imaged.



11

Louis de Broglie

In 1924 proposed

- Just as light has a particle-like nature

$$E = h\nu \qquad \text{Planck}$$

- Matter has a wave-like nature

$$\lambda = \frac{h}{p} \qquad \text{de Broglie}$$

$$= \frac{h}{mv}$$

12

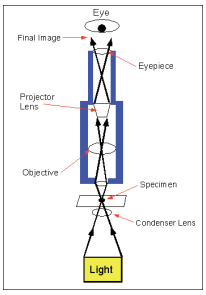
High Energy Electron Beam

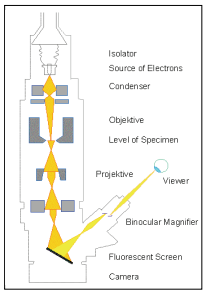
- For a 100 *kV* beam, in which an electron is accelerated through a 100,000 *V* potential field.

$$\lambda = \frac{h}{mv}$$
$$= 0.004 \ nm$$
$$= 0.04 \ \text{Å}$$

Microscopy

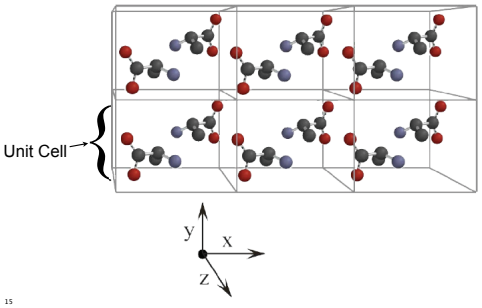
Comparing light microscopy to electron microscopy





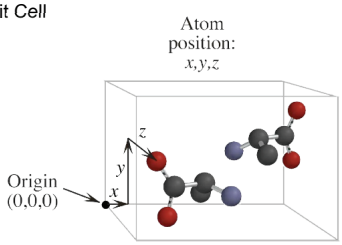
X-ray Crystallography

The Crystal Lattice



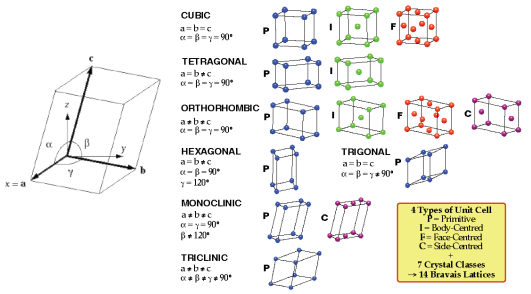
X-ray Crystallography

The Unit *Cell*



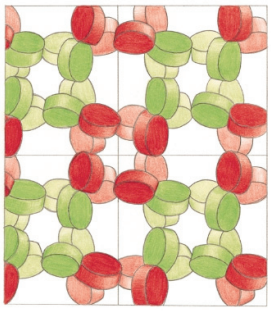
X-ray Crystallography

The Bravais Lattice Types



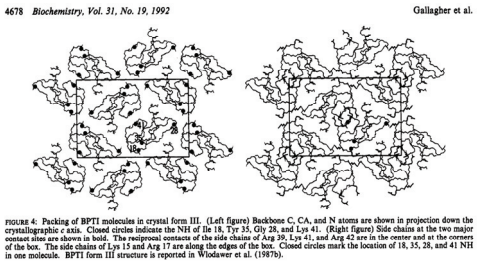
X-ray Crystallography

Protein Crystals



X-ray Crystallography

Protein Crystals

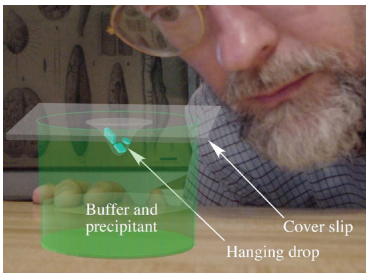


4678 *Biochemistry, Vol. 31, No. 19, 1992* Gallagher et al.

FIGURE 4: Packing of BPTI molecules in crystal form III. (Left figure) Backbone C, CA, and N atoms are shown in projection down the crystallographic *c* axis. Closed circles indicate the NH of Ile 18, Tyr 35, Gly 28, and Lys 41. (Right figure) Side chains at the two major contact sites are shown in bold. The reciprocal contacts of the side chains of Arg 39, Lys 41, and Arg 42 are in the center and at the corners of the box. The side chains of Lys 15 and Arg 17 are along the edges of the box. Closed circles mark the location of 18, 35, 28, and 41 NH in one molecule. BPTI form III structure is reported in Wlodawer et al. (1987b).

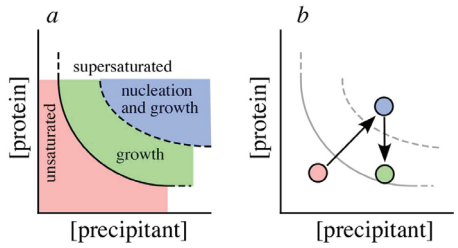
X-ray Crystallography

Growing Crystals



X-ray Crystallography

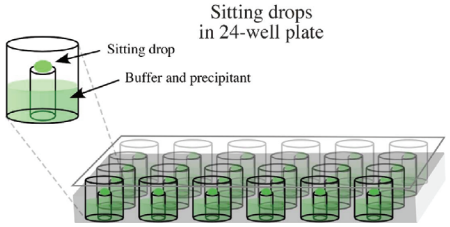
Growing Crystals



21

X-ray Crystallography

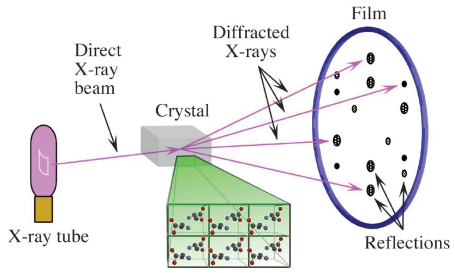
Growing Crystals



22

X-ray Crystallography

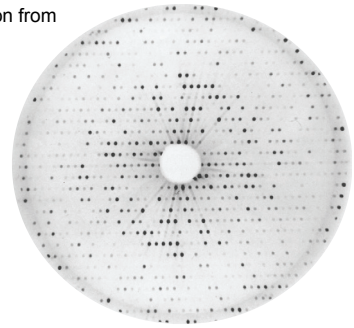
Diffraction from Crystals



23

X-ray Crystallography

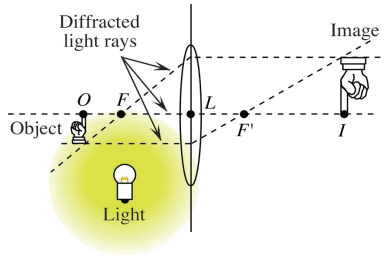
Diffraction from Crystals



24

X-ray Crystallography

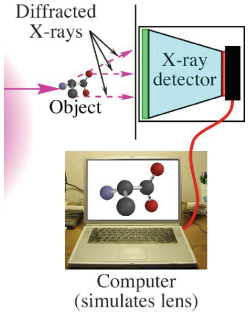
In light microscoscopy, lenses are used to collect the light scattered from an object and use it to form an image.



X-ray Crystallography

There are no X-ray lenses

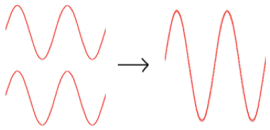
- Computers are used to simulate a lens
- To understand how this works, we need to understand the elements of a diffraction pattern



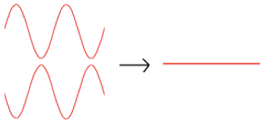
The Diffraction Pattern

Constructive and Destructive Interference

Constructive Interference

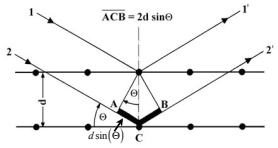


Destructive Interference



The Diffraction Pattern

Bragg's Law describes the conditions required for obtaining a spot.



- Constructive interference occurs when

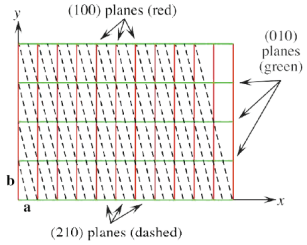
$$2d\sin(\Theta) = n\lambda \quad \text{(Bragg's Law)}$$

  - William Henry Bragg, and his son William Lawrence Bragg, shared the Nobel Prize in Physics in 1915.

The Diffraction Pattern

Crystal Diffraction

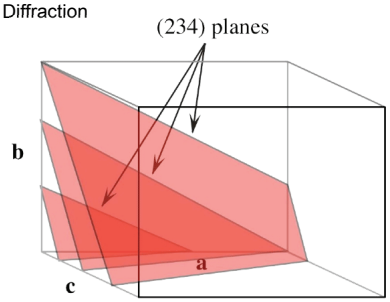
- The spots arise from reflections off of the crystal planes



29

The Diffraction Pattern

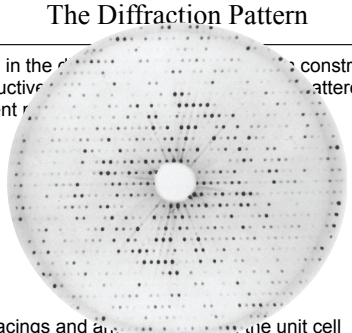
Crystal Diffraction



30

The Diffraction Pattern

The spots in the d… constructive and destructive… attered from the different …



- The spacings and a… …e unit cell dimensions and angles (*a, b, c, α, β* & *γ*)

31

The Diffraction Pattern

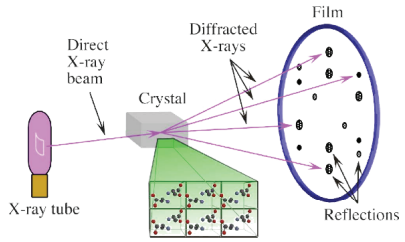
Demos

- XRayView 3.0
  - Simulates diffraction from a crystal
- Optical Diffraction
  - Diffraction patterns are reciprocally related to the object responsible for the diffraction

32

The Diffraction Pattern

The diffraction pattern is also referred to at the *reciprocal lattice*.



33

The Fourier Series

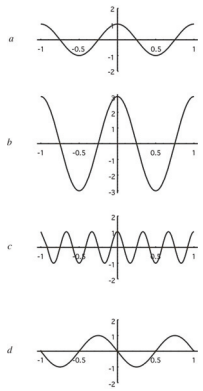
Fourier Series

- Sine and Cosine waves have
  - Amplitude, *F*
  - Frequency, *h*
  - Phase, *α*

$$f(x) = F_o \cos(2\pi[hx + \alpha])$$

or

$$f(x) = F_o \sin(2\pi[hx + \alpha])$$



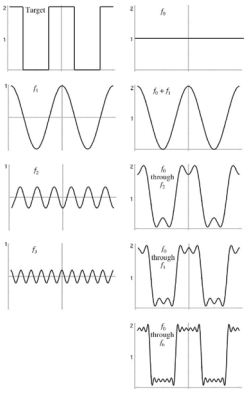
34

The Fourier Series

Fourier Synthesis

- Approximating a square wave by a Fourier Series

$$\begin{aligned} f(x) &= F_o \cos(2\pi[0x + \alpha_o]) \\ &+ F_1 \cos(2\pi[1x + \alpha_1]) \\ &+ F_2 \cos(2\pi[2x + \alpha_2]) \\ &+ F_3 \cos(2\pi[3x + \alpha_3]) \\ &\ldots \\ &+ F_n \cos(2\pi[nx + \alpha_n]) \\ &= \sum_{h=0}^{n} F_h \cos(2\pi[hx + \alpha_h]) \end{aligned}$$



35

Fourier Series

An alternative to designating a phase angle, α, is to combine cosine and sine wave

- The sine is equal to a cosine which has been shifted by an angle of $\pi/2$ (90°)

$$\sin(x) = \cos(x + \pi/2)$$

- This allows us to replace

$$f(x) = \sum_{h=0}^{n} F_h \cos(2\pi[hx + \alpha_h])$$

with

$$f(x) = \sum_{h=0}^{n} \left(F_{h,\cos} \cos(2\pi hx) + F_{h,\sin} \sin(2\pi hx)\right)$$

36

Fourier Series

It is also mathematically convenient to represent this relationship using complex numbers.

$$x = a + ib, \text{ where } i = \sqrt{-1}$$

$$f(x) = \sum_{h=0}^{n} F_h \left(\cos(2\pi hx) + i\sin(2\pi hx)\right)$$

where the amplitude, F_h, is also a complex number

- The cosine portion represents the *Real* part, while the sine portion represents the *Imaginary* part, of the complex number

Fourier Series

Another mathematical convenience comes from making use of Euler's formula:

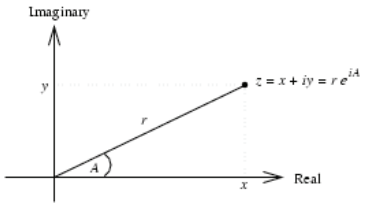
$$\cos(x) + i\sin(x) = e^{ix}$$

- This allows us to cast the Fourier series into a nice, compact, complex exponential form:

$$f(x) = \sum_{h=0}^{n} F_h e^{i2\pi hx}$$

Fourier Series

The complex polar coordinate system



Applying the Fourier Series to Diffraction

- The diffraction pattern also called the *reciprocal lattice*.
- Each spot in the diffraction pattern is called a structure factor
  - The structure factors are represented in three dimensions by the symbol

$$F_{h,k,l}$$

  - Where *h, k,* and *l* are the *Miller indices*, which describe the sets of planes that can be used to divide up the unit cell in the crystal lattice.

Applying the Fourier Series to Diffraction

- The diffraction pattern also called the *reciprocal lattice*.
- Each spot in the diffraction pattern is called a structure factor
  - The structure factors are represented in three dimensions by the symbol

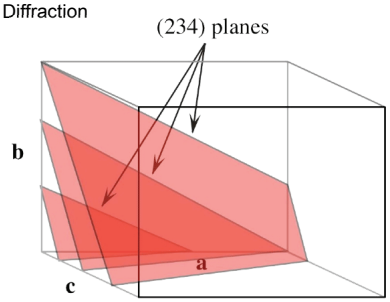
$$F_{h,k,l}$$

  - Where *h, k,* and *l* are the *Miller indices*, which describe the sets of planes that can be used to divide up the unit cell in the crystal lattice.
    - Remember, According to Bragg's Law, it is the reflections from these planes that give rise to the diffraction pattern.

41

Applying the Fourier Series to Diffraction

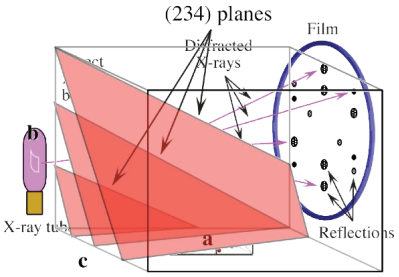
Crystal Diffraction



42

Applying the Fourier Series to Diffraction

The diffraction pattern is also referred to at the *reciprocal lattice.*



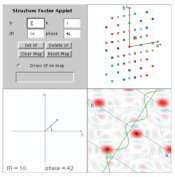
43

The Interactive Structure Factor Tutorial

Welcome to the structure factor tutorial! The aim of this tutorial is to learn about structure factors, phases, symmetry, and the relationship between the structure factors and the electron density map. The tutorial uses an interactive tool, the 'structure factor applet'.

The structure factor applet has four windows. Top left is the control window, in which you can select reflections and set their values by entering values in the text boxes. Several buttons allow you to modify the structure factor values in the map and the diffraction pattern. Top right is the structure factor window, which shows the current diffraction pattern. Bottom left is the reflection window, with which you can view or modify magnitudes and phases. Bottom right is the map window, showing the unit cell and electron density.

Magnitudes and phases are encoded as colours using the same scheme as the Book of Fourier.



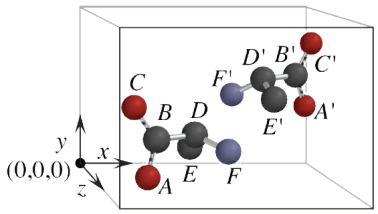
Small screen *or* slow browser	There are two versions of the structure factor tutorial. If you have a small screen (less than 1280x1024), or if you have a browser with a slow Java virtual machine (e.g. SGI machines), pick the links on the left.	**Big screen *and* fast browser**
Miller Indices. Structure factors(1). Structure factors(2). Symmetry(1). Symmetry(2). Patterson(1). Patterson(2). Direct Methods.	If you have a big screen (at least 1280x1024), and a browser with a fast Java virtual machine (e.g. Sun, Linux, Windows), pick the links on the right. If you are not sure about your machine, try both versions and use whichever works better for you.	Miller Indices. Structure factors(1). Structure factors(2). Symmetry(1). Symmetry(2). Patterson(1). Patterson(2). Direct Methods.

Bugs and fixes

44

Applying the Fourier Series to Diffraction

The structure factors (spots), $F_{h,k,l}$

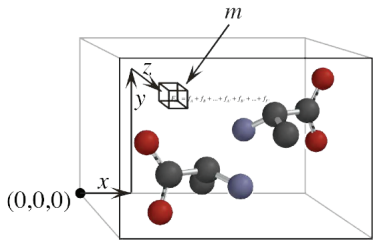


$$F_{hkl} = f_A + f_B + ... + f_{A'} + f_{B'} + ... + f_{F'}$$

Sum over atoms

Applying the Fourier Series to Diffraction

The structure factors (spots), $F_{h,k,l}$



$$F_{hkl} = f(\rho_1) + f(\rho_2) + ... + f(\rho_m) + ... + f(\rho_n)$$

Sum over volume elements

Applying the Fourier Series to Diffraction

The structure factor can be related related to the electron density using a complex exponential Fourier Series

- Each spot in the diffraction pattern is associated with a particular structure factor, *F*(*hkl*)

$$F(hkl) = \sum_x \sum_y \sum_z \rho(xyz) e^{2\pi i(hx+ky+lz)}$$

- Every location in the the unit cell contributes to each structure factor
  - If the electron density fluctuates with position at a frequency corresponding to *h*, *k* and *l*, Then this will produce a strong intensity for this particular structure factor, *F*(*hkl*)

Fourier Transforms

When one function represents the amplitude for the Fourier series of a second function then the dependent variables for the two functions are reciprocally related.

- In X-ray diffraction, the dependent variables for the electron density function, *x*, *y* and *z*, have dimensions of distance, where as the dependent variables for the structure factor, *h*, *k* and *l*, have dimensions of frequency, or one over distance.
- When this situation exists, there exists a mathematical operation called the *Fourier Transform*, can be used to convert between the two functions

Fourier Transforms

For example, for a one dimensional, discrete function, *F*(*h*), which is represented as the Fourier series of a second function, *f*(*x*):

$$F(h) = \sum_{x=0}^{\infty} f(x) e^{i2\pi hx}$$

- The *Fourier transform* is given as

$$f(x) = \sum_{h=0}^{\infty} F(h) e^{-i2\pi hx}$$

Fourier Transforms

When this is extended to 3-dimensions:

$$F(hkl) = \sum_{x}\sum_{y}\sum_{z} f(xyz) e^{i2\pi(hx+ky+lz)}$$

- The *Fourier transform* is given as

$$f(xyz) = \sum_{h}\sum_{k}\sum_{l} F(hkl) e^{-i2\pi(hx+ky+lz)}$$

Fourier Transforms

Substituting *ρ*(*xyz*) for *f*(*xyz*) we get:

$$F(hkl) = \sum_{x}\sum_{y}\sum_{z} \rho(xyz) e^{i2\pi(hx+ky+lz)}$$

- The *Fourier transform* is given as

$$\rho(xyz) = \sum_{h}\sum_{k}\sum_{l} F(hkl) e^{-i2\pi(hx+ky+lz)}$$

- This provides us with a method for determining the electron density function, *ρ*(*xyz*), from the structure factors, *F*(*hkl*)

Solving for the Electron Density Functions

$$\rho(xyz) = \sum_{h}\sum_{k}\sum_{l} F(hkl) e^{-i2\pi(hx+ky+lz)}$$

As indicated earlier, each structure factor *F*(*hkl*) is itself a complex number that is given by an amplitude, phase and frequency

- The Interactive Structure Factor tutorial demostrates this.

Solving for the Electron Density Functions

- The frequency is determined by the Miller indices (*h,k,l*), which determine the frequency of the planes cutting through the unit cell.
- The amplitude of each structure factor |*F*(*hkl*)|, can be determined from the intensity of each spot *I*(*hkl*):

$$\left|F(hkl)\right| = \sqrt{I(hkl)}$$

- What's missing is the phase information
  - We have essentially taken a black-and-white photo instead of a colored one

Solving the Phase Problem

For unit cells with small numbers of atoms, a Patterson Map can be used to determine the distances and directions between the atoms in the unit cell.

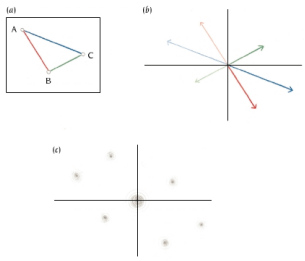
- A Patterson Map is constructed by assuming the phases are all zero

$$P(xyz) = \sum_h \sum_k \sum_l \left|F(hkl)\right| e^{-i2\pi(hx+ky+lz)}$$

- The Interactive Structure Factor tutorial demostrates this.

Solving the Phase Problem

- Patterson Map



Solving the Phase Problem

For protein crystals, can approximate this situation by using isomorphic replacement

- A small number of metal ions are introduced into the crystal.

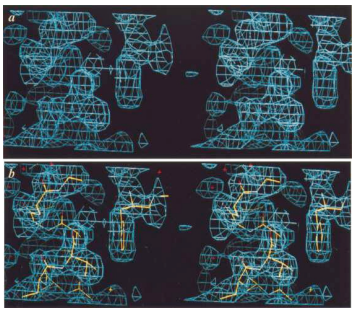
Another method used to solve for the structure factor phoases is *Molecular Replacement*

- In this method a homologous protein with a known structure is packed into the cell and used used to determine the phases using an *Inverse Fourier Transform*

$$F(hkl)_{calc} = \sum_x \sum_y \sum_z \rho(xyz) e^{i2\pi(hx+ky+lz)}$$

Model Building

A model is fit to the electron density:



Model Building

The model can be used to calculate a new set of phases:

$$F(hkl)_{calc} = \sum_x \sum_y \sum_z \rho(xyz)_{\text{model}}\, e^{i2\pi(hx+ky+lz)}$$

- And this is used to determine a calculated structure factor that can be compared to the observed structure factor:

$$R = \frac{\sum \left| |F(hkl)_{obs}| - |F(hkl)_{calc}| \right|}{\sum |F(hkl)_{obs}|}$$

Model Building

This is an iterative process