X-ray Diffraction and the Discovery of the Structure of DNA

A Tutorial and Historical Account of James Watson and Francis Crick's Use of X-ray Diffraction in Their Discovery of the Double Helix Structure of DNA

David T. Crouse
Department of Electrical Engineering, City College of New York, City University of New York, New York, NY 10031;
dcrouse@ccny.cuny.edu

> We wish to suggest a structure for the salt of deoxyribonucleic acid (DNA). This structure has novel features which are of considerable biological interest *(1)*.

The above statement is the first paragraph in James Watson and Francis Crick's 1953 article in *Nature* in which they described the correct structure of DNA. To recognize the recent 50th anniversary of this historic discovery, I decided to have undergraduate students perform the X-ray diffraction analysis of DNA and verify Watson and Crick's proposed structure. Unfortunately, even though Watson and Crick's article is clear and readable by anyone with a basic science degree, the X-ray diffraction analysis of the proposed structure by Maurice Wilkins, A. R. Stokes, H. R. Wilson, Rosalind Franklin, and Raymond Gosling in the two articles that immediately follow Watson and Crick's article in *Nature* may be difficult for undergraduate and graduate students to grasp. This is because even though X-ray diffraction is a topic covered in undergraduate and graduate physics, chemistry, and materials science and engineering courses and applied to various crystal structures, the X-ray diffraction analysis of DNA is more complicated and usually not analyzed in detail. In this article, I introduce a method of teaching the analysis of the X-ray diffraction of DNA through a series of steps using the original methods employed by Watson, Crick, Wilkins, Franklin, and Gosling. Each step is challenging but can be performed by upper-level undergraduate and graduate students and each step builds upon the previous steps, ending with the verification of Watson and Crick's proposed structure.

With this assignment, I give the students a copy of Watson and Crick's 1953 *Nature* article along with Maurice Wilkins's et al., and Rosalind Franklin's et al. 1953 *Nature* articles *(1–3)*. I also encourage them to read Watson's book *The Double Helix* that provides a riveting narrative of their struggles and ultimate success *(4)*. Throughout the steps of the X-ray diffraction analysis, I make connections between the results the students obtain and parts of the original 1953 articles. In this way, I reinforce the historical aspects of the project and perhaps provide a little more motivation for the students with them thinking that the analysis they are doing could have won them a Nobel Prize if they had done it prior to 1953. It is important to note that there are many ways to perform the X-ray diffraction analysis of DNA including modern day computer programs that do all of the work and produce the answer within minutes! However, the purpose of this work is to have the students read Watson and Crick's, Wilkins's et al., and Franklin and Gosling's articles, get into their mindset at the time of their discovery, and use the methods they used to make their groundbreaking discovery of the structure of DNA. There are several works by A. A. Lucas, Phillipe Lambin, and others that provide complementary historical accounts of these events with less emphasis on the mathematical details of the X-ray diffraction analysis compared to the current work *(5–7)*. Also, techniques to experimentally demonstrate in the classroom the X-ray diffraction of double helixes or other periodic patterns is outside the scope of this work but can be obtained from the Institute for Chemical Education and their excellent DNA Optical Transform Kit *(8)*.

The rest of this article is organized in two sections. Along with the introduction, the second and third sections are given to the students with the second section including the step-by-step calculations (one step assigned each week) and the third section including a discussion of the project and historical aspects. All of the questions posed in these sections are to be answered by the students. Additional hints and helpful information for the students are given in the Supplemental Material.[W] Complete solutions including all intermediate steps are also included as Supplemental Materials.[W]

The Assigned Project

X-ray Diffraction Background Information

Because this project involves the X-ray diffraction analysis of DNA, it is necessary to give a brief summary of the most important concepts of X-ray diffraction. A detailed description of X-ray diffraction is included in textbooks on the subject *(9)*. The electric field $\mathbf{E}$ of a diffracted X-ray beam can be written as a Fourier transform of the X-ray scattering elements (e.g., atoms, molecules, …),

$$\mathbf{E}(\mathbf{K}) = \int_{\substack{\text{all}\\ \text{space}}} \mathbf{A}(\mathbf{R},\mathbf{K}) \exp(i\mathbf{K}\cdot\mathbf{R})\,dV \tag{1}$$

where $\mathbf{K} = (\mathbf{k}_{inc} - \mathbf{k}_{diff})$ with $\mathbf{k}_{inc}$ and $\mathbf{k}_{diff}$ being the incident and diffracted X-ray wavevector, respectively. It is assumed that scattering of X-rays involves elastic scattering events and therefore $|\mathbf{k}_{diff}| = |\mathbf{k}_{inc}|$. $\mathbf{A}(\mathbf{R},\mathbf{K})$ is a term that is dependent on the density of scattering elements and the ability of an

infinitesimally small scattering element (of volume dV centered about **R**) to scatter an X-rays as a function of **K**. If we assume that each scattering element scatters X-rays equally and isotropically, then **A**(**R**, **K**) can then be replaced by the product of the density of scattering elements ρ(**R**) and a form factor F that is assumed in this work to be independent of **K**, leading to

$$E(\mathbf{K}) = \int_{\substack{\text{all}\\ \text{space}}} F\rho(\mathbf{R})\exp(i\mathbf{K}\cdot\mathbf{R})dV \qquad (2)$$

For the analysis of a crystalline structure with a lattice and a basis, eq 2 can be modified to be the product of two terms. The first term of the product, called the structure factor, is an integral over the scattering elements within one unit cell of the lattice (i.e., the basis):

$$S(\mathbf{K}) = \int_{\substack{\text{unit}\\ \text{cell}}} F\rho(\mathbf{r})\exp(i\mathbf{K}\cdot\mathbf{r})dV \qquad (3)$$

where **r** is the displacement vector representing the position of the infinitesimal scattering element relative to the lattice point. If the basis is composed of a finite number of discrete scattering elements, each with a displacement vector $\mathbf{r}_j$, then eq 3 becomes a summation:

$$S(\mathbf{K}) = \sum_j F\exp(i\mathbf{K}\cdot\mathbf{r}_j) \qquad (4)$$

The second term of the product, called the interference function, involves a summation over all lattice points with translation vectors **T** relative to an arbitrarily chosen origin:

$$I(\mathbf{K}) = \sum_{\mathbf{T}} \exp(i\mathbf{K}\cdot\mathbf{T}) \qquad (5)$$

For X-ray diffraction by fibers, such as multiple strands of randomly oriented strands of DNA in this work, E(**K**) is averaged over all orientations (described in more detail later). Generally, what is now done is to construct the reciprocal space lattice (RSL) using the restrictions on **K** that eqs 4 and 5 produce. Following this, the Ewald sphere is constructed. This Ewald sphere construction provides a convenient tool to extract information about the RSL from the angles of the diffracted X-ray; the reader is referred to elementary textbooks on X-ray diffraction for a detailed description of the Ewald sphere construction *(10, 11)*: the necessary equations from this technique for the analysis of the X-ray diffraction of DNA are included in a later section. After the Ewald sphere is constructed, angles of diffraction are obtained experimentally that give us information about the RSL that in turn, gives us information about the coordinate space lattice.

Project Game Plan

Because Watson and Crick (let us denote Watson and Crick as W&C from here on) believed that the phosphates were responsible for the observed X-ray diffraction pattern *(1)*, let us construct a model of the double helix structure of discrete phosphates as the product of a double *uniform* helix multiplied by a lattice of infinite continuous planes (Figure 1). The double *uniform* helix will provide the x and y coordinates

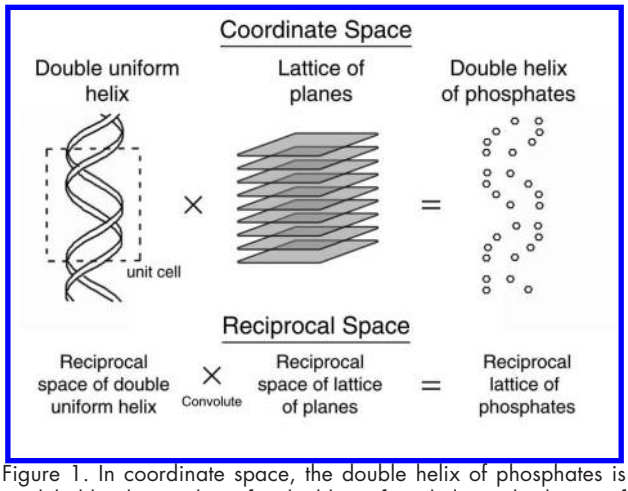

Figure 1. In coordinate space, the double helix of phosphates is modeled by the product of a double uniform helix and a lattice of planes. In reciprocal space, the complete reciprocal space lattice (RSL) of the double helix of phosphate is obtained by convoluting the RSLs of a double uniform helix and a lattice of planes.

of the phosphates and the infinite continuous planes will provide the z coordinates of the phosphates. Therefore, what will be done first is the calculation of the RSL of double *uniform* helices. Then the RSL of an array of infinite planes will be calculated. These two RSLs then will be convoluted to find the complete RSL of discontinuous double helix structure. Finally, the X-ray diffraction pattern will be deciphered using this complete RSL. This analysis of the X-ray diffraction of DNA will be done in a series of steps as shown in Figure 2. Many of the more difficult steps in this analysis have additional information and helpful hints online as supplementary material.[1]

Step 1a: Interference Function for a 1D Lattice of Points

Consider a 1D lattice of points with a spacing d. Show that the interference function, eq 5, is zero unless the following equation for K_z is satisfied (where m is an integer):

$$K_z = \frac{2\pi}{d}m \qquad (6)$$

Show that there are no such restrictions on K_x and K_y.

Step 1b: Structure Factor of a Double Uniform Helix

Both W&C and Franklin and Gosling (denoted from here on as F&G) stated in their articles that the phosphate groups are responsible for the X-ray diffraction and that they have to be on the outside of the helix *(1, 3)*. W&C state, "We believe that the material which gives the X-ray diagrams is the salt, not the free acid *(1)*." With the above opinions of W&C and F&G in mind, let us take one twist of a double *uniform* helix as the basis associated with the 1D lattice studied in step 1a. The structure factor of this basis will now have to be calculated.

Consider first just one *uniform* helix and then add in the second uniform helix. Use eq 6 stating that $K_z = 2\pi m/d$ and the definition of Bessel functions[W],

$$J_m(x) = \frac{i^{-m}}{2\pi}\int_0^{2\pi}\exp\{i[x\cos(\theta) + m\theta]\}d\theta \qquad (7)$$

Step 1. Calculate the RSL of a double uniform helix

1a. Calculate the interference function for a 1D lattice of points with spacing *d*.

1b. Calculate the structure factor for a single uniform helix with a length *d* per twist of the helix and radius *r*. Then add a second identical single uniform helix shifted relative to the first helix in the way that W&C, Franklin, and Gosling proposed and calculate the resulting structure factor.

1c. Using the results from steps 1a and 1b, construct the RSL of the double uniform helix.

Step 2. Calculate the RSL of an array of infinite continuous planes spaced a length *c* apart.

Step 3. Use the Convolution theorem for Fourier transforms and the RSLs obtained in steps 1 and 2 to construct the complete RSL of the discontinuous double helix lattice.

Step 4. Construct the Ewald sphere and calculate the theoretical comparison ratios that will be compared with their experimental counterparts.

Step 5. Compare the theoretical results with the experimental X-ray diffraction pattern and verify that the structure of DNA has the following properties:

Property 1. The structure is a double helix.

Property 2. The structure has 10 base pairs per twist.

Property 3. One helix is displaced relative to the other along their axes by $\delta = (3/8)d$.

Property 4. The helices have a periodicity of 34 Å.

Property 5. The radii of the helices are approximately 10 Å.

Figure 2. The steps for the analysis of the X-ray diffraction of DNA.

to justify F&G's assertion that the structure factor[ω] for the single uniform helix is *(3)*

$$S_m\left(K_\rho, \theta_K, K_z = \frac{2\pi}{d}m\right) = J_m\left(r K_\rho\right) \exp\left[i m \left(\theta_K + \frac{\pi}{2}\right)\right] \quad (8)$$

where r is the radius of the helix, d the length per twist of the helix, K_ρ is the radial component of the reciprocal space vector **K** (i.e., $K_\rho = [(K_x)^2 + (K_y)^2]^{1/2}$ and θ_K is the azimuthal angle in the K_xK_y plane.

Now add the second helix in the same way as is done in W&C's proposed structure for DNA *(1–3, 12–16)*. W&C, and F&G describe the proposed structure in the following paragraph of their articles[2]:

> W&C: This structure has two helical chains each coiled round the same axis… Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions *(1)*.
>
> F&G: The structure is probably helical… The structural unit probably consists of two co-axial molecules which are not equally spaced along the fibre axis, their mutual displacement being such as to account for the variation of observed intensities of the innermost maxima on the layer lines; if one molecule is displaced from the other by about three-eights of the fibre-axis period, this would account for the absence of the fourth layer line maxima and the weakness of the sixth *(3)*.

Let us verify F&G's statement that the second helix is displaced relative to the first helix by three-eights of the period of the helix or $\delta = (3/8)d$. Show that by adding in the second uniform helix, the structure factor becomes

$$\tilde{S}_m\left(K_\rho, \theta_K, K_z = \frac{2\pi}{d}m\right) = \left[1 + \exp\left(i\frac{3\pi}{4}m\right)\right] J_m\left(r K_\rho\right) \exp\left[i m \left(\theta_K + \frac{\pi}{2}\right)\right] \quad (9)$$

Do certain K_z layers drop out that were originally present for the single uniform helix? How about the $K_z = 2\pi 4/d$ [or the "fourth layer line" *(2, 3)*]? It will be clear later in the project why this displacement also produces a weak "sixth layer line" (i.e., $m = 6$) as noted by F&G and a strong "fifth layer line" (i.e., $m = 5$) as noted by Wilkins *(2, 3)*.

Step 1c: The Reciprocal Space Lattice of the Double Uniform Helix

Because we used the relation $K_z = 2\pi 4/d$ derived in step 1a, eq 9 is the final RSL of the double *uniform* helix and no further work needs to be done for this step.

Step 2: The Reciprocal Space Lattice for an Array of Infinite Continuous Planes with Spacing c

As was stated previously, the infinite array of uniform planes determines the z coordinates of the phosphate groups and will be used in conjunction with the double *uniform* helix to model the DNA structure. For this step, first calculate the interference function for a 1D lattice of points along the z axis with a spacing c. Then calculate the structure factor for an infinite plane oriented in the xy plane. Finally, combine these two results and show that the complete RSL is a 1D lattice of points[ω] with

$$K_x = 0; K_y = 0; K_z = \frac{2\pi}{c}m \quad (10)$$

Step 3: Constructing the Complete Reciprocal Space Lattice Using the Convolution Theorem: Fourier Transforms for the Birdwatcher[3] (17) or The Cochran–Crick–Vand Theorem (18)

So far in steps 1 and 2 the RSLs for two uniform helices (step 1) and an array of infinite planes (step 2) have been calculated. To find the RSL of the type of structure proposed by W&C, these two RSLs have to be convoluted. One of the first articles that addressed this issue was Cochran's et al. earlier article in 1952 that provided W&C, Wilkins, and Franklin with important theoretical information on the RSL of a discontinuous lattice *(18)*. Cochran et al. describes his

Table 1. K_z Values, and Local Maxima $(rK_\rho)_m^{\max}$ of $(|\Omega_{ave}(K_\rho, K_z = 2\pi m/d)$ for m = 0 → 5

m	Peak 1 $(rK_\rho)_m^{1\text{st max}}$	Peak 2 $(rK_\rho)_m^{2\text{nd max}}$	Comparison Ratios $(rK_\rho)_m^{1\text{st max}}/(rK_\rho)_1^{2\text{nd max}}$
0	0		0
1			1
2			
3			
4			
5			

NOTE: The theoretical comparison ratios will be compared with the experimentally obtained values.

use of the convolution theorem as it applies to the discontinuous helix *(18)*:

> Consider a function *H* which is zero everywhere except on a continuous helix, where it assumes the value unity, and another function *K* which is zero everywhere except on a set of horizontal planes of spacing *p*, where it assumes the value unity. The product *KH* of these two functions is a discontinuous helix. It follows that the transform of *KH* is the transform of *K*, convoluted with that of *H*.

Restating this in more mathematical terms, the convolution theorem states that if $\rho(\mathbf{r})$ is the product of $f(\mathbf{r})$ and $g(\mathbf{r})$ with Fourier transforms $F(\mathbf{K})$ and $G(\mathbf{K})$, respectively, the Fourier transform of $\rho(\mathbf{r})$ is

$$\Omega(\mathbf{K}) = \int G(\tilde{\mathbf{K}}) F(\mathbf{K} - \tilde{\mathbf{K}})\, d\tilde{\mathbf{K}} \tag{11}$$

Use the convolution theorem, your results for the two RSLs obtained in steps 1 and 2, and the assumption that there are an integer multiple *P* phosphate groups per twist of the helix (i.e., assume $d = Pc$) to show that the total RSL can be expressed as

$$\Omega\left(rK_\rho, \theta_K, K_z = \frac{2\pi}{d} m\right) = \sum_{n=-\infty}^{\infty} \left\{1 + \exp\left[i\frac{3\pi}{4}(m - Pn)\right]\right\} J_{(m-Pn)}(rK_\rho) \times \exp\left[i(m - Pn)\left(\theta_K + \frac{\pi}{2}\right)\right] \tag{12}$$

The quantity of interest is $|\Omega(K)|^2$ because it will be proportional to the intensity of the scattered X-ray beams. However, eq 12 predicts oscillatory behavior for $|\Omega(K)|^2$ about the K_z axis (i.e., as a function of θ_K) that will not be experimentally observed because the DNA strands may be rotating or you have many DNA strands with random orientations. Therefore, the true quantity of interest is the average value of $|\Omega(rK_\rho, \theta_K, K_z)|^2$ in the range of $\theta_K = 0 \rightarrow 2\pi$. Show that this quantity $|\Omega_{avg}(rK_\rho, K_z)|^2$ is

$$\left|\Omega_{avg}\left(rK_\rho, K_z = \frac{2\pi}{d} m\right)\right|^2 = 4\sum_{n=-\infty}^{\infty} \cos^2\left[\frac{3\pi}{8}(m - Pn)\right] J^2_{(m-Pn)}(rK_\rho) \tag{13}$$

Also, show that the following relations hold:

$$\left|\Omega_{avg}\left[rK_\rho, K_z = \frac{2\pi}{d}(-m)\right]\right|^2 = \left|\Omega_{avg}\left(rK_\rho, K_z = \frac{2\pi}{d} m\right)\right|^2 \tag{14}$$

$$\left|\Omega_{avg}\left[rK_\rho, K_z = \frac{2\pi}{d} m\right]\right|^2 = \left|\Omega_{avg}\left[rK_\rho, K_z = \frac{2\pi}{d}(m+P)\right]\right|^2 \tag{15}$$

Also at this stage, assume that the number of phosphates is as W&C predicted, that is, let $P = 10$. Equations 14 and 15 allow you to express $|\Omega_{avg}(rK_\rho, K_z = 2\pi m/d)|^2$ for *m* outside the range of $m = 0 \rightarrow 5$ as one of the $|\Omega_{avg}(rK_\rho, K_z = 2\pi m/d)|^2$ quantities in this range, for example,

$$\left|\Omega_{avg}\left(rK_\rho, K_z = \frac{2\pi}{d} 6\right)\right|^2 = \left|\Omega_{avg}\left[rK_\rho, K_z = \frac{2\pi}{d}(-4)\right]\right|^2 = \left|\Omega_{avg}\left(rK_\rho, K_z = \frac{2\pi}{d} 4\right)\right|^2$$

$$\vdots$$

$$\left|\Omega_{avg}\left(rK_\rho, K_z = \frac{2\pi}{d} 10\right)\right|^2 = \left|\Omega_{avg}\left(rK_\rho, K_z = 0\right)\right|^2$$

The above results show why the 6th layer line is as weak as the 4th layer line as F&G stated *(3)*. How do the above results also verify Maurice Wilkins's statement that the 5th layer line is strong *(2)*?

Now plot the functions $|\Omega_{avg}(rK_\rho, K_z = 2\pi m/d)|^2$ versus rK_ρ with $0 \le rK_\rho \le 15$ for $m = 0 \rightarrow 5$. Then fill in Table 1 with the values $(rK_\rho)_m^{1\text{st max}}$ and $(rK_\rho)_m^{2\text{nd max}}$ that produce the first and second local maximum for $|\Omega_{avg}(rK_\rho, K_z = 2\pi m/d)|^2$, respectively, for $m = 0 \rightarrow 5$. Also, calculate the theoretical comparison ratios $(rK_\rho)_m^{1\text{st max}}/(rK_\rho)_{m=1}^{1\text{st max}}$ and put these values in Table 1 as well.

Step 4: Construct the Ewald Sphere

This step has been done for you with a portion of the complete RSL of W&C's proposed DNA structure as shown in Figure 3. In this figure, the $K_z = 0$ and $K_z = 2\pi 5/d$ layer lines of the RSL are shown, the other K_z planes are left out for the sake of clarity. Then, assuming that Copper *K*α X-rays with a wavelength of 1.54 Å are being used, the Ewald sphere can be drawn (only a small portion of the Ewald sphere is shown in Figure 2). The bottom part of Figure 3 gives important angles and coordinates in coordinate space that will be of help when comparing the theoretical and experimental results.

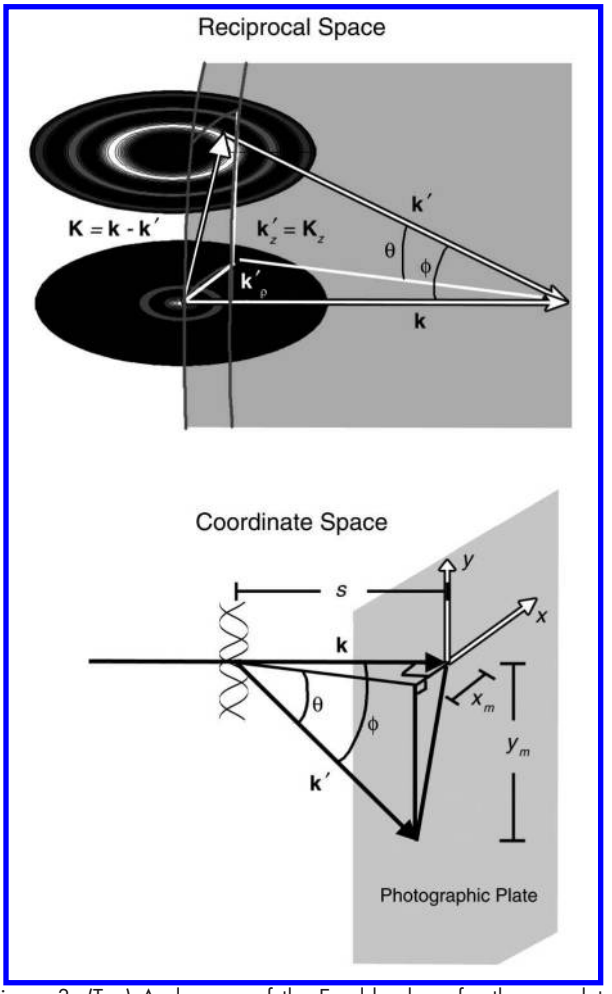

Figure 3. (Top) A close up of the Ewald sphere for the complete RSL with only the $K_z = 0$ and $K_z = 2\pi 5/c$ shown (i.e., the 0th and 5th layer lines). The incident wavevector **k**, diffracted wave vector **k′** and the reciprocal lattice vector **K** are shown. (Bottom) The experimental setup with the photographic plate, **k**, **k′**, and various angles and lengths including the camera constant *s*.

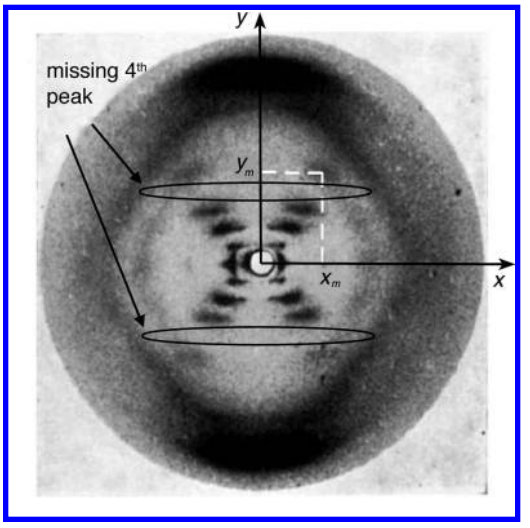

Figure 4. An X-ray diffraction pattern from Rosalind Franklin's 1953 article discussing the structure of DNA *(3)*.

Table 2. Diffraction Pattern Values x_m and y_m and Other Important Values

Peak	x_m/cm	y_m/cm	$(K_z)_m$/Å^{-1}	$(K_\rho)_m$/Å^{-1}	$(K_\rho)_m/(K_\rho)_{m=1}$
0	0	0	0	0	0
1	0.195	0.230			1
2	0.335	0.455			
3	0.520	0.715			
4	Missing	Missing			
5	0.800	1.190			

NOTE: Assume that $s = 5$ cm.

Step 5: Comparison of Theory and Experiment

You now have everything you need to verify that the X-ray diffraction pattern shown in Figure 4 corresponds with a structure that has properties 1–5 (see Figure 2). Properties 1–3 are verified in the following way. First, assume that you use your ruler and measure the x_m and y_m values that I have conveniently listed for you in Table 2. From these values of x_m and y_m, you will need to calculate the corresponding values of $(K_z)_m$, $(K_\rho)_m$, and the experimental comparison ratios $(K_\rho)_m/|K_\rho)_{m=1}$. To do this, you will need to use the following equations gleaned from Figure 4,

$$(K_z)_m = |\mathbf{k}|\sin(\theta_m); \quad \theta_m = \arctan\left(\frac{y_m}{\sqrt{x_m^2 + s^2}}\right) \tag{16}$$

where s is the camera constant and can be assumed to be 5 cm, and $|\mathbf{k}| = 2\pi/\lambda_{\text{X-ray}}$. Also, using the relations $|\mathbf{K}|_m = 2|\mathbf{k}|\sin(\phi_m/2)$ and $|\mathbf{K}|_m^2 = (K_\rho)_m^2 + (K_z)_m^2$, $(K_\rho)_m$ can be written as

$$(K_\rho)_m = \sqrt{4|\mathbf{k}|^2 \sin^2\left(\frac{\phi_m}{2}\right) - (K_z)_m^2}$$

$$\phi_m = \arctan\left(\frac{\sqrt{x_m^2 + y_m^2}}{s}\right) \tag{17}$$

Calculate the experimental values of $(K_z)_m$, $(K_\rho)_m$, and the comparison ratios $(K_\rho)_m/(K_\rho)_{m=1}$ for the m peaks and list these values in Table 2. Do the theoretical and experimental comparison values (i.e., $(rK_\rho)_m{}^{1\text{st max}}/(rK_\rho)_{m=1}{}^{1\text{st max}}$ and $(K_\rho)_m/(K_\rho)_{m=1}$, respectively) listed in Table 2 agree? If they do to within an acceptable error, then you will have verified properties 1–3.

After verifying properties 1–3, properties 4 and 5 are easily verified. Use eq 16 and the fact that $(K_z)_{m=1} = 2\pi/d$ to

calculate d. Use eqs 16 and 17 to calculate $(K_\rho)_{m=1}$ and the value of $(rK_\rho)_{m=1}{}^{1st\ max}$ listed in Table 1 to calculate r.

Congratulations, you have verified that the structure of DNA is a discontinuous double helix with properties 1–5! If you could have done this prior to W&C's 1953 discovery, you would have won the Nobel Prize!

Project Discussion and Conclusion

This project is suitable for upper-level undergraduate students or graduate students in a physics, chemistry, or materials science and engineering course who have learned the basics of X-ray diffraction. The time duration of each of the project sections and the entire project will depend on the specifics of the course but may be a two-month project with sections of the project assigned each week.

It is interesting to note that both the X-ray diffraction pattern analysis and chemical principles were used equally by W&C to make the final determination of the DNA structure. Without the use of either X-ray diffraction or the application of basic chemical principles, the discovery of the DNA structure as quickly as it was, would not have been possible. The X-ray diffraction pattern led to W&C's conclusion of the basic helical structure of DNA and its dimensions *(12)* whereas basic chemical principles led Watson to verify this structure along with the counter-oriented nature of the two helices, the hydrogen bonding between purines (adenine and guanine) and pyrimidines (thymine and cytosine) *(13),* and the implications to gene replication *(19).* This is an important aspect of the history of the discovery of the DNA structure because W&C and the Cavendish group at Cambridge University were in a race for the discovery with Linus Pauling at Caltech quick at their heels. Watson noted this in his *The Double Helix (20)*:

> Our first principles told us that Pauling could not be the greatest of all chemists without realizing that DNA was the most golden of all molecules. Moreover there was definite proof. Maurice had received a letter from Linus asking for a copy of the crystalline DNA X-ray photographs.

W&C, as well as the entire Cavendish group had further concerns when in 1951, Pauling discovered the α-helix describing the structure of proteins *(21).* With such a tremendous mind and scientific skills that Pauling had, they believed that his discovery of the structure of DNA was imminent. Their concerns were well founded and in late 1952, Pauling et al. put forth a hypothetical structure for DNA and published it in *Nature* in 1953 *(14).* While this caused the Cavendish group some initial concern, they disliked the proposed structure and very soon realized that the configuration of phosphates was impossible *(22).* Before Pauling realized his mistake and before he could propose another structure, W&C proposed their structure for DNA *(1).* W&C's proposed structure of DNA was quickly seen by all to agree with the X-ray diffraction data and chemical principles as well as having dramatic biological implications *(12–13, 19, 23).* It seems that most people at the time thought that Pauling would be upset that he had been beaten; however, it turned out that his feeling of "thrill" for the biological implications was stronger. After hearing about the proposed structure and seeing a model of it during a visit to England, Pauling's "reaction was one of genuine thrill" and he "effectively conceded the race *(24).*"

It is seen from the step-by-step approach to the X-ray diffraction analysis of DNA in this article that the analysis is not difficult. In fact Watson admits that he did not even know Bragg's law when he arrived in England in 1951 *(25).* Furthermore, he also admits that his chemistry skills were not what they should be *(26).* A combination of many factors including Watson and Crick's passion and intelligence, Watson's insights, Crick's and Wilkins's mathematical and scientific skills, Franklin and Gosling's X-ray diffraction patterns, and the competitive atmosphere produced an environment conducive to making this great discovery.

Notes

1. Whenever there is additional information about a specific step in the Supplemental Material, its existence is denoted by ω. The Supplemental Material also has information for instructors including all of the intermediate steps leading to the solutions for each step.

2. Linus Pauling and R. B. Corey at Caltech and Watson and Crick both seriously considered a three intertwined chains with the phosphates near the fiber axis *(27).* This hypothetical structure was shown to be incorrect and was a bit of an embarrassment for Pauling who published the faulty structure in *Nature (14).*

3. With the X-ray diffraction pattern and the Cochran–Crick–Vand theorem, Watson commented to Crick that the analysis of the X-ray diffraction pattern of DNA was so easy it could be analyzed by a "former birdwatcher". Crick had made this statement earlier with respect to Watson's analysis of the X-ray diffraction of the tobacco mosaic virus (TMV) *(17).*

ωSupplemental Material

Instructions for the students and notes for the instructor are available in this issue of *JCE Online*.

Literature Cited

1. Watson, J. D.; Crick, F. H. C. *Nature* **1953,** *171,* 737–738.
2. Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953,** *171,* 738–740.
3. Franklin, Rosalind E.; Gosling, R. *Nature* **1953,** *171,* 740–741.
4. *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980.
5. Lucas, A. A.; Lambin, Ph.; Mairesse, R.; Mathot, M. *J. Chem. Educ.* **1999,** *76,* 378–383.
6. Lucas, A. A. *Int. J. Quantum Chem.* **2002,** *90,* 1491–1504.
7. Lucas, A. A.; Lambin, Ph. *Rep. Prog. Phys.* **2005,** *68,* 1181–1249.
8. DNA Optical Transform Kit. *http://mrsec.wisc.edu/Edetc/background/DNA/DNA_OTK.htm* (accessed Feb 2007).
9. Cullity, B. D. *Elements of X-Ray Diffraction,* 2nd ed.; Addison-Wesley: Reading, MA, 1978.
10. Cullity, B. D. *Elements of X-Ray Diffraction,* 2nd ed.; Addison-Wesley: Reading, MA, 1978; pp 488–490.
11. Warren, B. E. *X-Ray Diffraction;* Dover Publications: Mineola, NY, 1990; pp 10–22.
12. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 98–108.

13. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 110–115.
14. Pauling, L.; Corey, R. B. *Nature* **1953,** *171,* 346–347.
15. Watson, J. D.; Crick, F. H. C. The Structure of DNA. In *Cold Spring Harbor Symposia on Quantitative Biology;* Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1953; Vol. XVIII, pp 123–131.
16. Watson, J. D.; Crick, F. H. C. *Nature* **1953,** *171,* 964–967.
17. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 69, 101.
18. Cochran, W.; Crick, F. H. C.; Vand, V. *Acta Cryst.* **1952,** *5,* 581–586.
19. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 108, 114, 123.
20. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; p 15.
21. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 24–25.
22. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 93–94.
23. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 122–131
24. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; p 127.
25. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; p 28.
26. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; p 17.
27. Watson, J. D. The Double Helix. In *The Double Helix,* A Norton Critical Edition; Stent, Gunther S., Ed.; W. W. Norton and Co.: New York, 1980; pp 56–57.