Harnessing the Power of Structural Oncology

Maria J. Solares\textsuperscript{1,2,3} and Deborah F. Kelly\textsuperscript{1,3*}
\textsuperscript{1}Center for Structural Oncology, Pennsylvania State University, University Park, PA 16802
\textsuperscript{2}Molecular, Cellular, and Integrative Biosciences Graduate Program, Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA 16802
\textsuperscript{3}Department of Biomedical Engineering, Pennsylvania State University, University Park, PA 16802
*dfk5267@psu.edu

Abstract: Microscopy has evolved through centuries of desire to see the unseen world. Modern technologies have pushed the limits of what's possible, inspiring the field of structural biology. Coupled with advances in biochemistry, scientists can now study the fine details of tiny intruders in the body as well as the molecular culprits of disease. Despite these monumental achievements, human health still suffers from persistent dysfunctions in biological molecules. Through cryo-electron microscopy (cryo-EM) imaging technology we can create more translatable models to uncover the atomic details of macromolecules gone awry. For instance, the tumor suppressor p53 is one of the most popular proteins in cancer research, yet there is little structural information about its complete 3D organization. The Center for Structural Oncology (CSO) at Penn State University is working to address these impactful uncertainties. Using a wealth of cryo-EM resources in the CSO, we capture dynamic snapshots of p53 assemblies while preserving their disease-related context for drug discovery and therapeutic development. With these tools in hand, the CSO is working to engineer new molecular paradigms in creating a world without cancer.

Keywords: cryo-EM, p53, tumor suppressor, structural oncology, cancer research

The Secret Life of Molecules

It happened on a Martin Luther King, Jr. (MLK) holiday. The usual buzz of the lab was absent; it was just professor, graduate student, and the hiss of machinery coming from the transmission electron microscope (TEM). It was still early in my lab rotation, and my soon-to-be mentor, Dr. Deb Kelly, was walking me through the minutia of cryo-electron microscopy (cryo-EM). She was imparting previous knowledge that had been cultivated in the lab, ranging from how to prepare samples for the microscope to the art of identifying regions of interest. It was supposed to be another training session, another day putting samples in the TEM, and another day completing proof-of-concept experiments. Little did we know that this day would change our entire career paths.

"Look at all of those molecules!" The air felt electric in that moment as we both huddled around the computer screen. The cryo-EM images revealed a molecular species that could no longer evade discovery. There on the screen, Dr. Kelly and I observed elusive molecules at the root of many human cancers, the tumor suppressor, p53.

They appeared to be beads on a string, and DNA strands in between the globular protein assemblies were clearly visible. Subsequent analysis later proved the samples contained different forms of the p53 complex. We have since resolved 3D structures for the different p53 assemblies naturally formed in cells derived from cancer patients. The weight of this discovery was even larger than either of us could have imagined. Those micrographs and subsequent models were the first step to creating a new scientific field named structural oncology.

As Dr. Kelly imparted the technical knowledge that set this new work into motion, I could not help but think about all the scientists that came before us. Our efforts to decipher new structures were not novel concepts. Rather, they were built on the knowledge of scientists who lived centuries before us. We were merely expanding their efforts, with the hopes of combating human disease by training the next generation of cancer researchers (Figure 1).

Let There be Light

Within the last few centuries, through the spirit of innovation, the microscopy field has advanced light years (Figure 2). The invention of the light microscope is attributed to several academic contributions originating in the 17th century, with each step along the way advancing the inherent capabilities of the instrument. By mid-century, simple microscopes were ubiquitous, which aided significant breakthroughs in biology and medicine. Giovanni Battista Odierna's work in 1644 produced the first insights of tissue anatomy through comprehensive descriptions of the fly's visual system [1]. Within a decade Robert Hooke published the work, Micrographia, and coined the term "cell" [2].

The imaging technology that fostered and advanced science eventually caused a plateau in generalized knowledge. Early on, scientists such as Hooke used microscopes that could only reveal features at low magnification. This technical limitation hindered Hooke's studies and interpretations of specimens. For instance, as it was not possible to see organelles in living tissue with low magnifications, it was incorrectly surmised that cells were not living organisms. Further, Hooke noted in Micrographia that there was no evidence of "seeds" in mold samples and suggested that the theory of spontaneous generation was behind the growth of these specimens [2,3]. Not until Antonie van Leeuwenhoek pushed the boundaries of light microscopy did scientists debunk such theories and replace them with new, more accurate ones.

Leeuwenhoek, with the encouragement of his colleagues, submitted over 200 letters to the Royal Society in support of the existence of unicellular organisms and common microbes [4]. Equipped with a single, high-quality lens of his own design, Leeuwenhoek could magnify his samples as much as 250–300 x [3]. At this level of resolution, the self-made scientist could see "animalcules," as small as 1–2 μm [3,5]. His imaging work also described the life cycle of fleas and maggots, pinpointing their origin from eggs, unlike previously believed [3]. This critical discovery was met with so much skepticism that the Royal Society issued an independent review by its members to verify the findings [4]. It would take another 200 years until Pasteur definitively debunked the theory of spontaneous...
creation. However, the use of microscopy had cemented its place at the center of new discoveries.

It's All a Matter of Waves

Light waves are the perfect gateway to understanding the macroscopic world. Early researchers could easily produce light waves for microscopes, and they had some of the best detectors available: their eyes. Other components and hurdles took time to overcome. Improvements in sample preparation were needed, hence the introduction of staining to improve contrast in specimen features. Enhanced contrast permitted scientists Camillo Golgi to study nervous tissue in 1873 and Hans Christian Gram to investigate bacterial species in 1884 [6]. To help with the problem of chromatic aberration, Ernst Abbe improved on an earlier lens design to ensure that all light waves would reflect to the same point [6].

Still, there was a theoretical resolution limit to light microscopes. Photon wavelengths average about 500 nm, which are small enough to visualize micro-organisms, whole cells, or large cellular features. If scientists wanted access to smaller units, like proteins, it was critical to use smaller wavelengths. The discovery of X-rays was a fortuitous event to accomplish this goal. While Wilhelm Roentgen’s accidental discovery of X-ray radiation led to a medical revolution, it was Max Laue and his team’s observations in 1912 that kickstarted the new imaging field, settling the controversy that X-rays were made of waves [7]. Theoretically, these wavelengths of ~0.1 nm could be used to resolve macromolecules at atomic detail.

In the early stages of X-ray technology development, there was no useful way to interpret diffraction patterns produced from bombarding minerals such as zinc sulphide (ZnS) with X-rays [7]. Lawrence Bragg proposed two theories to explain the diffraction data, which were used to write Bragg’s Laws. Guided by these principles, Bragg used diffraction data to demonstrate the alternating patterns of sodium atoms and chloride atoms in rock salt, and that there were no sodium-chloride complex bonds in the material [7,8]. Currently, X-ray crystallographers commonly use Bragg’s Laws to analyze complex patterns of repeating structures.

By adopting X-ray diffraction technology, meaningful strides were taken to determine the structures of macromolecules. Up until the 20th century, the scientific world regarded proteins as loose, heterogenous aggregates with no strong, inherent order. Protein studies were ingrained in “colloidal theory” [8–10]. Pepsin was the first enzyme to change this perspective. The work of J.D. Bernal and eventual Nobel Laureate Dorothy Crowfoot Hodgkin led to the first X-ray photographs of a protein crystal [11,12]. Another prime example of this groundbreaking technology was the famous X-ray photograph of DNA collected by Rosalind Franklin, which gave rise to Watson and Crick’s
hypothetical DNA structure [13,14]. As data in this area began to accumulate, the importance of biological macromolecules was recognized. The colloidal theory behind the existence of proteins gave way to new, modern descriptions of them, and they were identified as essential molecular machines and the pillars of the cellular world.

However, one problem persisted in the field, in that it was not mathematically easy to reconstruct a protein structure from X-ray diffraction data. Although it was possible to obtain diffraction patterns, it was difficult to calculate an electron density map of complex entities. This issue was further complicated by the fact that essential information was lost during data collection, as the phases of waves cannot be collected. By solving the “phase” problem by introducing heavy atoms into protein crystals along with the use of Fourier methods, Max Perutz and colleagues calculated the first electron density map for myoglobin [15,16]. After nearly 30 years of research, the atomic model was built into the density map to yield the first protein structure. Following this discovery, reports of other protein structures started to emerge. To date, over 180,000 different protein structures have been published and cataloged under the Protein Data Bank (https://www.rcsb.org).

The Use of EM Made Everything Crystal Clear

Almost in parallel to the discovery of X-rays, electrons were discovered by J.J. Thomson in 1897. He found that cathode rays were composed of charged particles, which were later shown to move in wave patterns by Schrödinger in 1926, as described by his famous wave propagation equations. Soon after, Ernst Ruska developed the first electron microscope by creating an apparatus that focused an electron beam. By the 1940s, electron microscopes were commercially available.

If X-ray crystallography already existed, why did electron microscopy (EM) become so popular within the same time period? The modern use of electron microscopes overcome many of the technical limitations required by crystallography, primarily the need to crystallize proteins. Proteins do not naturally exist as crystals, therefore creating samples for X-ray diffraction can be challenging. Not to mention, flexible regions in proteins are often excluded from crystallographic models due to their inability to crystallize. Cryo-EM allows scientists to omit this rigorous step by relying on vitreous ice encapsulation. Vitrified samples are somewhat protected from severe radiation damage by the nature of the frozen specimen, and results are quickly apparent in the form of micrographs (Figure 3). However, many specimen-related issues needed to be addressed for cryo-EM to yield atomic-level details comparable to crystallography samples. Enter the resolution revolution.

Initially, cryo-EM research was lovingly called “blobology.” The typical EM samples were larger macromolecules, such as viruses, embedded in heavy metal salts. This technique yielded 3D structures at lower resolutions when compared to results from X-ray crystallography. In 1975, Richard Henderson and Nigel Unwin published a 7Å structure of bacteriorhodopsin depicting its overall architecture. The structure included the seven transmembrane helices, which researchers were unable to resolve using other techniques [17]. In a moment of inspiration, Henderson and Unwin developed a novel way to collect images and diffraction patterns of unstained proteins placed on a metallic grid [18]. Although these models could not compete with resolutions obtained with crystallographic studies, scientists kept pursuing means to improve their imaging techniques. Jacques Dubochet and Alasdair McDowall were experimenting with cryogenic cooling methods to suspend particles in a glassy form of vitreous ice and published the first cryo-EM images of viruses [19,20]. Vitrification is still used today to prepare cryo-EM samples. In 1990, Henderson and colleagues used this new cryo-EM technique to resolve a high-resolution density map of bacteriorhodopsin at 3.5Å [21]. What was once considered a niche method was shaping up to be the future of structural biology.

Sample preparation was not the only improvement in the field. Joachim Frank and co-workers created a data analysis approach that would later give rise to “single-particle” EM. By using computational alignments of particle images via cross-correlation strategies, Frank’s team created statistically relevant particle averages with reduced noise to reveal better resolved features of the particles [22]. Particle averages were then classified based on different orientation values and used to calculate image reconstructions of the original 3D objects [23]. By publishing structures of asymmetric bacterial ribosomes, Frank’s team demonstrated the new computational algorithm without the use of inherent symmetry operators [24,25].
Although single-particle cryo-EM studies can provide meaningful snapshots of molecules in action, scientists are now seeking to apply the technique to more complicated and flexible proteins. Conformationally diverse entities limit the resolution of EM maps, while being under-appreciated mechanistically. This is an unfortunate reality, since many biological processes utilize the messy parts of molecules to complete their functions. Flexible stretches of individual proteins or multicomponent complexes are challenging to resolve due to the innate low contrast and high background noise inherent in the images. It is these proteins, however, that give rise to diseases such as glioblastoma multiforme, for which there is no known cure and survival rates are very low (Figure 4).

A much less recognized limitation is the sourcing of biological samples. For high-resolution studies, thousands of particle images must be acquired, which is time-consuming and resource-intensive. These procedures, in turn, are dependent on sample concentration. Recombinant protein technology allows scientists to produce and purify large volumes of individual proteins of interest. However, recombinant technology relies on its use of bacterial expression systems or yeast models. These organisms can serve as excellent sources for initial structural observations. In the age of personalized medicine and drug discovery, however, researchers seek to understand how structural changes can influence disease origins.

Hence, focusing on only model systems to study human health can lead to incomplete or incorrect interpretations. Studying structures of full-length proteins from native sources can yield more rigorous and critically relevant data. The lack of important regulatory components as structural considerations also diminishes the impact of healthcare science. As such, there is a clear need to examine complex diseases that have haunted humankind since their discovery.

**Viewing Cancer Through a New Lens**

The Center for Structural Oncology (CSO) is located at the Huck Institutes for the Life Sciences at Penn State University, where we use cryo-EM to fight cancer (Figure 5). Thanks to unfettered access to advanced technologies, we use materials science to address long-standing questions in medical research. Three main elements have spurred this pioneering effort: state-of-the-art imaging, complex materials, and novel biochemical approaches. Highly specialized electron microscopes with automated data collection routines cut down on the hours spent glued to the microscope, allowing us to advance at an unprecedented pace.

An image, however, is only as good as its sample. This is why CSO members focus on perfecting new materials for cryo-EM research. The perfect sample requires a conductive, flat surface with reliable physical properties under a variety of temperature extremes (Figure 6). In addition to conventional materials, we have identified silicon-based microprocessor materials as an ideal substrate. In particular, silicon nitride (SiN) and silicon dioxide (SiO) seem to fit the role. They can be produced in university nanoscience centers, and they are commercially available from different sources (for example, Protocips, Inc., SimPore, LLC). By using microchip materials, we are also developing systems that can be used for bioelectronics and biosensor applications.

Last, but certainly not least, we study protein assemblies sourced from patient-derived cancer cells. Developing translatable models for drug discovery is a top priority, although obtaining clean and highly concentrated native proteins is challenging. We have learned to adapt unconventional biological techniques to harvest cancer-related proteins of interest. By studying proteins such as the Breast Cancer-Related protein (BRCA1) and its binding partners, we determined in molecular detail new modification sites that diminish the function-mutated BRCA1 (Figure 7) [26]. Using current drug therapies, already approved by the FDA to treat other cancers, we can somewhat remedy the dysfunctional properties of mutated BRCA1 [26]. These treatments were shown to restore much of BRCA1’s original function in human cells. This is a major achievement for a burgeoning new center.
Revealing the Molecular Culprit of Cancer, p53

Often referred to as the “guardian of the genome,” the tumor suppressor, p53, is pivotal to many cell processes including cell cycle arrest, DNA repair, and apoptosis [27,28]. It is also one of the most deregulated proteins in all human cancers [29,30]. Surprisingly, the structure of p53 remains incomplete, limiting insights into target-based therapeutics. Full-length models have not been achieved since half of p53’s structure is disorganized and flexible. p53 can be broadly described as having three main regions, the N-terminal domain (NTD), the DNA-binding domain (DBD), and the C-terminal domain (CTD) (Figure 8A). The only region that has been studied successfully by X-ray crystallography is the DBD [31,32]. Although mutations in the DBD region are synonymous with cancer, the NTD and CTD are key regulatory regions that are heavily under-investigated in structural biology. Constructs of p53 produced in model cell lines are likely devoid of native modifications that influence p53 function. This type of information is also unresolved structurally.

We use a rapid biochemical procedure that exploits natural characteristics in the p53 molecule to harvest it from the nuclear material of cancer cells and enrich for it using immobilized metal affinity chromatography (IMAC) strategies (Figure 8B). SiN microchips are coated with special materials to further concentrate the native p53 proteins. Cryo-EM samples are then prepared using the microchip samples and imaging experiments conducted using our high-powered electron microscopes (Figure 8C, 8D). What started on that fateful MLK day as promising data was further brought to life by our team.

The combination of innovative biochemical strategies, advanced materials, and high-resolution imaging permitted us to determine the first full-length structures of p53 in different oligomeric states including monomers, dimers, and tetramers [33,34]. This work was revolutionary to us as we surpassed many of the limiting factors in the field. Each structure was derived from human glioblastoma cancer cells, and accurate modifications to the structure could be further analyzed. Most importantly, our EM maps included flexible regions, such as the NTD or CTD, that were previously unresolved (Figure 9).
Figure 7: EM analysis of wild-type BRCA1 and its binding partner, the BRCA1-Associated Ring-Domain protein 1 (BARD1). (A) EM image with inset of wild-type BRCA1 in association with its protein partner, BARD1 (left panel). Statistical averages computed for the BRCA1-BARD1 complexes (right panel) show features consistent with the molecules present in the images. Scale bar = 50 nm. (B) The EM structure of BRCA1-BARD1 is shown in different rotational views and assumes a clamp-like motif. Existing atomic models were used to interpret the structure. These included the BRCA1-BARD1 RING (really interesting new gene) domain (magenta; pdb code, 1JM7) and the BRCT (BRCA1-C-terminus) motif of BRCA1 (gray; pdb code, 1JNX). Scale bar = 1.5 nm. Adapted from Liang et al. [26].

Figure 8: The cell-to-structure pipeline. (A) The primary sequence of p53 is comprised of the N-terminal transactivation domain 1/2 motifs (TAD1/2), followed by a central DNA-binding domain (DBD). A predicted model for the N-terminus (cyan) is shown along with the DBD structure (blue; pdb code, 2AC0, A chain [32]). The oligomerization domain (OD) and C-terminal domain (CTD) (orange) constitute the most flexible region of the protein. (B) The isolation of p53 from human cancer cells is based on the protein’s innate ability to interact with metal cations through immobilized metal affinity chromatography (IMAC). Each fraction obtained from a biochemical preparation is analyzed biochemically to determine purity and oligomerization states of p53. (C) SiN microchips are used to further concentrate the native p53 proteins, which are imaged using cryo-EM (D). Scale bar = 50 nm. Adapted from Solares et al. [34].
At CSO, We Divide and Conquer

Humans were meant to be explorers. This can be seen in the development of modern technology that allows us to discover a whole new microscopic world. Every step of the way, we have overcome the physical barriers that have stopped us from seeing the unseen. We have molded light and electron waves, bent elements to form new materials, and even tamed molecules to aid us in isolating our particles of interest.

Above all, humans were meant to overcome disease. We find ways to apply our exploration and discoveries to better society. The CSO now harbors research teams that go above and beyond to determine the structure and function of disease culprits (Figure 10). By having each team focus on an individual component, we are now able to capture dynamic snapshots of macromolecules by using shared powerful resources. Together we are laying the groundwork for drug discovery and therapeutic development so...
that, as a society, we may rise up from a disease that robs us from our dignity. Together we shall begin ending the war against cancer.

References


Fast Focusing Motors & Nanopositioning Stages

7mm travel, nanometer resolution, msec response time, voice coil motor

Pinano® II XYZ piezo stage with advanced controller
High-stability XY piezo motor stage, 100mm
Low cost XY stages for instrumentation OEMs

AskPI@pi-usa.us | www.pi-usa.us | 50 Years Experience | Global Support