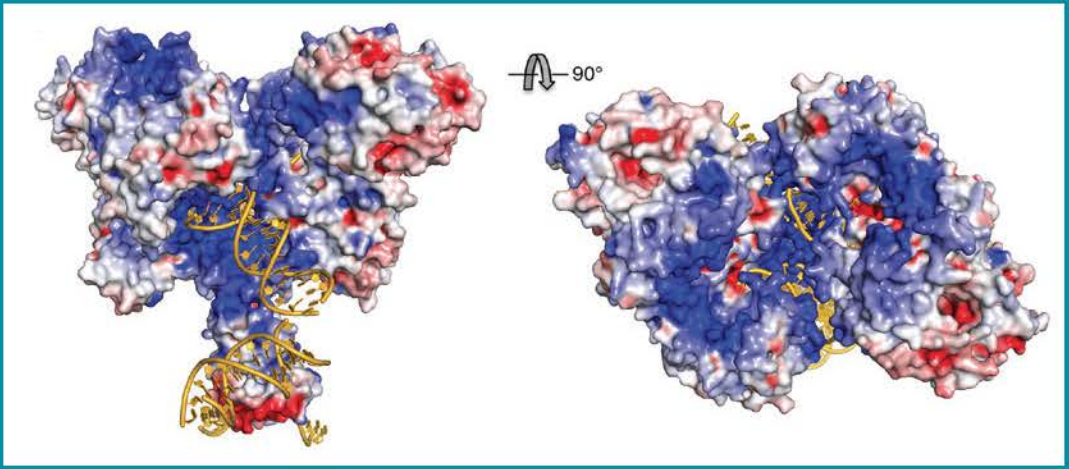

Determining the three-dimensional structure of a protein allows researchers to better understand its role in health and disease and identify the regions that are important for function. As described in this chapter, the RAG1 and RAG2 protein complex is critical to generating a diverse population of antibodies that defend us against pathogens but could also be culprits in autoimmune diseases. The images above show two views of the computer modeling of the RAG1-RAG2 complex associated with its DNA targets. The blue color represents regions of the protein that are positively charged, red represents negatively charged regions, and the bound DNA is in yellow. Modeling the complex revealed that its surface is highly positively charged, promoting its interaction with the negatively charged backbone of DNA. Understanding the three-dimensional structure of RAG1-RAG2 provides insight not only into how antibodies are made, but into human disease as well. Antibodies are essential for recognizing pathogens, while misguided antibodies can recognize the body s own proteins and lead to diseases that affect the pancreas, gastrointestinal tract, kidneys, and immune system.

Image courtesy of Drs. Wei Yang and Martin Gellert, NIDDK. Reprinted by permission from Macmillan Publishers Ltd: Nature, Kim M-S, Lapkouski M, Yang W, and Gellert M. Crystal structure of the V(D)J recombinase RAG1-RAG2. 518: 507-511, copyright 2015.

For an animated view of this image, please see http://www.nature.com/nature/journal/v518/n7540/fig_tab/nature14174_SV1.html

Cross-Cutting Science

Medical advances are not always achieved in great, intuitive leaps. More often, new prevention strategies, treatments, and cures result from a long, gradual accumulation of new knowledge from years of scientific research. Insights into the fundamental biologic building blocks and processes of an organism—its genes, the proteins they encode, the inner workings of cells, and the ways cells communicate with each other—can have broad and far-reaching implications. Indeed, many significant advances in our knowledge of disease and disease treatment can be traced to laboratory studies whose relevance to health could not have been fully known or appreciated at the time they were conducted.

With the development of innovative scientific technologies and the emergence of new scientific disciplines as talented and creative research teams join together to tackle ever more complex challenges, new opportunities to make exciting discoveries arise each day. The insights gained through this research can be expected to further scientific progress in many research areas, for today's discoveries may hold the seeds of tomorrow's cures.

Described in this chapter are several recent studies that illustrate the Institute's commitment to basic and applied research relevant across a broad spectrum of scientific disciplines. Also featured are the NIDDK's transdisciplinary research efforts to understand how microbial communities in the intestine affect health and diseases/conditions, especially obesity, digestive, and liver diseases. Another feature describes a new NIH program, co-led by the NIDDK, that aims to advance neuromodulation therapies toward precise neural control of organ function to treat diseases and conditions of interest to the NIDDK. This chapter also highlights research in the NIDDK's Intramural Research Program that is determining the structure of molecular machines and thus providing insights into human health and disease.

INSIGHTS INTO PROTEIN STRUCTURE AND FUNCTION

Solving Three-dimensional Structure of Immune System Proteins: Scientists in the NIDDK Intramural Research Program have determined the three-dimensional structure of a protein complex known as RAG1-RAG2 that is critical to the immune system. Because animals encounter a wide variety of potential infectious agents, the immune system must be able to generate a large and diverse population of antibodies to recognize these various agents. This diversity is generated by a process that joins different gene segments together in different combinations, a process known as "V(D)J recombination."

The proteins RAG1 and RAG2 are critical to this process; over 60 mutations that lead to severe combined immunodeficiency (SCID) in humans or a milder type of immunodeficiency called Omenn syndrome have been mapped to changes in the genes coding for RAG1 and RAG2. Understanding the three-dimensional structure, therefore, could provide insight not only into the critical process of V(D)J recombination, but into human disease as well. Solving the structure of mouse RAG1-RAG2 revealed that the disease-involved components of the proteins (amino acids), which are identical between mouse and human RAG proteins, fell into four classes based upon how the changes would disrupt the structure and/or affect RAG1-RAG2 activity. These data highlighted the amino acids critical to RAG1-RAG2, and may help explain how RAG1-RAG2 carries out its role in V(D)J recombination. *(See also the Scientific Presentation later in this chapter.)*

Kim M-S, Lapkouski M, Yang W, and Gellert M. Crystal structure of the V(D)J recombinase RAG1-RAG2. Nature 518: 507-511, 2015.

Protein Folding Machinery—GroEL Gives the Assist: Scientists from the NIDDK's Intramural Research Program have designed a new experimental approach to visualize better how a bacterial protein called GroEL helps other proteins to fold properly. Proteins carry out many of the biological functions underlying normal health and disease, and these molecules need to be "folded" into the right shape to properly function. In many instances, disease-causing mutations alter a protein's function by disrupting its three-dimensional structure, resulting in a misfolded protein that may accumulate in the cell and/or function poorly or not at all. Thus, an important part of biomedical research is to understand the relationship between the structure and biological function of proteins and how cells promote proper protein folding. GroEL, a member of the chaperonin family of proteins, assists other proteins in the folding process and is also suspected of having the ability to recognize misfolded proteins and unfold them to allow proper refolding. However, getting a clear picture of how GroEL interacts with the proteins it is folding has been difficult, because GroEL essentially engulfs the protein. Conventional techniques for determining molecular structures are unable to "see" the engulfed protein or how it is being manipulated in GroEL's folding chamber. As a result, very little is known about precisely how GroEL folds or unfolds proteins.

To obtain more information about how GroEL works, researchers took advantage of several different nuclear magnetic resonance (NMR) techniques. These techniques, instead of looking at the outside of GroEL as it folds a protein, examine the activity of GroEL and of the folding protein itself to collect detailed information about molecules' structures and how they are moving. Researchers used these techniques to investigate how GroEL interacts with a modified form of the protein Fyn SH3. Fyn SH3 exists largely in the folded state but also naturally unfolds a small amount of the time. Scientists monitored the transition of Fyn SH3 from unfolded to folded and back again, demonstrating that GroEL accelerates this transition by about 20-fold. Notably, these events took place in the absence of any cofactors or accessory proteins, giving conclusive evidence that GroEL possesses an innate or intrinsic ability to both fold and unfold proteins. Furthermore, the researchers were able to reconstruct how Fyn SH3 interacts with the inside of the GroEL folding chamber, identifying for the first time what parts of the two structures interact as Fyn SH3 folds and unfolds. These results provide new information about how GroEL (and possibly other protein-folding proteins) functions and how it helps protect cells from the detrimental effects of misfolded proteins.

Libich DS, Tugarinov V, and Clore GM. Intrinsic unfoldase/foldase activity of the chaperonin GroEL directly demonstrated using multinuclear relaxation-based NMR. Proc Natl Acad Sci USA 112: 8817-8823, 2015.

Friends of NIDDK Coalition Holds Congressional Reception

On June 23, 2015, the Friends of NIDDK held their inaugural event, a Congressional Reception. The Friends of NIDDK is a coalition of approximately 40 organizations who share the Institute's commitment to research to improve the lives of people living with, or at risk for, diseases and conditions within the NIDDK's mission. Through the coalition, these organizations have come together to speak with one voice, with the goal of promoting and sustaining the vital research activities of the NIDDK. Their collective interests span the Institute's research portfolio, including diabetes and other endocrine and metabolic diseases, liver and other digestive diseases, obesity, nutrition, kidney diseases, urologic diseases, and hematologic diseases.

At the June event, NIDDK Dr. Griffin P. Rodgers and other NIDDK staff had the opportunity to speak with leaders from many of the Friends organizations. Welcoming remarks were provided by Ms. Shereen Arent, Chief Advocacy Officer at the American Diabetes Association, who expressed the shared interests of the Friends organizations in research supported by the NIDDK. She also commented that nearly everyone's family is affected by a disease within the NIDDK's mission. Dr. Rodgers then addressed the group, highlighting the NIDDK's research areas and expressing appreciation to the attendees for their efforts in support of biomedical research.

Three Congress Members then provided remarks: Senator Dick Durbin (D-Illinois), Senator Jeanne Shaheen (D-New Hampshire), and Representative Robin Kelly (D-Illinois). Senator Durbin spoke of his legislative work to promote biomedical research, and of the new bipartisan Senate NIH Caucus, which he co-launched with Senator Lindsey Graham (R-South Carolina) earlier in 2015. Senator Shaheen, who had extended the invitation to Dr. Rodgers for this event, also spoke of the many benefits of research. Representative Kelly expressed her support and spoke in particular about health disparities, a topic relevant to many of the diseases in the NIDDK's mission. Congressional staff also attended the event, representing both Republican and Democratic members.

The NIDDK greatly appreciates the support and passion for biomedical research that was clearly evident at this event, and looks forward to the continued efforts of the Friends of NIDDK.

Left to right: Dr. Rodgers addressing the Friends of NIDDK; Dr. Rodgers with Sen. Durbin (D-Illinois), with Sen. Shaheen (D-New Hampshire), and with Ms. Arent (American Diabetes Association) and Rep. Kelly (D-Illinois). Photo credit: Hilary Schwab Photography

NIDDK Workshop: How Microbes Residing in the Gut Influence the Health of Their Hosts

Humans host trillions of microbes in the gut, including bacteria and viruses and other microorganisms, referred to collectively as the "gut microbiota" or "gut microbiome." These diverse intestinal inhabitants process dietary components and contribute to nutrient absorption and metabolism, interact with the immune system, and produce molecular signals that affect the intestines and other organs throughout the body. The types and proportions of different microbes in the microbiome can vary, and these differences may confer different effects on health. Although distinct from the bacteria and viruses that typically cause infectious disease, various communities of resident gut microbes have been associated with other diseases and conditions, such as obesity, diabetes, and inflammatory bowel disease.

In September 2014, the NIDDK convened a workshop to advance understanding of how disease pathways and normal biological processes are affected by gut microbes and to define research needs and opportunities for further investigating host-microbiome interactions. These include both exploration of how gut microbes influence the health of their hosts, and how differences among hosts might affect the composition of the gut microbial community, for example, through diet, inflammation, or genetics. As researchers have been cataloguing the numerous types of bacteria in the gut, it will be valuable to identify the many molecules these microbes produce and what functions they may have. While scientists have observed correlations between the gut microbiome and multiple diseases, future research would help determine whether certain microbial communities actually cause these diseases, and if so, how. These research avenues would benefit from collaborations among researchers from diverse disciplines, including basic science and clinical investigation, the study of microbes and animal

models, and the analysis of large amounts of complex data (bioinformatics). Finally, there is considerable interest in manipulating the microbiome for health benefits. For example, there is a clinical application of fecal microbiome transplantation to restore a normal microbiome for treatment of recurrent infections with a type of bacteria called *Clostridium difficile*, which can be severe. However, it is not clear whether manipulation of the normal gut microbiome may affect other diseases processes, or whether any of the molecules these microbes produce might be useful as therapeutics.

Following this workshop, the NIDDK released Funding Opportunity Announcements to encourage further research: "A Community Research Resource of Microbiome-Derived Factors Modulating Host Physiology in Obesity, Digestive and Liver Diseases, and Nutrition" (*RFA-DK-15-012*) and "Exploratory Studies for Delineating Microbiome: Host Interactions in Obesity, Digestive and Liver Diseases and Nutrition" (*RFA-DK-15-013*).

SPARC and NIDDK Organs of Interest

NIDDK leadership provides direction to the NIH Common Fund Program "Stimulating Peripheral Activity to Relieve Conditions" (SPARC). Peripheral nerves, the nerves outside of the brain and spinal cord, make connections with and influence the function of every organ in the body. Modulation of peripheral nerve signals (neuromodulation) to control the functions of the organs they supply has been recognized as a potentially powerful way to treat many diseases and conditions, such as hypertension, heart failure, gastrointestinal disorders, diabetes and its complications, inflammatory disorders, and more. However, the underlying physiology and mechanisms of action for neuromodulation therapies are poorly understood. The design of more effective and minimally invasive neuromodulation therapies requires knowing exactly how to stimulate nerves to achieve the desired effect on organ function. It also requires knowing exactly what nerves must be avoided to prevent unwanted side effects.

SPARC is uniquely positioned to serve as a community resource that provides the broader public and private research communities with the scientific foundation necessary to advance neuromodulation therapies towards precise neural control of end-organ system function to treat diseases and conditions. This high-risk, goal-driven program is structured as a consortium of four distinct research areas that will function in an integrated and iterative way, fostering discovery and broad dissemination of the fundamental physiology and biological mechanisms underlying peripheral control of internal organ function (autonomic nervous system) and sensory control, along with changes attributable to disease states and conditions. In turn, these discoveries will enable development of next generation closed-loop neuromodulation therapies, investigation of approved devices for new indications, and adoption of improved computational tools and modeling methods.

To spur applications to develop new and/or enhance existing tools and technologies tailored to elucidate the neurobiology and neurophysiology underlying autonomic control of internal organs in health or disease, which will ultimately inform next-generation neuromodulation therapies, the SPARC program released a research solicitation (Funding

Opportunity Announcement), "Exploratory Technologies to Understand the Control of Organ Function by the Peripheral Nervous System for SPARC," in January 2015. Of 12 projects funded in September 2015, eight related to NIDDK organs of interest:

Gastric and Intestinal Motility Disorders

- An implantable wireless system to study gastric (stomach) neurophysiology
- Defining gastric vagal mechanisms underlying nausea and vomiting using novel electrophysiological and optical mapping technology
- Establishment of *in vitro* and *in vivo* models of human gastrointestinal organoids with a functional enteric nervous system

Bladder Dysfunction and Pain

- Generate novel viral tools that target sensory neurons for control of bladder function and pain
- Develop a novel multi-modal spinal root interface to better understand nervous system control of the bladder
- Integrate optogenetic technologies with fully implantable wireless systems for inhibition of neurons innervating the bladder

Colon and Bladder Dysfunction

- Selective peripheral neuromodulation through transfer of genes designed for this purpose into organ-specific nerve cells

Acute Kidney Injury

- Tailoring ultrasound technology and imaging to explore mechanisms of neuromodulation of inflammation to protect kidneys from acute kidney injury

For more information on SPARC, please see: *http://commonfund.nih.gov/sparc/index*

Dr. Wei Yang—
Molecular Machines in DNA Biology

Dr. Wei Yang is the Chief of the NIDDK Intramural Research Program's Mechanism of DNA Repair, Replication, and Recombination Section in the Laboratory of Molecular Biology. Dr. Yang received her Ph.D. in Biochemistry and Molecular Biophysics from Columbia University and conducted her postdoctoral work at Columbia and Yale Universities. She joined the NIDDK's Intramural Research Program in 1995 and was tenured in 2000. A fellow of the American Association for the Advancement of Science since 2012, Dr. Yang was also elected to the National Academy of Sciences in 2013 and the American Academy of Arts and Sciences in 2015. Her research focuses on DNA recombination, repair, and replication using X-ray crystallography, molecular biology, and other biochemical and biophysical approaches to understand the mechanisms underpinning these biological processes. At the September 2015 meeting of the National Diabetes and Digestive and Kidney Diseases Advisory Council, Dr. Yang presented a lecture highlighting two recent advances to understand how DNA processes are carried out at the molecular level.

Critical Molecular Processes Protect DNA

Cancers are initiated by changes (mutations) or aberrations in DNA—changes that could be inherited or could result from exposure to a cancer-causing agent. To prevent these changes, cells have protective mechanisms, the three "Rs" of DNA as Dr. Yang calls them: replication, repair, and recombination. Cells replicate their own DNA to create new, identical, healthy cells. Repair pathways correct mistakes or damage to DNA. Recombination is a process that can repair damaged DNA as well, and enables genetic diversity by allowing pieces of DNA to move in the genome; as discussed below, this process is critical for the immune system. When these protective pathways become inactive or malfunction, disease can result. For example, a genetic cause of hereditary colon cancer is mutations that affect the mismatch pathway by which cells reduce DNA mutation rates. Dr. Yang's research investigates these pathways that maintain genomic stability. Although it may seem as though she is studying many different pathways, all of these processes converge on DNA. The processes are linked and share commonalities, as Dr. Yang and her colleagues have discovered.

Dr. Yang utilizes a variety of biochemical and biophysical strategies to study these pathways. In particular, she and her colleagues utilize techniques in structural biology—a field that focuses on elucidating the three-dimensional configurations of molecules, and how the

details of molecular structures influence their functions. Determining how a molecule is built at the atomic level provides insight into how it works mechanically and reveals how mutations disrupt its function. Structural biology provides detailed portraits of molecules, identifies their dynamic states, and reveals the molecular properties that allow them to function. With recent improvements in techniques and methodology, Dr. Yang's research has been able to answer more challenging questions in DNA biology and to probe the mechanics of DNA processes more deeply.

Key Insights into DNA Recombination in the Immune System

In her first example, Dr. Yang discussed her research that determined the structure of the RAG1-RAG2 protein complex, an important enzyme in the immune system. The RAG1-RAG2 complex helps create diversity in antibodies, proteins that are produced by the immune system to identify and target pathogens. Because the body can encounter a large number of different pathogens, this ability to create antibodies that recognize diverse pathogens is critical to survival. The RAG1-RAG2 complex helps create this diversity by recombining—cutting and rejoining—segments of DNA that encode the pathogen-recognizing sites of antibodies. This process is called "V(D)J recombination" for the different V, D, and J segments of DNA that are shuffled to form a large variety of combinations.

Twenty years after the discovery of V(D)J recombination, Dr. Martin Gellert's group in the NIDDK's Intramural Research Program reconstituted a critical step *in vitro* (in a test tube) and showed that the RAG1-RAG2 complex cleaves the DNA at specific DNA sites, initiating the recombination process. When Dr. Yang joined the NIDDK in 1995, she and Dr. Gellert launched a collaboration to study how the RAG1-RAG2 complex functions by understanding its three-dimensional properties. It took many different innovations and another 20 years of effort to determine the structure of RAG1-RAG2; the breakthrough research paper describing their findings was published in 2015 (see also "Solving Three-dimensional Structure of Immune System Proteins" in the Cross-Cutting Science chapter). Dr. Yang explained that one of the challenges they faced was that only 1 percent of the purified protein was functional from the insect cells they were using to make the protein. Although insect cells are often used for protein production, the researchers eventually found a mammalian cell type that was more permissive—generating over 50 percent functional purified protein—which was the quantum leap they needed.

Dr. Yang and her colleagues utilized a method called X-ray crystallography to determine the structure of the RAG1-RAG2 complex. In this technique, the arrangement of atoms in a crystallized protein or protein complex can be determined by the pattern in which x-ray light is diffracted from the crystal. The structure, determined by Drs. Yang and Gellert's group, provided important insights into how RAG1 and RAG2 combine to form the active complex. In addition, the structure explained over

60 mutations in human immune disease. These changes in the DNA sequences of the *RAG1* and *RAG2* genes (which encode the RAG1 and RAG2 proteins) cause severe combined immunodeficiency, a disease in which people are highly susceptible to severe, recurrent infections, or a milder form known as Omenn syndrome. Dr. Yang and her colleagues were able to predict how the disease-causing mutations affect the RAG1-RAG2 complex's structure and function. This important discovery not only provides answers to one of the biggest and most important questions in biology about how V(D)J recombination proceeds, it also reveals critical information about human disease and could lead to improved diagnostics and treatments.

Watching DNA Repair in Action

In a second example, Dr. Yang described her research visualizing the process of DNA synthesis, by studying a protein called "DNA polymerase η." DNA polymerase η replicates DNA that has been damaged by ultraviolet (UV) light. This process is important; people who have defects in this protein have a severe sensitivity to sunlight—a heritable disease called xeroderma pigmentosum. Sunlight causes some of the building blocks of DNA to become linked aberrantly, making a "UV lesion." This linkage blocks DNA and RNA synthesis at that site; DNA polymerase η allows bypass of the blockage. Dr. Yang and her colleagues were able to determine the structure of DNA polymerase η, revealing that that the UV lesion fits snugly into DNA polymerase η, whereas a normal DNA building block does not fit well; therefore DNA polymerase η is prevented from inappropriately working on undamaged DNA.

This discovery answered questions about DNA repair and synthesis and, not surprisingly, generated many more. DNA polymerase η is just one of the many polymerases found in human cells. Although Dr. Yang and her colleagues had determined how DNA polymerase η functions in bypassing UV lesions, they were interested in probing more deeply into the chemical mechanism for DNA synthesis, which is similar among all DNA polymerases. They also wanted to explore fundamental questions, such as how the many polymerases of human cells slightly differ from those found in bacteria and viruses, in order to design inhibitors to eliminate pathogens without preventing our own enzymes from working. To answer these questions, Dr. Yang needed to see DNA polymerase η in action. Solving a single crystal structure is like taking a still photograph, she noted. While looking at a photograph can give an idea of how something might work, a series of photographs of a dynamic process is needed to reveal the sequence of steps and transiently associated key ingredients. As an analogy, Dr. Yang told a famous story: in 1878, the English photographer Eadweard Muybridge was hired to help solve a popular question of the day—when a horse runs, are all four feet off of the ground? Muybridge took a series of photographs of a trotting horse. In many of the photographs, the horse had at least one foot on the ground. However, in one photograph all four feet were off the ground. If Muybridge had only looked at one photograph in the series, he would have likely answered "no" to the question. The series of

photographs, however, revealed the true answer of the dynamic process.

Dr. Yang and her colleagues took the same approach to watching DNA polymerase η in action. In an amazing feat, they were able to control the activity of DNA polymerase η and obtain sequential crystals of the DNA polymerase η in successive stages in action, and determine their structures. One of the observations they made was that a previously unknown third magnesium ion was essential to the reaction, and that this third magnesium ion is essential for every DNA polymerase. However, each polymerase has a different environment surrounding the third magnesium ion, providing a handle to design unique inhibitors for individual proteins. The identification of the critical role of this metal ion could lead to new drug targets and provides a new understanding of how this important protein functions.

Conclusion

Structural biology, as Dr. Yang discussed, provides key information to determining the causes and treatments for disease beyond identification of the genetic mutation. Understanding the structure of RAG1-RAG2 revealed how specific mutations in this complex lead to severe combined immunodeficiency and Omenn syndrome. Being able to watch DNA polymerase η in action exposed a critical metal ion that had not been previously appreciated and opened new avenues for developing potential new treatments for cancers and infections. A protein's structure tells the story of its function, pinpoints the key regions, and explains how mutations lead to disruption of its activity. These details are critical to designing new, efficacious drugs and revealing new targets. As Dr. Yang demonstrated, structural biology is an essential tool and continues to play a central role in biomedical research.