Protein Structure, Dynamics, and Function: A Philosophical Account of Representation and Explanation in Structural Biology

by

Jacob P. Neal

BS, Biochemistry & Molecular Biology, University of Richmond, 2008

BA, Women, Gender, & Sexuality Studies, University of Richmond, 2008

MA, Philosophy, University of Memphis, 2010

MS, Chemistry, University of Maryland, Baltimore County, 2012

MSc, Philosophy of Science, London School of Economics and Political Science, 2013

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Jacob P. Neal

It was defended on

July 27, 2021

and approved by

Sandra Mitchell, Distinguished Professor,
Department of History and Philosophy of Science

Anya Plutynski, Associate Professor, Department of Philosophy,
Washington University in St. Louis

Dissertation Co-Director: Michael Dietrich, Professor,
Department of History and Philosophy of Science

Dissertation Co-Director: James F. Woodward, Distinguished Professor,
Department of History and Philosophy of Science

Protein Structure, Dynamics, and Function: A Philosophical Account of Representation and Explanation in Structural Biology

Jacob P. Neal, PhD

University of Pittsburgh, 2021

Most philosophical work in molecular biology has historically centered on DNA, genetics, and questions of reduction. My dissertation breaks from this tradition to make proteins the object of philosophical and historical analysis. The recent history of structural biology and protein science offers untapped potential for history and philosophy of science. My ultimate goal for this dissertation therefore is to identify and analyze some of the key historical and philosophical puzzles that arise in these fields. I focus primarily on the shift from the static to the dynamic view of proteins in the late twentieth century. The static view treated proteins as stable, rigid structures, whereas the dynamic view considers proteins to be dynamic molecules in constant motion. In the first half of the dissertation, I develop a historical account of the origins of the static view of proteins. I show how this view led molecular biologists to adopt mechanistic explanation as their preferred strategy for explaining protein function. I then develop an account of the emergence of the dynamic view of proteins, arguing that thermodynamic theory and the theoretical commitments of scientists played an important and often overlooked role in driving this change. In the second half of the dissertation, I analyze the epistemological relationship between the static and dynamic concepts of the protein and argue that conceptual replacement is occurring. I then develop an account of ensemble explanation, a new type of explanation introduced to highlight the role of dynamics in protein function. I show that these explanations fail to fit existing philosophical accounts of explanation, ultimately concluding that my account is required to capture their epistemic structure.

Table of Contents

List of Tables

List of Figures

Preface

While only my name appears on the cover of this dissertation, many colleagues, friends, and family have contributed to its successful completion. I have been extremely lucky to receive guidance and support from many within the Pitt HPS family during my doctoral studies and while drafting this dissertation. I could not have finished it without the help and encouragement of Mike and Jim, my co-advisors. Mike has been my cheerleader throughout this process, and Jim has consistently offered speedy and thoughtful comments on my work. I am grateful for the intellectual guidance of my other two committee members, Sandy and Anya. This dissertation, and my scholarship, are better for their input.

The Pitt HPS graduate community has fostered my growth as a scholar and as a person. I count my cohort-mates Siska de Baerdemaeker and Zina Ward among my closest friends and am grateful that Pitt HPS brought us together. Their friendship and support mean the world to me, and I cannot imagine a world in which I successfully completed this journey without them. I am grateful to the entire Pitt HPS grad community, especially Gal Ben Porath and Mahi Hardalupas. During the drafting of the dissertation, I have had the pleasure of co-working either in-person or remotely with many friends, colleagues, and even family, including Michael Begun, Yudi Feng, J.P. Gamboa, Gus Law, Siddharth Muthu Krishnan, Carrie Neal, Elizabeth Oestreich, and Ellen Smith. I am thankful for their encouragement and companionship.

I owe a tremendous debt of gratitude to the friends and family who have supported and nurtured me during my long and winding academic journey, from a bachelors in biochemistry, through multiple masters in science and philosophy, until now, with the completion of my doctorate. Special thanks are due to Katie Campbell-Proszowska, Allison DuVal, Rachel Elaine

Cronmiller, and Franzi Seeger, who joined me at various points along the way and stood by me through the good times and the bad. I consider myself incredibly lucky to count two of my undergraduate professors, Jane Geaney and John Pagan, as friends and mentors for the past 15 years. In addition, my partner, Kevin Vergara, has supported me through the rather grueling process of drafting this dissertation, helping me stay sane and grounded.

I am forever grateful for my family for always being there for me. I am saddened that my dad, Peter Neal, did not live to see me finish this PhD, but I know he would have been proud of me. I am happy to share credit for my success with him, my mom, Kathleen Poll, and my great big sister, Rebekah Neal Hensley, as well as my sister's family, Greg, Suzanne, and Oscar Hensley. I do not know if it takes a village, but in my case, it certainly took a loving and supportive family to help me achieve this dream.

1.0 Introduction

We start from the premise that proteins are at the heart of all living processes, uniquely versatile in their capability to do whatever is needed, and with the knowledge that proteins are increasingly being harnessed to serve practical needs of society, to cure disease, to safeguard crops, etc. As a result, proteins are now moving to centre stage in the theatres of medicine and biology.

Charles Tanford and Jacqueline Reynolds (2003), *Nature's Robots: A History of Proteins*, p. 6

Since the middle of the twentieth century, our understanding and perceptions of life, health, and disease have become thoroughly molecular. During this time, the funding for molecular biology and biomedical research has soared, and the US and other governments have funded "big science" projects in molecular biology, such as the human genome project (Green et al. 2015). The faith in the molecularization of biology has always rested on the belief that molecular studies will reveal the underlying structures that control diverse biological processes from heredity to metabolism. By discovering the molecular structures and their clockwork mechanisms, researchers would bring order to the seemingly complex and disorderly biological world. In his Nobel lecture in 1965, Jacques Monod voiced this belief: "The ambition of molecular biology," he claimed, "is to interpret the essential properties of organisms in terms of molecular structures" (207). But in the late twentieth century, there was dramatic shift in molecular worldview from static to dynamic. It was no longer enough to characterize molecular structures in order to understand biological functions. One also needed to understand the dynamics of biological molecules. Nowhere did this transition from the static to dynamic worldview play out more visibly than in structural biology with the study of proteins. Over the last half century, scientists have modified their concept of the protein to include both structural and dynamic properties. They have developed new models to

better represent protein dynamics, and they have replaced explanations that invoked static structures obtained from x-ray crystallography with explanations that capture the dynamic fluctuations within these structures in living organisms.

As the epigraph from eminent scientists Charles Tanford and Jacqueline Reynolds suggests, proteins are becoming more and more important as an object of scientific inquiry. Within the molecular life sciences, many disciplines, including molecular and structural biology, biochemistry, enzymology, and molecular genetics, count proteins as among their objects of study. Proteins figure in explanations of almost all biological processes, from DNA replication and repair to cellular metabolism to innate immunity, as well as explanations of diseases such as sickle cell anemia, HIV, and COVID-19. Proteins also serve as sites of intervention in both laboratory techniques and medical treatments. Over the past century, the importance of proteins in these and other biological processes has become clearer, and many changes have occurred as scientists have shifted from thinking of proteins as rigid and static structures to dynamic molecules always in motion (Motlagh et al. 2014; Morange 2017, 2020). This shift from static to dynamic has led to changes in the methods scientists use to investigate proteins, the ways they conceptualize and represent them, as well as the explanations scientists give of their behaviors and functions.

This recent history of protein science is ripe for historical and philosophical analysis. Most philosophical work in molecular biology has historically centered on DNA, genetics, and questions of reduction. Early work focused on genetics, debating the possibility of intertheoretic reduction, while more recently some philosophers of molecular biology have discussed explanatory reduction and developed accounts of mechanistic explanation. This dissertation breaks from this tradition to make proteins the object of philosophical analysis. It focuses specifically on the shift from the static to dynamic view of proteins in the late twentieth century. In the dissertation, I examine the

historical reasons for this shift. I also address epistemological questions that arise in the context of representing proteins, rather than genes, and I characterize a novel type of non-mechanistic explanation that scientists have developed to explain dynamic protein behavior.

My dissertation therefore advances the scholarly literature in important ways. It develops a historical account of the emergence of the dynamic view of proteins, which has been largely overlooked by historians and philosophers of biology. Although the origins of molecular biology have been the focus of much historical research, little has been done on this recent history in protein science (Morange 2018, Sarkar 2008). It also contributes to the literature on scientific explanation, introducing and characterizing a novel type of explanation in the biological sciences. Finally, my analysis of protein concepts, representations, and explanations has important implications for philosophical debates about unification in the biological sciences.

Scientific explanation has been a central topic in philosophy of science, at least since Hempel's characterization of the deductive-nomological model (Hempel 1965, Hempel and Oppenheim 1948). Recently, philosophers interested in the biological sciences have offered new accounts of explanation that aim to capture the explanatory practices of molecular biologists, geneticists, neuroscientists and others working in related fields. These philosophers have argued that mechanistic explanations, which explain the behavior of a system by describing the organization, activities, and causal roles of its component parts, dominate in these fields (Machamer et al. 2000, Bechtel and Abrahamsen 2005, Craver 2007). Although mechanistic explanation certainly plays an important role in the explanatory practices of the life scientists, my account of ensemble explanations of protein behavior demonstrates that mechanistic explanations do not have a monopoly on explanatory practices in protein science. In fact, the recognition of the importance of protein dynamics for protein function has caused a major shift in explanatory

practices in the field because mechanistic explanations, which are modeled after machines, cannot easily accommodate the fast timescale dynamics that cause proteins to undergo constant structural fluctuations. My account of ensemble explanation of protein behavior therefore undermines the claims that mechanistic explanation is the only or dominant type of explanation in the molecular life sciences. It shows that scientists have had to develop a new type of explanation—ensemble explanation—to account for the dynamic properties of proteins.

Throughout the dissertation, I focus primarily on one protein function, namely, allostery, or the phenomenon whereby binding at one site on protein affects binding at a second site (Cui and Karplus 2008). I analyze older mechanistic explanations as well as newer ensemble explanations of allostery, which aim to account for the role of dynamics in enabling proteins to carry out this function. Interest in ensemble explanations, however, extends far beyond the community of biologists studying allostery. The dynamic behavior that these explanations aim to capture is common; for this reason, ensemble explanations are useful for explaining other protein functions as well as the behavior of other biomolecules. Indeed, historian and philosopher of biology, Michel Morange (2017) has expressed excitement that explanations of this sort have the potential to provide unified explanations of many phenomena, for example, explaining protein folding, catalysis, and allostery in one fell swoop.

This dissertation is comprised of four substantive chapters, and although they all share a common theme, each chapter is self-contained. In the first two chapters, I develop a history of proteins in the twentieth century, as studied by molecular and structural biology, focusing on the shift from the static to dynamic view of proteins. In the final two chapters, I develop a philosophical analysis of this shift. I first analyze the epistemological relationships between the

static and dynamic view of proteins and then develop a novel philosophical account of explanation associated with the dynamic view.

In Chapter 2 ("Proteins in Classical Molecular Biology: The Static View, Protein Structure, and Mechanistic Explanations of Allostery), I describe the origins of the static view of proteins in the early twentieth century, which held that proteins were compact, rigid, and static molecules. Because of this view, protein dynamics were largely ignored, since they were not thought to play a role in protein function. In the first part of the chapter, I consider the development of two accounts of protein function in the 1950s and 1960s that seemed to break from the static view: Daniel Koshland's induced-fit model of catalysis and Jacques Monod's model of allostery. I analyze the role of protein dynamics in these two accounts of protein function and assess their compatibility with the static view. I argue that neither account of protein function represented a serious challenge to the static view of proteins. In the second part of the chapter, I focus my analysis on the development of the Monod-Wyman-Changeux (MWC) model of allostery, since this episode played a central role in the birth of structural biology. While some recent scientific commentators have claimed that the MWC model is not explanatory, I argue that it fits philosophical accounts of mechanistic explanation. Although it is an abstract, how-plausibly model, it nevertheless reveals the underlying molecular mechanism that explains the allosteric effect.

In Chapter 3 ("From Static to Dynamic: A Historical Account of the Emergence of the Dynamic View of Proteins"), I shift my attention to the dynamic view of proteins. Here, I develop a historical account of the origin and spread of the dynamic view that answers two historical puzzles: (1) why did the dynamic view emerge in the 1970s and 1980s as an alternative to the dominant static view? and (2) what explains the long time lag between the origin of the view and its eventual acceptance? The answer to both questions, I contend, reveals the importance of theory

in this history. I argue that the theoretical understanding of protein dynamics was the primary driver behind the emergence of this new view of proteins. The application of thermodynamic principles to proteins led to the development of the dynamic view, and commitment to these principles led a small group of scientists to seek out anomalous cases of protein dynamics that could not be explained by the static view. These researchers slowly accumulated empirical evidence for the dynamic view, and their empirical findings eventually convinced the majority of protein scientists. Explaining the uptake of the dynamic view therefore involves the discovery of anomalies aided by new technologies, but the commitment to treating proteins as small thermodynamic systems ultimately drove that search for anomalies. For this reason, I conclude that the emergence and spread of the dynamic view of proteins view was, at its core, theory-driven.

In Chapter 4 ("Protein Concepts, Representations, and their Epistemic Relations"), I turn my attention from historical analysis to consider the epistemic relations between protein concepts. Specifically, I characterize the epistemic relationships that obtain between the structural and dynamic protein concepts and their associated representations. A major innovation of my philosophical approach to this topic is the identification of two levels of analysis. I argue that we must distinguish the epistemic relationship that obtains between the two protein concepts from the epistemic relationships that obtain among their associated representations. From my analysis of the concepts, I argue that conceptual replacement, rather than reduction or integration, is occurring in protein science. The conceptual core of the structural concept includes beliefs about proteins that privilege structure over dynamics that are rejected by the dynamic concept. The dynamic concept replaces those beliefs with ones that recognize structure and dynamics as co-determinants of protein function. On my view, this revision in the inferential role of the concept is too drastic to count as reduction and instead constitutes conceptual replacement. At the level of representations,

however, I argue that dynamic and structural representations of proteins are typically related via abstraction. In other words, a structural representation can be obtained by abstracting away dynamic information from a representation that includes both dynamic and structural properties of proteins. Because protein dynamics and structure are largely partitionable subfeatures of proteins, the inclusion or exclusion of one of them in a representation has little effect on the other. Therefore, scientists can often transition between structural and dynamic representations of proteins via processes of abstraction or de-idealization.

In Chapter 5 ("Ensemble Explanation of Protein Function: A Philosophical Account of a Novel Type of Explanation in the Biological Sciences"), I develop an account of ensemble explanation of protein behavior. These explanations deserve philosophical attention, since they are becoming increasingly common in structural and molecular biology as scientists recognize the importance of dynamic properties of proteins. To capture protein dynamics, these explanations represent a population of protein molecules as an ensemble of structurally distinct microstates. They cite changes in the equilibrium distribution of protein molecules across microstates, caused by perturbations, to explain protein functions. After developing my account, I compare ensemble explanations to two different philosophical accounts of explanation: equilibrium explanation and causal explanation in thermodynamics. I argue that ensemble explanations do not fit the model of equilibrium explanations because the equilibria in ensemble explanations do not play the right explanatory role. Instead, I argue that ensemble explanations are a species of causal explanation, similar to but distinct from causal explanations in thermodynamics. They are, moreover, distinct from mechanistic explanations of protein behavior. My account shows that scientists have not simply modified mechanistic explanations to include protein dynamics. Instead, they have

developed ensemble explanations as a novel type of explanation to deal with the added complexity of protein dynamics.

Ultimately, this project aims to contribute to the still nascent history and philosophy of the molecular life sciences. In my view, the molecular life sciences—biophysics, biochemistry, and structural and molecular biology—offer untapped potential for history and philosophy of science. By examining historical, epistemological, and methodological questions that arise in these sciences and relating them to debates in general philosophy of science, my dissertation will hopefully be of interest to historians and philosophers of science, as well as those with more specialized interest in structural and molecular biology.

2.0 Proteins in Classical Molecular Biology: The Static View, Protein Structure, and Mechanistic Explanations of Allostery

2.1 Introduction

Because of their role in carrying out most biological functions, proteins have been an important object of scientific inquiry since the late nineteenth century (Fruton 1999). Research in the twentieth century quickly gave rise to the static view of proteins. On this view, globular proteins, such as enzymes, were thought to be compact, rigid, and static molecules. When coupled with the structure-function rule, which held that protein structure determined protein function, the static view offered a heuristic for explaining the behavior and function of proteins (Sarkar 2008). In the 1970s and 1980s, a new dynamic view of proteins gradually emerged as a challenge to the prevailing static view. It held that proteins were dynamic molecules undergoing constant structural fluctuations. The second chapter of this dissertation recounts the emergence of this new view and its slow adoption by molecular and structural biologists. In this chapter, however, I begin by attending to the previous period—the late 1950s through the 1960s—when the static view of proteins remained largely unchallenged. This was the heyday of classical molecular biology, what one commentator has called "the decade of the rigid macromolecule" (Phillips 1981, 497).

During the reign of the static view, however, not all molecular biologists were committed to the view that macromolecules were entirely rigid; they developed two new accounts of protein function that posited a limited role for protein dynamics. First, in 1959, Daniel Koshland presented his induced-fit model of enzyme catalysis, which held that an enzyme changed conformation upon binding its substrate. Then, in the early 1960s, Jacques Monod and his colleagues at the Pasteur

Institute developed a model of allostery, which posited that allosteric proteins could adopt one of two conformations. My first objective in this chapter is to provide a historical account of the development of these two cases. I analyze the role of protein dynamics in these two accounts of protein function and assess their compatibility with the static view of proteins. Ultimately, I argue that although they did posit a limited role for protein dynamics in protein function, neither account of protein function represented a serious challenge to the static view of proteins.

My second objective leads me to focus more closely on Monod's discovery and characterization of allostery, or what he called "the second secret of life" (Ullman 1979, 167). Important for cellular and metabolic regulation, allostery is the phenomenon whereby binding at one site on a protein affects binding at a distant site. Through a series of important papers in the early 1960s, Monod and his colleagues established a structural framework for explaining allostery that set the research agenda for the field until the turn of the century.[1] But, despite the foundational role of allostery within structural biology, its history has been largely overlooked by historians of science. Angela Creager and Jean-Paul Gaudilliere's (1996) account of the early evolution of the concept is one noteworthy exception. The recent 50-year anniversary of the discovery of allostery and the development of influential Monod-Wyman-Changeux (MWC) model has produced a flurry of scientific retrospectives recounting the history of allostery.[2] Using these new sources, as well as previously published papers and unpublished archival sources, I revisit this history, tracing the development of the concept of allostery from 1961 through the publication of the MWC model in 1965. In addition to fleshing out earlier historical accounts, my main aim here is more

[1] See Monod and Jacob (1961), Monod et al. (1963), Monod et al. (1965).
[2] For recent scientific retrospectives of the history of allostery, see Changeux (2011), (2012), and (2013) and Gerhart (2014) as well as the special issue of the *Journal of Molecular Biology* devoted to the 50-year history of this topic (Edelstein 2013).

philosophical. Recent scientific commentators have claimed that the MWC model was merely phenomenological and did not explain allostery. I disagree. I analyze the model and argue that it fits philosophical accounts of mechanistic explanation. On my view, the MWC model is an abstract model as well as a how-plausibly model, since many of its structural posits were not yet confirmed in 1965, but it is nevertheless a mechanistic explanation of allostery.

In what follows, I first describe the development of the static view of proteins in the early twentieth century (Sec. 2.2). I then discuss Koshland's induced-fit model of catalysis (Sec. 2.3) and Monod's account of allostery (Sec. 2.4) within the context of the static view of proteins, arguing that both accounts can be accommodated within the static view. I next turn my attention more squarely to the MWC model of allostery (Sec. 2.5). I present and analyze the model and then argue that it should be construed as a mechanistic explanation of allostery.

2.2 The Static View of Proteins

For the first half of the twentieth century, the dominant view of protein structure held that globular proteins were rigid, compact, and largely static molecules. Experimental evidence and theoretical support for this characterization of proteins came from the work of organic chemists and protein scientists since the end of the nineteenth century. By 1900, Gerrit Mulder and other chemists had established that proteins were primarily composed of amino acids (Fruton 1999, Tanford and Reynolds 2001). Concurrent work by Emil Fischer and Franz Hofmeister in Germany suggested that the amino acids in proteins were joined together into long chains by a single bond type, named the "peptide" bond by Fischer (Fischer and Fourneau 1902, Fischer 1906, Hofmeister 1902). Although there were some early holdouts, the theory of the peptide bond and the linear

structure of amino acids would become textbook science by 1920. While it was a significant advance in understanding protein structure, the peptide bond did not explain how these long chains of amino acids were folded into 3D structures. Experimental results soon indicated the diameters of globular proteins in solution were much smaller than the linear amino acid chains of which they were comprised. This finding, along with other results measuring the electric charge on proteins, led many researchers in the 1920s and 1930s to adopt a view of proteins as compact, spherical molecules, which carried a high density of surface charges that made them water soluble (Tanford and Reynolds 2001).

Early work on protein crystallography, beginning in the mid-1930s, lent additional support to this static view of proteins. In 1934, John Bernal and Dorothy Crowfoot Hodgkin published the first x-ray crystal structure of a globular protein—pepsin (Bernal and Crowfoot 1934). The crystal structure revealed that pepsin had an ordered structure. Bernal later described those first crystal structures as being of "exceptional perfection" (Bernal 1939, 663). He claimed that the regularity of the crystals "indicated that not only were the molecules of the proteins substantially identical in shape and size, but also that they had identical and regular internal structures right down to the atomic dimensions" (Ibid.). In other words, this early crystallographic evidence was taken to show both that protein molecules had rigid structures, with each molecule fixed in place, and that all protein molecules of the same type had the same 3D structure.

The crystallographic work of John Kendrew on myoglobin and Max Perutz on hemoglobin in the 1950s and 1960s further bolstered the rigid view of proteins. Their studies showed that these proteins were composed of densely packed polypeptide chains (Green et al. 1954, Kendrew et al. 1958, Perutz et al. 1960). In her contribution on protein structure to a four-volume compendium entitled *The Proteins* (1953), which aimed to review the current knowledge of the topic, Barbara

Low highlighted the structural findings from crystallography that supported the static, structural view of proteins: x-ray diffraction patterns indicate that many proteins have high electron density rods in a "close-packed array," which "probably correspond to coiled peptide chains in the molecular structure" (Low 1953, 240). However, she acknowledged the limitations of these findings and the need to obtain more and better high-resolution protein structures:

> At best, a "bird's eye" long-distance view of some protein molecules has been derived. It is, however, far from a detailed or precise description of the molecular architecture from which may be identified the sequence of amino acids along the chains, the nature of the intrachain coils or folds, interchain packing, and finally, the 3N Cartesian coordinates of the equilibrium positions in space of the N atoms in the molecule. (Ibid.)

As this passage from *The Proteins* (1953) shows, structural biologists recognized the promise of x-ray crystallography for the study of proteins by the 1950s. Although the available structures at that time provided only coarse-grained information about the interior structure of globular proteins, the evidence they did provide corroborated the static view of proteins. Moreover, crystallographers and protein scientists were confident that improvements in x-ray crystallography would eventually reveal the complex molecular architecture that enabled proteins to maintain stable 3D structures.

Contemporaneous theoretical work on protein structure reveals that scientists working on protein structure from a theoretical perspective were also committed to the static view of proteins. Perhaps the most famous of those working on the problem of proteins structure in the 1950s was Linus Pauling. Using approximations of interatomic distances and bond angles obtained from x-ray crystal structures, Pauling and colleagues developed a plausible model for the secondary structure of amino acids in the interior of a globular proteins: the α-helix (Pauling et al. 1951). To build the model, they adopted a principle of maximum hydrogen bonding. Maximizing the number of hydrogen bonds between amino acids in the chain restricts the rotational degrees of freedom available to a polypeptide chain without any additional structure. The structure that maximized the

number of these internal bonds would consequently be the most rigid and stable helix possible. Thus, in theoretical work of this sort, the rigidity of globular proteins was an assumption the modelers built into their models.

This experimental and theoretical work on protein structure was complemented by earlier work on protein function that also endorsed a rigid and static view of proteins. The dominant theory of enzyme catalysis, from the 1890s until the 1950s, was the lock-and-key, or template, model of catalysis. Introduced by Fischer in 1890 to explain the specificity of enzymes for their substrates, this model of catalysis hypothesized a rigid fit between an enzyme (lock) and its specific substrate (key) (Fischer 1890, 1894). Daniel Koshland offered a description of Fischer's lock-and-key model in 1958: "In essence this theory said that the enzyme was a rather rigid negative of the substrate and that the substrate had to fit into this negative to react" (1958, 99). According to Koshland, with minor modifications and refinements, this model was the accepted view through the mid-1950s, since it offered "the best framework for explaining" the phenomenon of enzyme-substrate specificity (Ibid., 99). On this view, enzymes can carry out their catalytic function only if they maintain a fixed 3D structure that complements their substrates.

From the end of the nineteenth century up until the 1950s, research on protein structure and function converged on a static view of proteins. With the acceptance of this static, structural view of proteins, protein scientists in the mid-twentieth century naturally expected the emerging technology of x-ray crystallography, which offered detailed but static representations of proteins, to be the key to elucidating protein structure and answering questions about protein function.

2.3 A Functional Role for Protein Dynamics: Koshland's Induced-Fit Model of Catalysis

The account of protein structure generally accepted by the 1950s did not recognize a functional role for protein flexibility or dynamics. That changed when Daniel Koshland introduced his induced-fit model of enzyme catalysis at the "Symposium on Amino Acid Activation" at the meeting of the National Academy of Science in November 1957 (Koshland 1958). Presented as a correction to Fischer's lock-and-key model of protein catalysis, the induced-fit model proposed that substrate binding causes an enzyme to change conformation. According to the model, it is this induced change in enzyme conformation that brings the catalytic groups on the enzyme into the proper orientation with the substrate in order to catalyze the reaction. Koshland's induced-fit model was a major advance in the history of protein science and the understanding of protein structure. Proteins, which were previously viewed as wholly rigid and static, were for the first time thought to undergo conformational changes essential to the performance of their function. Nevertheless, I argue that the protein flexibility proposed by Koshland's model did not significantly challenge the static view of proteins; instead, the model proposed a modification that was easily incorporated by it.

The induced-fit model of enzyme catalysis was designed to overcome certain explanatory problems that beset the lock-and-key model (Koshland 1958, 1959, 1963). According to the lock-and-key model, the enzyme contains "a relatively hard and inflexible active site" that directly complements the stereospecific structure of the substrate (Koshland 1959, 245). While this explains why larger compounds analogous to the substrate show no catalytic activity (i.e., they are "keys" too large to fit into the "lock" formed by the active site), it fails to explain why smaller compounds also fail to react. That is, the lock-and-key model offers no explanation for why smaller substrate analogues with the same reactive sites as the substrate would not react. These molecules

are small enough to fit into the active site cavity, yet kinetic studies had shown that they also fail to react (Koshland 1958, 1959). Koshland introduced his induced-fit theory to account for this puzzling experimental observation.

His theory had three main postulates: (1) catalysis requires "precise orientation" of the substrate and catalytic groups on the enzyme, (2) substrate binding causes significant changes "in the three-dimensional relationship of amino acids at the active site," and (3) these substrate-induced changes in protein structure "bring the catalytic groups into the proper orientation for reaction," while the binding of non-substrate molecules will not (Koshland 1958, 100). According to the induced-fit model, only the substrate or substrate analogues of similar size will be able to react with the enzyme. A substrate analogue that is larger or smaller than the actual substrate will fail to establish the necessary linkages with the active site on the enzyme and will consequently fail to react.

This model of catalysis rejected the static picture of the enzyme of the lock-and-key model. According to Koshland, the results of enzyme modification studies suggest that it is unlikely that "a rigid 'positive' substrate fits a rigid 'negative' template" (Koshland 1963, 1540). To replace the lock-and-key metaphor, Koshland introduced a new metaphor for the induced-fit model: "the substrate induces a structural change in the molecule, as a hand changes the shape of a glove" (1963, 1540). As the glove metaphor illustrates, Koshland's theory retained the notion of fit, or structural complementarity, between the enzyme and substrate, but it emphasized that the fit is obtained "*only after* the changes induced by the substrate itself" (Koshland 1958, 100, emphasis in original).

In recounting the history of protein dynamics, the importance of the induced-fit model cannot be overstated. For the first time, proteins were recognized as flexible molecules that

undergo structural changes in order to perform their functions. Nevertheless, in the initial presentations of the theory, Koshland (1958, 1959) downplayed its novelty, arguing that earlier studies had already demonstrated that proteins were flexible.[3] Citing studies showing the reversible denaturation of certain enzymes when exposed to urea, Koshland argued that they provided evidence for a structural change in proteins, since the denatured proteins exhibited changes from their native counterparts in their optical rotation, sedimentation rates, and other properties that depended on the 3D structure of the protein (Koshland 1958). While these studies do indeed show that enzyme structure can change under certain conditions (e.g., transitioning from a folded active conformation to an unfolded, denatured conformation and back again), they do not show that this flexibility is a part of normal protein function. Urea denaturation is a laboratory technique involved in protein purification, and the action of an enzyme under these conditions has little bearing on normal enzyme function (Bonner 2019). The novelty of the induced-fit model is that the enzyme flexibility and conformational change it postulates are essential to the normal functioning of enzymes.

It would be a mistake, however, to think that Koshland's induced-fit theory represented a significant challenge to the prevailing static of protein structure. Rather, the induced-fit model was a friendly amendment that was easily accommodated within the static, structural view of proteins. Although Koshland's theory introduced dynamics and flexibility into a molecule previously considered static and rigid, that motion was carefully circumscribed in a way that allowed it to fit

[3] In a much later commemorative lecture, Koshland provides his rationale for downplaying the novelty of the induced-fit theory with its postulation of functionally relevant protein motions: "The theory of Emil Fischer was deep in the hearts of scientists and journal editors, so I had great difficulty getting the original ideas published or convincing skeptics, but we did obtain more evidence from my own laboratory, and soon others joined in" (Koshland 1994, 2378). Koshland's memory of publishing his theory reinforces my claim that the static view of proteins, which acknowledged no place for enzyme flexibility or dynamics, was still the prevailing view in the late 1950s.

within the existing static framework. The dynamics that occurred upon substrate binding merely enabled the enzyme to switch between two structures—one conformation for the enzyme-substrate complex and one for the free enzyme. These structures were hypothesized to be discrete and fully determined by the binding context: "the protein changes shape under the influence of the substrate and returns to its original shape after the products have been released from the enzyme surface" (Koshland 1959, 247). Hence, each of these two possible enzyme conformations—bound and free—were assumed to have structures that were stable and rigid.

Koshland himself highlights the conformity of this limited kind of protein dynamics with the static view of proteins. In his discussion of protein structure-function relationships, he approvingly cites Kendrew's crystallographic work on myoglobin, which he claims showed that the "folding of linear arrays of amino acids" is "precisely defined," forming a "unique three-dimensional" structure for this protein (Koshland 1963, 1533). Moreover, he is hopeful that the "'moment of truth' is not far off" when crystallographers will be able to determine the 3D structure of enzymes, in both their bound and free forms, in much the same way (Ibid. 1540). If Koshland saw the protein flexibility postulated by his theory as a refutation of the static view of proteins, he would certainly not cite these structural studies so enthusiastically. Although he considered the induced-fit model of catalysis as a replacement for Fischer's lock-and-key model, he also saw his account of limited enzyme flexibility as compatible with the prevailing static view of protein structure. Whereas before 1958 every protein had one rigid conformation, after the introduction of the induced-fit model, some proteins—viz., enzymes—would be assumed to have two conformations, each of which would have its own precise, well-defined, and stable structure.

Although Koshland's work in the late 1950s and early 1960s highlights the "dynamic interaction" between the enzyme and substrate and the "flexible nature" of the enzyme (Koshland

1963, 1540; Koshland 1958, 101), the protein flexibility postulated by the induced-fit theory is strikingly different from the dynamic motions relevant to the dynamic view of proteins that first appeared in the late 1970s. The flexibility Koshland describes in his 1959 paper entitled "Enzyme Flexibility and Enzyme Action" did not arise from the inherent "jigglings and wigglings of atoms" (Feynman et al. 1963). Instead, the enzyme is flexible because it responds to an exogenous substance—viz., the substrate—in order to adopt a different conformation. The two structures the enzyme adopts are internally stable; they are not hypothesized to undergo any motion except when acted upon by the substrate. Koshland's primary goal is to distinguish his theory from Fischer's lock-and-key theory, rather than the static view of protein structure. Hence, he claims that the conformational changes postulated by his induced-fit model should typically be "big" and "significant" structural changes (Koshland 1994, 2377). Moreover, he explicitly argues that the induced-fit model does not apply to the sort of protein dynamics that would arise from random thermal fluctuations that cause the protein to deviate from its average structure:

> The important feature from the induced fit theory is that the alignment of catalytic groups and binding groups must be optimized for the transition state, and the attainment of the state is unfavorable energetically unless it is supplied with the energy of the substrate binding. If the protein movements were easy to attain, they would occur spontaneously often enough to have little effect on catalysis. (Ibid., 2378)

Even in 1994, when accounts of protein dynamics were already beginning to appear, Koshland confirmed that the induced-fit model aimed to describe relatively large translational motions that enable an enzyme to switch between two discrete structural conformations and not the small, fast timescale motions that are at the heart of dynamic accounts of proteins. In sum, Koshland's work on enzyme flexibility in the late 1950s and early 1960s represented an important advance because it was the first to argue that protein motion could be relevant for protein function. However, it did not pose a significant challenge to the prevailing view of protein structure.

2.4 Another Role for Protein Dynamics: Monod's Discovery of Structural Allostery

Around the time that Koshland was advocating a new role for protein dynamics in catalysis, another set of scientists studying enzymes were considering a potential role for dynamics in a different protein function, namely, allostery. Allostery, or the phenomenon whereby binding at one site on a protein affects binding at a distant site, was discovered and characterized by Monod and his colleagues at the Pasteur Institute in the early 1960s. Building upon previous work on feedback inhibition, Monod introduced the term "allosteric inhibition" in 1961 to generalize the phenomenon, focusing on features of protein structure that enable a ligand that is structurally dissimilar from the substrate to bind an enzyme in order to regulate its function.

In this section, I draw upon archival and published sources, scientific retrospectives, and the historical analysis of Angela Creager and Jean-Paul Gaudilliere (1996) to document this early history of allostery. I focus primarily on the role Monod and his colleagues at the Pasteur Institute played in characterizing and explaining allostery from 1961 until the publication of the MWC model in 1965. My periodization largely tracks their three major publications during this time. However, I begin by describing the biological research in the 1950s on bacterial feedback, or end-product inhibition, since Monod drew heavily upon this earlier research (Sec. 2.4.1). I then analyze Monod's concluding remarks to the 1961 Cold Spring Harbor meeting in which he first introduced and characterized the concept of "allosteric inhibition." I argue that the concept shifted attention from the biological function of feedback to the physical structure of enzymes (Sec. 2.4.2). I then show how Monod and his colleagues at the Pasteur Institute refined the concept in a review article, published in the *Journal of Molecular Biology* in 1963. The review sketched a molecular mechanism that posited changes in protein structure to explain allostery (Sec. 2.4.3). Because of

its importance in this history, I reserve my discussion of the MWC model itself for the following section (Sec. 2.5).

2.4.1 Feedback Inhibition

Monod coined the term "allosteric" in his concluding remarks to the Cold Spring Harbor meeting in 1961, but the history of allostery actually began the decade before. It originated with the discovery of "feedback inhibition" in the laboratories of two American researchers investigating cellular regulation in bacteria. H. Edwin Umbarger at Harvard and Arthur Pardee at Berkeley both independently discovered the phenomenon of feedback inhibition in the mid-1950s (Umbarger 1956; Yates and Pardee 1956b). They found that the enzyme at the beginning of a biosynthetic pathway could be inhibited by the end product of the pathway. Umbarger, for example, showed that the enzyme threonine deaminase, the first step in the biosynthesis of the amino acid isoleucine, was inhibited by isoleucine, its end-product.

The discovery of feedback inhibition occurred as these researchers were attempting to map out different metabolic and biosynthetic pathways in *E. coli*. Although Umbarger studied amino acid synthesis and Pardee investigated nucleic acid synthesis, both researchers used similar experimental methods, which were becoming common in the burgeoning fields of bacterial genetics and cell regulation (Dixon and Webb 1964). Researchers in these fields induced mutations into *E. coli* and isolated mutants that required certain metabolites in order to survive (e.g., particular amino or nucleic acid intermediates). They hypothesized that the mutant strains were lacking an enzyme in a particular biosynthetic pathway. The absence of a functional enzyme would lead to the accumulation of the metabolites from the preceding steps in the pathway, but the bacteria could overcome this "enzymatic blockade" when the metabolite normally synthesized by

the defective enzyme was added to the cell culture. Researchers were thus able to identify the intermediate steps in various biosynthetic pathways by identifying the metabolites that accumulated. While conducting research of this sort, Umbarger and Pardee confirmed earlier work indicating that particular amino acid and nucleic acid biosynthetic pathways were inhibited by the presence of the end-product of the pathway. What made their discoveries noteworthy, however, was the fact that they identified the first enzymatic step in the respective pathways as the point of inhibition and then sought to characterize the mechanism of inhibition of these enzymes under negative feedback control.[4] To accomplish this aim, they turned to the methods of biochemistry and enzyme kinetics.

By the 1950s, kinetic assays were a standard tool for the study of enzymes. Reflecting on his time as Monod's graduate student in the late 1950s and early 1960s, Jean-Pierre Changeux recalls how Malcolm Dixon and Edwin Webb's *Enzymes* quickly became "the bedside book of biochemistry students" such as himself (Changuex 1993, 625). This textbook summarized the great progress on enzymes and enzyme kinetics in the first half of the twentieth century. During that time, work in enzyme kinetics focused on the formalization of reaction kinetics, identifying how to mathematically describe and experimentally quantify the process of enzyme-substrate formation and transformation into products.

The simplest kinetic assays, both in the 1950s and still today, measure either the appearance of product or the disappearance of substrate over time. From this data, researchers calculate certain catalytic properties of enzymes, such as their maximum velocity (V_{max}). Typically, the values of

[4] When Novick and Szilard (1954) showed that the synthesis of a tryptophan precursor was inhibited by tryptophan, they hypothesized that the end-product was inhibiting an enzyme early in the pathway. Umbarger, however, was the first to work "directly at the enzyme-level" in order to show that the first enzyme in the pathway was "strongly and specifically inhibited" by the metabolic end-product (Monod and Jacob 1961, 390-1).

these properties are estimated from Michaelis-Menten plots. Named after the pioneering work of Leonor Michaelis and Maud Menten (1913), the Michaelis-Menten plot is a substrate saturation curve that depicts the velocity of the reaction at different substrate concentrations with fixed enzyme concentration. Standard Michaelis-Menten plots take the form of a rectangular hyperbola (Fig. 1).

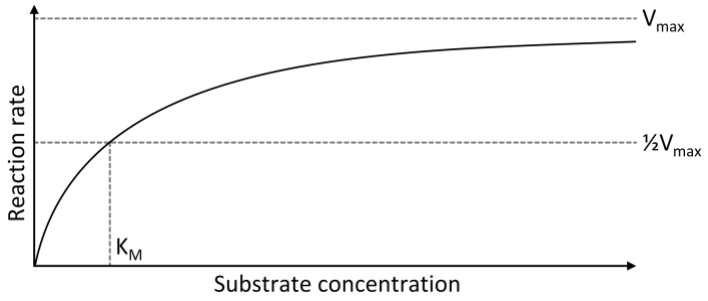

Figure 1 Idealized Michaelis-Menton Plot. Enzyme activity plot, showing dependence of enzyme velocity (V) on substrate concentration. Image by Thomas Shafee, distrubuted under a CC BY 4.0 license <https://creativecommons.org/licenses/by/4.0>, via Wikimedia Commons.

They enable the estimation of V_{max}, which is the asymptote of the hyperbolic function, as well as the Michaelis constant (K_m), which is the substrate concentration at half V_{max} (K_m describes the binding affinity of the substrate for the enzyme). Different enzymes, or even the same enzyme under different conditions, will have different values for V_{max} and K_m reflected in changes in the slope of the curve or the value of the asymptote. However, the rectangular hyperbola of the Michaelis-Menten plot was thought to be a largely invariant feature of enzyme kinetics that could accurately describe most enzymes under most conditions (Dixon and Webb 1964).

The presence or absence of inhibitors, for example, was a condition often interpreted within the Michaelis-Menten framework. Many biochemists and enzymologists in the early twentieth century were interested in understanding the effects of inhibitors on enzyme kinetics. At that time, competitive inhibition, in which the inhibitor competes with the substrate for binding at the active site, was the best understood type of inhibition. Researchers were able to identify and characterize

these inhibitors by comparing Michaelis-Menten plots conducted in the presence and absence of the inhibitor. Although the features of the saturation curves for an enzyme would differ depending on the presence and relative concentration of inhibitor, they would all be accurately represented by rectangular hyperbolas. In this case, the asymptote (V_{max}) would remain the same, but the slope of the curve would become less steep in the presence of the inhibitor. Inhibitors caused changes in the Michaelis-Menten plots that altered the plots in predictable ways but left the mathematical structure of the hyperbola intact.

When Umbarger and Pardee identified enzymes with unique inhibitory properties at the beginning of biosynthetic pathways, they were able to turn to these standard tools of enzyme kinetics to characterize their behavior. The earliest experiments in both labs used crude cell extracts rather than purified enzymes to measure the kinetic properties of these enzymes. Despite this limitation, Umbarger's and Pardee's laboratories were able to approximately quantify the inhibition of threonine deaminase and ATCase, respectively, by measuring the V_{max} and K_m of these enzymes in the presence and absence of their inhibitors. Based on the substrate saturation curves, which showed that the inhibitor did not change the maximum velocity but required higher substrate concentrations to reach it, they initially classified feedback inhibition as a type of competitive inhibition. In his one-page letter to *Science* in 1956 and again in his follow-up study with Barbara Brown in 1958, Umbarger noted that the inhibition of threonine deaminase did not follow standard Michaelis-Menten kinetics; in other words, the saturation curve in the presence of the inhibitor could not be described by a rectangular hyperbola.[5] In presenting and discussing the novelty of his discovery, however, Umbarger (1956) downplayed the heterodox kinetics. Instead,

[5] In a paper drafted in 1961, Gerhart and Pardee (1962) would also note the heterodox kinetic behavior of ATCase, but, as Umbarger did in his earlier work, they downplayed these findings.

he emphasized the fact that threonine deaminase was specifically and strongly inhibited by the end-product of the pathway, isoleucine, rather than by intermediate metabolites in the pathway or other amino acids.[6] Similarly, Pardee and Yates included a table showing that it was only the end-product and not other pyrimidine intermediates that inhibited ATCase.[7] In these early publications, both Umbarger and Pardee emphasize this peculiar feature of enzymes at the beginning of biosynthetic pathways and the biological importance of feedback inhibition.

Within this scientific context in the 1950s, the concept of feedback inhibition referred to a biological phenomenon that was an important regulatory property of certain biological systems. The concept emphasized the role of feedback inhibition in the regulation of the cell and the maintenance of normal functioning. Strictly speaking, feedback inhibition was not a property of a particular enzyme, but rather it was a feature of a metabolic regulatory system. Umbarger (1956), for instance, justifies studying feedback inhibition in *E. coli* "because of the complexity of so many biological systems" (848). In contrast, he claims *E. coli* offers the researcher "less complex systems for study of internal regulation" (Ibid.). Although they studied the kinetic properties of particular enzymes, Umbarger and Pardee used the concept of feedback inhibition to refer not to an enzyme in isolation, but instead to an enzyme embedded within a particular regulatory context—viz., a biosynthetic pathway.

The discoverers of feedback inhibition did not speculate on the molecular basis of inhibition. Their early work focused on the biological role of feedback inhibition. They did not study threonine deaminase and ATCase to determine their structural properties or any other features that might explain how their respective end-products inhibited these enzymes. The

---

[6] In fact, Umbarger (1956) includes only two sets of empirical data, and one is a table including exactly this information: "Table 1. Specificity of Inhibition of Threonine Deaminase."
[7] See Table I in Yates and Pardee (1956, 764).

concept of feedback inhibition in the 1950s, which was to become the precursor for Monod's new concept of "allosteric inhibition," was thus a thoroughly biological concept.

2.4.2 The Birth of Allosteric Inhibition

In the summer of 1961, researchers working on cellular and metabolic regulation gathered in Long Island, New York, for the annual Cold Spring Harbor Symposium on Quantitative Biology. The topic of "Cellular Regulatory Mechanisms" brought together two groups of scientists studying protein synthesis and regulation. The first group focused on elucidating the genetic basis for protein synthesis, while the second studied the biochemical regulation of proteins. Many in both groups focused their attention on the synthesis and regulation of enzymes, in particular. Among those studying enzyme regulation were scientists who analyzed the genetic control of enzyme activity as well as those who studied the biochemical regulation of enzymes. The contingent from the Pasteur Institute was well represented in both of these camps: Jacob and Monod presented their model of gene control of enzyme synthesis (i.e., the *lac* operon model), work that would win them the 1965 Nobel Prize, and Changeux, Monod's graduate student, presented his work on threonine deaminase in the section on feedback control. It is no wonder, then, that Arthur Chovnik and Umbarger, the main organizers of the conference, gave Monod and Jacob the prime speaking slot at the close of the program on July 12th.

In their concluding remarks, Monod and Jacob sought "to reconsider the problem of cellular regulation as a whole" (389). Since Bernard Davis had already emphasized the ubiquity and physiological importance of regulatory mechanisms in his opening remarks eight days earlier, Monod and Jacob chose to "center attention on the mechanisms" of control (389). Their presentation was abstract and programmatic. They classified distinct types of cellular control

mechanisms and then presented "possible, plausible, and actual" mechanisms that fit their classificatory scheme (389). It is in the published version of these "General Conclusions" that Monod first introduces the concept of "allosteric inhibition" (390). [8] He presents this new concept as arising from an identification and high-level generalization of the key features of feedback inhibition, and he justifies the need for it by claiming that this generalized phenomenon is likely to be common.

In the published conclusion to the Cold Spring Harbor meeting, Monod distinguishes allosteric inhibition from feedback inhibition, providing ample evidence that he intends for "allosteric inhibition" to refer to structural and mechanistic features of proteins rather than functional properties of biological systems. Through Monod's summary of research on feedback inhibition, along with his discussion of the other mechanisms of cellular regulation discussed at the meeting, we can identify at least three ways that he distinguishes "allosteric inhibition" from "feedback inhibition." First, Monod uses this new concept to signal a shift in focus from physiological effects to causal mechanisms. Both Umbarger and Pardee had discussed how feedback inhibition functioned in the "cellular economy" to promote the efficient use of energy and resources, but neither of these researchers had speculated on the molecular basis for the inhibitory mechanism in the late 1950s. Second, Monod suggests that the new concept of allosteric inhibition will require different research strategies. Whereas the physiological effects of feedback inhibition could be investigated via kinetic analysis, understanding the mechanisms of allosteric inhibition would require structural studies of proteins. Relatedly, a third distinguishing feature of allosteric inhibition was that it referenced the three-dimensional structure of an individual protein.

[8] Changeux (1993) claims this new terminology was added only in the published version.

A constant theme throughout Monod and Jacob's summary of the state of research on cellular regulation is that understanding the observed regulatory effects would require the discovery of underlying causal mechanisms. They begin their remarks by attempting to "classify and define *a priori* the main types of cellular regulatory mechanism" (Ibid., 389). What they conclude from this logical exercise is that "all these mechanisms…are directedly related to the specific molecular structure of the enzymes, or other proteins, concerned" (Ibid., 390). Thus, their focus on understanding molecular mechanisms, rather than physiological effects, brought with it a new emphasis on protein structure that was absent from previous work on feedback inhibition. Understanding molecular mechanisms, on this view, entailed understanding certain aspects of protein structure. The new concept of "allosteric inhibition" was no exception: it referred to structural and mechanistic features of certain proteins.

In the first mention of "allosteric inhibition" in the published conclusion to the 1961 symposium, Monod presents it as a synonym for feedback or endproduct inhibition. However, the discussion in the section entitled "Endproduct or 'Allosteric' Inhibition" reveals that Monod intends for this change to be more than a mere change in name. In this section, he begins by summarizing the results of previous research on feedback inhibition, highlighting Umbarger's work on threonine deaminase and praising it for revealing that the inhibitory relationship only exists between the endproduct and the *first* enzyme in a biosynthetic pathway. He then notes the functional importance of this phenomenon before introducing the new concept:

> As the reports here have shown, endproduct inhibition is extremely widespread in bacteria, insuring immediate and sensitive control over the rate of metabolite biosynthesis in most, if not all, pathways. From the point of view of mechanisms, the most remarkable feature of the Novick-Szilard-Umbarger [feedback] effect is that the inhibitor *is not a steric analogue of the substrate*. We propose to designate this mechanism as "allosteric inhibition." (Ibid., 391, emphasis in original)

While introducing this new concept of allosteric inhibition, Monod draws a distinction between the observed feedback effect, on the one hand, and the mechanism that accounts for it, on the other. In this passage, he transitions from consideration of the biosynthetic "pathways" under feedback control to "the point of view of mechanisms," and he introduces the concept of "allosteric inhibition" to refer to "this mechanism" that explains feedback inhibition rather than the feedback effect itself.

Further support for this interpretation of "allosteric inhibition" comes from Monod's identification of the concept with structural features of proteins. In the above passage, Monod emphasizes the fact that the allosteric inhibitor is "*not a steric analogue of the substrate.*" This finding is of critical importance, according to Monod, because it suggests that the inhibitor does not bind the enzyme at the active site. Because specific binding requires structural complementarity between the enzyme and ligand, if two ligands—e.g., substrate and inhibitor—were structurally dissimilar, then they would be unlikely to bind the enzyme at the same site. This inference leads to the conclusion that the enzyme has two binding sites, the active site where the substrate binds and the allosteric site where the inhibitor binds.

Earlier at the 1961 symposium, Changeux had presented evidence from his work on threonine deaminase in support of this two-site model for allosteric enzymes. He argued that the "dissimilarity of structure" between threonine (substrate) and isoleucine (inhibitor), along with experimental evidence that certain treatments of the native enzyme would lead to loss of inhibitor but not substrate binding, supported a model of this enzyme in which it has two distinct, nonoverlapping binding sites (Changeux 1961, 316). Monod expressed his support for this hypothesis in his concluding remarks: "This [i.e., the desensitization experiments] leads to the conclusion that two distinct, albeit interacting, binding sites exist on native threonine deaminase"

(Monod and Jacob 1961, 391). Monod's endorsement of the two-site model reveals that he was already considering the potential for structural features of enzymes to explain the observed feedback effect. This focus on the molecular mechanism of the effect, and protein structure in particular, differs from earlier work on feedback inhibition. Umbarger's and Pardee's research on feedback inhibition did not so much as speculate about the mechanism of the effect. In fact, in their publications in the late 1950s, these researchers did not even mention the structural dissimilarity between the substrates and inhibitors that Monod identifies in 1961 as the "most remarkable feature."

The centrality of the heterodox reaction kinetics further distinguishes Monod's account of allosteric inhibition from the previous work on feedback inhibition. In their published reports from the meeting, both Changeux and Monod emphasized how the desensitization of the enzyme led to the loss of heterodox kinetics, leading them to conclude that the sigmoidal reaction curve of the enzyme was "directly related to its competence as a regulatory enzyme" (Monod and Jacob 1961, 391). Although Umbarger and Pardee had noted the heterodox kinetics in their publications in the 1950s, both had classified the inhibition as competitive without much further consideration.[9] Monod, in contrast, stressed the novelty of the mechanism in his concluding remarks: "Competitive inhibition, in this system, therefore is not due to *mutually exclusive* binding of the inhibitor and substrate, as in the classical case of steric analogies" (Monod and Jacob 1961, 391, emphasis in original). This novel molecular mechanism that Monod identifies then forms part of his justification for introducing "allosteric inhibition." He hypothesizes that this mechanism "may

[9] Umbarger did postulate that the nonstandard kinetics curve might indicate the reaction was bimolecular, requiring two substrate molecules to bind the enzyme (Umbarger 1956, Umbarger and Brown 1958). By the time of the 1961 meeting, Umbarger also entertained the two-site hypothesis, albeit less enthusiastically than either Changeux or Monod: "The site at which catalysis and endproduct inhibition occur need not be and perhaps never are, identical" (Umbarger 1961, 306).

therefore be a general one for enzymes subject to allosteric inhibition" (Ibid.). With the endorsement of the two-site model, Monod thus distinguishes "allosteric inhibition" from "feedback inhibition" in two ways. First, he makes heterodox reaction kinetics, depicted in the sigmoidal curve, a crucial feature of allosteric enzymes, and second, he shifts the focus from biological pathways to molecular mechanisms, identifying those mechanisms with certain features of protein structure.

While the new concept of allosteric inhibition refers to the molecular mechanism responsible for the observed feedback effect, Monod does not claim to have offered a fully worked out account of that mechanism. In fact, his discussion of future research suggests that he thinks more empirical work is necessary in order to make this abstract concept more concrete. He presents the concept as a reasonably high-level and abstract concept—one which unlike feedback inhibition refers to protein structure and molecular mechanisms—while recognizing that the lower-level structural and mechanistic details still needed to be filled in.

2.4.3 The Allosteric Transition and Changes in Protein Structure

After the Cold Spring Harbor meeting, Changeux had great difficulty obtaining any useable new experimental data on threonine deaminase. Consequently, Monod and his colleagues at the Pasteur Institute shifted their attention to theoretical considerations, publishing a review article on allostery in 1963 (Creager and Gaudillere 1996). With this review, they reinforced and expanded their commitment to the two-site model and to structural explanations of allosteric enzymes more generally. In that paper, they considered three possible models for the interaction between the substrate and inhibitor on the enzyme (Monod et al. 1963): the first, in which the substrate and inhibitor binding sites are overlapping and inhibition occurs via steric hindrance (Model I); the

second, in which the two binding sites are nonoverlapping but sufficiently close that the substrate and inhibitor have direct attractive or repulsive interactions (Model II); or the third, in which the two binding sites are nonoverlapping and the interaction between the inhibitor and substrate is indirect, "mediated entirely by the protein" (Model III) (Ibid., 311). They argued that the available kinetics data for allosteric proteins were "incompatible with Model I and difficult to reconcile with Model II" (Ibid.). They thus conclude that allosteric proteins conform with Model III: they have two distinct binding sites. Since the substrate and inhibitor do not interact directly, they conclude that the allosteric interaction, which enables binding at one site to affect binding at a distant site, must be transmitted through structural changes within the protein itself.

Unlike in 1961 when Changeux first introduced the two-site model, the account of allostery in the 1963 review aimed to solve the problem of action at a distance posed by allostery. Although the two-site model emerged as the consensus from the 1961 meeting, none of its proponents so much as speculated as to how the binding of the inhibitor at the allosteric site led to an effect on catalysis at the active site. In the 1963 review, Monod and his colleagues introduced the allosteric transition as a solution to this problem.[10] On this account, binding of the inhibitor at the allosteric site caused the protein to undergo a conformational change—the allosteric transition—that altered the catalytic properties of the active site:

> The formation of the enzyme-allosteric effector complex…is assumed…to bring about a discrete reversible alteration of the molecular structure of the protein or *allosteric transition,* which modifies the properties of the active site, changing one or several of the kinetic parameters which characterize the biological activity of the protein. (Monod et al. 1963, 307, emphasis in original)

[10] In their contribution to an edited volume on cell differentiation, Jacob and Monod (1963) write, "The sum of these observations very strongly suggests that the action of allosteric inhibitors is not due to a *direct* interference, by steric hindrance, with the binding of substrate, but rather to *an induced alteration of the shape or structure of the enzyme protein, resulting in misfit or reduced fit of the substrate at the active site*." (33, emphasis added). It is possible that this chapter was drafted before the review article.

The allosteric transition thus explained how binding of one molecule at one site on a protein indirectly affected binding of another molecule at a distant site: the allosteric effect was "mediated entirely through the protein…by a conformational alteration" (Ibid., 310). Monod and his colleagues admitted that, at that time, the postulation of a "specifically inducible conformational alteration of protein structure" outstripped the evidence. Nevertheless, they argued that the available evidence, although limited, supported the conclusion that the binding of the effector caused an allosteric protein to switch between two distinct conformational states.

At the time, there was very limited structural data on the few known allosteric enzymes, and neither ATCase nor threonine deaminase had been crystallized. Therefore, in order to assess their structural claims, Monod and his colleagues included hemoglobin in their 1963 review as an "honorary" allosteric enzyme. First suggested by Bernard Davis in the discussion of Changeux's research at the 1961 Cold Spring Harbor meeting, this analogy between allosteric enzymes and hemoglobin was based upon striking qualitative similarities in the kinetics and oxygen-binding behavior of these proteins, as both exhibit the same heterodox sigmoidal saturation curve.[11] However, hemoglobin had a distinct advantage over bacterial enzymes. Because of its medical interest, hemoglobin had been the target of much x-ray crystallographic work. By 1963, Max Perutz had published the structure of hemoglobin, and his results suggested that oxygen-binding induced a conformational change in the protein (Perutz et al., 1960, Muirhead and Perutz 1963). Hence, by treating hemoglobin as an honorary allosteric enzyme, Monod and his colleagues saw

[11] Davis, "Discussion" in Changeux (1961). In an earlier paper, Wyman suggested that hemoglobin might be considered an "honorary" regulatory enzyme (Wyman and Allen 1951), but this paper was unknown to the Monod lab until Monod and Changeux had begun to collaborate with Wyman in 1963 (Changeux 1993). For a review of the history of this analogy, see Brunori (1999).

a way to provide some empirical evidence for the claim that allosteric proteins undergo a conformational change during the allosteric transition.

Although certain aspects of structural allostery would not be fully elaborated by Monod and his colleagues until they developed their formal model in 1965, they had already identified certain features of allosteric proteins and laid out a general molecular framework for explaining allostery in their 1963 review. Building on their previous work, they continued to identify the common features of allosteric enzymes under feedback control. They highlighted structural features in particular, aiming to "formulate certain generalizations concerning the functional structures responsible" for allosteric regulation (Monod et al. 1963, 306). The major innovation of the review was the introduction of the allosteric transition, which described a conformational shift in the protein between two distinct structures. Protein dynamics were therefore already implicit in this account, since the allosteric transition requires a protein to shift from one conformation to another. I contend that this limited role for protein dynamics was easily accommodated by the static view of proteins. However, I will wait to defend this claim until after I present the MWC model, since it retains the allosteric transition and further explicates the dynamics required for a protein to shift between conformational states (see Sec. 2.5.3).

2.5 The MWC Model: The First Explanation of Allostery

Monod's contribution to the history of allostery culminated in 1965 with the publication of the MWC model. In this paper, he and his colleagues completed the research program that he and Jacob had laid out in their concluding remarks to the Cold Spring Harbor meeting in July 1961. With Wyman and Changeux, Monod developed an explanatory model that revealed the molecular

mechanism responsible for the observed allosteric effect. In this section, I consider the model both historically and philosophically. I aim to accomplish three tasks. First, I briefly present the MWC model, arguing that it includes both a mathematical description as well as a particular physical interpretation (Sec. 2.5.1). I then argue that the MWC model offers an explanation for allostery (Sec. 2.5.2). Drawing upon philosophical work on explanation in the biological sciences, I argue that the model is best construed as an abstract mechanistic explanation. Based on the historical evidence, I further argue that Monod and his colleagues intended to offer the MWC model as an explanation for allostery. I then consider the relationship between the model, which posits a limited functional role for protein dynamics, and the static view of proteins (Sec. 2.5.3). Ultimately, I conclude that the features of proteins posited by the model do not require a significant departure from the dominant view.

2.5.1 The Model

In their highly influential 1965 paper, Monod, Wyman, and Changeux presented what would become known as the MWC model of allostery. The model has two parts. It includes a qualitative description of allostery based on the hypothesis that allosteric proteins have certain structural features, and it also includes a mathematical model. The mathematical model, which was derived from the application of principles of equilibrium chemistry to proteins with these unique structural features, was able to reproduce the sigmoidal kinetics behavior of allosteric proteins. Let us briefly consider these two parts of the MWC model in turn.

Allosteric proteins, according to the model, were oligomeric proteins, composed of at least two identical monomers, and each monomer contained one stereospecific binding site for the substrate and one for the allosteric ligand. The protein molecule was hypothesized to be capable

of adopting one of two conformations, and the affinity of the protein for its ligands differed in each conformation: in the R (relaxed) state, the protein bound the substrate with high affinity, while in the T (tensed) state, the protein bound the inhibitor with high affinity. Finally, the model posited that any change in conformation would preserve the molecular symmetry of the molecule, so that a dimeric protein could either be in the RR conformation or the TT conformation.

According to Monod and colleagues, when given a protein with these features, the allosteric effect will occur if the binding of an allosteric ligand shifts the equilibrium between the R and T conformations. For example, if the $R \rightleftharpoons T$ equilibrium favors the T conformation (i.e., most of the protein molecules are in the T state) and the allosteric activator is added, it will stabilize those protein molecules already in the R conformation and shift the equilibrium to increase the proportion of the protein molecules that are in the R conformation. By shifting the $R \rightleftharpoons T$ equilibrium towards the R conformation, the allosteric activator thereby increases the proportion of protein molecules capable of binding the substrate with high affinity, since the high affinity substrate binding site exists only in the R conformation.

Although the 1965 MWC paper does not include any figures depicting this process, Monod presented one during his Nobel lecture in December 1965 and Changeux included a more detailed one in his doctoral dissertation (Fig. 2).[12] This figure depicts the major posits of the model. It shows the preservation of symmetry through the transition between R and T states. Moreover, the visible changes in the binding sites show that the R state has higher affinity for both the allosteric activator and substrate, whereas the T state has higher affinity for the allosteric inhibitor.

[12] Similar pictorial representations of allostery can now be found in practically any biochemistry textbook (e.g., Garrett and Grisham 2011).

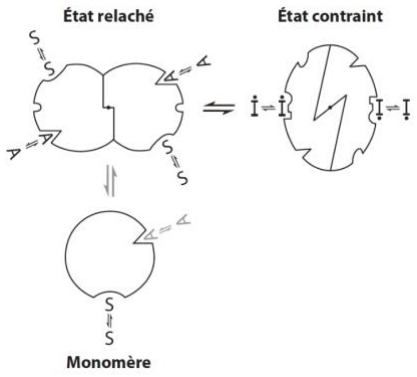

Figure 2 Allosteric Transition (Changeux 2012). Schematic diagram of an allosteric dimer undergoing an allosteric transition from the relaxed (R) state (*état relaché*, left) to the tensed (T) state (*état contraint*, right). The diagram shows the differential binding affinities of these two states for the three ligands: substrate (S), allosteric activator (A), and allosteric inhibitor (I). It also depicts the dissociation between the dimer and the monomer (*monomère*). Redrawn from the original diagram included in Changeux's 1964 doctoral thesis. Republished with permission from Changeux (2012); permission conveyed through Copyright Clearance Center, Inc.

Monod provides a clear statement of how this model accounts for the observed behavior of allosteric proteins in his lecture notes for a course at the Collège de France in 1967. In a dimeric allosteric protein, the two possible conformations have different affinities for the allosteric activator F, with the R state having greater affinity for the activator. Monod writes, "In the presence of F, the [R] state is stabilized." [13] In other words, the preferential formation of the protein-activator complex R-F shifts the equilibrium between the T and R states toward the latter. The cooperative effect then occurs because the "presence of an F associated with *one* of the dimers will increase the affinity of the *other* dimer for F."[14] Thus, the allosteric effect occurs because the binding of the effector to one subunit of a dimer in the R state effectively stabilizes that entire dimer in that state, leaving the second subunit primed to bind a second effector molecule.

In the 1965 paper, Monod and his colleagues develop this qualitative description into a formal, mathematical model. Using the methods and principles of thermodynamics and

[13] Fonds Monod, emphasis in original, translated by the author. I have changed Monod's notation for the two conformational states from A and A' to T and R, respectively, in order to match the modern notation used throughout this chapter.
[14] Ibid., emphasis in original.

equilibrium chemistry, they show that a protein will be allosteric if it has (1) two accessible conformational states, T and R, in which (2) the equilibrium between these two favors the T state, and (3) the R state has greater binding affinity for the allosteric ligand. Furthermore, they note that the cooperative effect will be "more marked…when the $R \rightleftharpoons T$ equilibrium is strongly in favour of T" and when the affinity of the protein for the allosteric ligand is much greater if the protein in the R state rather than the T state (Monod et al. 1965, 91).

From their mathematical treatment of the model posits, they derive a saturation function ($\overline{Y_F}$) for an allosteric protein given by the following equation:

$$\overline{Y_F} = \frac{Lc\alpha(1 + c\alpha)^{n-1} + \alpha(1 - \alpha)^{n-1}}{L(1 - c\alpha)^n + (1-\propto)^n}$$

Here, n is the maximum number of ligand molecules that can bind the protein when it is fully saturated. For example, n = 2 for a dimer, with one ligand binding site per protomer subunit, and n = 4 for hemoglobin, which is a tetramer, with one oxygen-binding site per subunit. The normalized ligand concentration, α is calculated from the actual ligand concentration, F, by the following equation:

$$\alpha = \frac{F}{K_R},$$

in which K_R is the dissociation constant of the protein in the R state for the ligand.[15] By plotting the saturation function over the normalized ligand concentration, Monod and his colleagues constructed binding plots that show the dependence of the allosteric effect on the allosteric constant (L) and the ratio of the microscopic dissociation constants ($c=K_R/K_T$). It replicated the

[15] The smaller the K_R the stronger the affinity of the protein in the R state for the ligand.

sigmoidal saturation curve discovered by Changeux, Pardee, and Umbarger for allosteric enzymes as well as the oxygen-binding curves of hemoglobin.

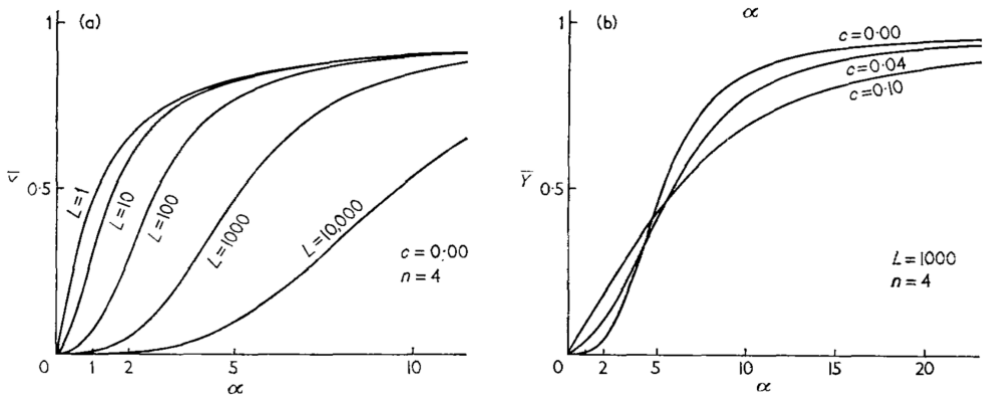

Figure 3 Allosteric Protein Ligand Binding Curves (Monod et al. 1965). Theoretical curves for ligand binding at different concentrations for allosteric proteins with different values for allosteric (L) and dissociation constants (c). Republished with permission from Monod et al. (1965); permission conveyed through Copyright Clearance Center, Inc.

These plots for a tetramer (n = 4), such as hemoglobin, depict the effect of both the allosteric constant (Fig. 3a) and ratio between dissociation constants (Fig. 3b) on the strength of the observed allosteric effect. When L = 1, the T and R conformations are in equilibrium, and there is no allosteric effect (Fig. 3a, leftmost curve). Instead, the saturation function simplifies to produce the hyperbolic curve of classic Michaelis-Menten binding kinetics. As L becomes larger and the equilibrium shifts to favor the T state over the R state, the sigmoidal character of the curve increases. The ratio between the dissociation constants, c, has a similar effect (Fig. 3b). If c = 1 (i.e., if $K_R = K_T$ and the ligand binds each protein conformation with the same affinity), then the saturation function once again reduces to the classic Michaelis-Menten equation. In contrast, as the value of c approaches 0, the allosteric ligand (F) preferentially binds the R state, and the allosteric effect—as evidenced by the sigmoidal character of the saturation curve—increases.

This mathematical model, along with the qualitative interpretation, aimed to reveal the molecular mechanism responsible for allostery. The cooperative effect, which produces the sigmoidal saturation curve, occurs when the binding of a certain ligand to one protein subunit increases the binding affinity for that same ligand on the other subunits. In the paper, Monod and his colleagues used hemoglobin as their primary example. According to the model, hemoglobin exists in two states, one in which all four subunits are in the T conformation and one in which they are all in the R conformation. These two states are in thermodynamic equilibrium, and in the absence of oxygen that equilibrium favors the T state, i.e., the conformation with low affinity for oxygen. When oxygen is present, it binds and stabilizes the R state, thereby shifting the equilibrium in favor of the R state. If a single oxygen molecule binds one of the four subunits of a molecule of hemoglobin that happens to be in the R state, it stabilizes the molecule in that conformation. Because of the symmetry constraint, the stabilization of one subunit of hemoglobin in the R state also stabilizes the other three in the same conformation, and these three subunits are now locked in the high affinity conformation, ready to bind additional oxygen molecules.

2.5.2 The MWC Model is Explanatory

The 1965 MWC model paper has been hugely influential. It is one of the most highly cited theoretical papers of all time, with over 7,500 citations, and the model itself is now textbook science, included in most introductory biochemistry texts.[16] Part of the reason for this success is the fact that the MWC model included both a mathematical model and a simple qualitative

[16] According to Web of Science, the citation count as of May 30, 2021 is 7,501. As a testament to its durability and contemporary relevance, it has garnered at least 100 citations each year for the past 15 years.

interpretation relating the model posits to various structural properties of proteins. In this section, I will argue that the MWC model provides an explanation of allostery. More specifically, I claim that when taken together, the mathematical model along with the physical interpretation of the model and model variables constitute an abstract mechanistic explanation for allostery. This claim is not uncontroversial. In recent scientific reviews, some allosteric researchers have claimed that the MWC model was not an explanatory model of allostery but was instead merely a phenomenological model. The MWC model, they suggest, was able to fit the binding curves for hemoglobin and allosteric enzymes, but it did not provide an explanation for this behavior of proteins.[17] Through a close reading of the 1965 paper and other historical sources, I aim to show that Monod, Wyman, and Changeux intended to offer their model as an explanation for allostery, and more importantly, I will argue that it fits contemporary philosophical accounts of mechanistic explanation. What has led recent scientific commentators astray, I suggest, is the failure to recognize that mechanistic explanations can vary along two different axes: (1) from concrete to abstract and (2) from how-possibly to how-actually explanations. I will argue that the MWC model of allostery is an abstract how-plausibly mechanism.

The primary aim of phenomenological models is to "save the phenomena." They aim to capture the phenomenon of interest, often providing a concise mathematical description of it, but they do so in a way that is largely "detached from theoretical considerations" (Bueno et al. 2012).

[17] For example, Hilser and his colleagues claim that the MWC model of allostery is merely phenomenological, and the first explanation for allostery was not offered until Perutz filled in the mechanistic details in the 1970s: "It is well known that…the MWC…model [is] phenomenological, and consequently, [does] not provide insight into how the structure facilitates allosteric communication between sites. It was the birth of structural biology and the development of the influential stereo-chemical model by Perutz that first addressed this issue and set the course for future allosteric studies" (Motlagh et al. 2014, 331). Similarly, Cui and Karplus (2008) claim that the MWC model is "phenomenological" and does "not answer the fundamental question of how the binding of a ligand or its modification yield the allosteric effect at an atomic level of detail" (1296).

Thus, a phenomenological model might bring two seemingly different phenomena together "under a common principle" or a shared mathematical description, but it will not go much beyond what is given directly by the phenomenon (London 1950, 30).[18] Insofar as a mathematical model is merely a matter of fitting a curve to the data, it can be construed as phenomenological. No matter how well the curve fits the data or how accurately the model predicts new data, such a model is not taken to be explanatory. Especially if a mathematical model has many unconstrained parameters, it might be able to achieve good fit with the data, but the parameters in the model responsible for the goodness of fit have no straightforward physical or realist interpretation and thus fail to be explanatory (Epstein and Forber 2013).

Key to my argument, then, is the following: the MWC model was not primarily a phenomenological model, even though it shared some features with such models. Monod and his colleagues show that the mathematical model—viz., the saturation function ($\overline{Y_F}$)—fits the available binding and kinetics data from hemoglobin and allosteric enzymes. But their primary goal is not to develop a concise mathematical description of the data. Instead, they aimed to explain the observed data by revealing the underlying molecular mechanism that was responsible for the behavior. The authors themselves downplay the importance of the mathematical description and the goodness of fit between their model and the data and instead highlight its abstract explanatory features:

> We feel…that the main interest of the model which we have discussed here does not reside so much in the possibility of describing quantitatively and in detail the complex kinetics of allosteric systems. It rests rather on the concept, which we have tried to develop and justify, that a general and initially simple relationship between symmetry and function may explain

[18] This is how Fritz London described a phenomenological model in order to claim that the London and London model of superconductivity was not a phenomenological model, since it was not limited in this way. In addition to Bueno et al. (2012), see Cartwright et al. (1995) and Suarez and Cartwright (2008) for differing interpretations of this particular model and a fuller discussion of phenomenological and theoretical models.

> the emergence, evolution and properties of oligomeric proteins. (Monod et al. 1965, 116-7)

What is important to the scientists themselves is their physical interpretation of the model—i.e., the fact that relatively simple structural posits about proteins and their binding properties can account for the observed data.

The construction of the mathematical MWC model is also quite different from the construction of phenomenological models, as described by Nancy Cartwright and her colleagues (Cartwright et al. 1995, Suarez and Cartwright 2008). Monod and his colleagues built their model following what Cartwright calls the "theory-driven view of models." They simply applied the principles of equilibrium chemistry to proteins that had the particular structural features and binding properties that they had postulated. Moreover, the only unconstrained variable in the quantitative model is the allosteric constant L. (The ratio between the dissociation constants c is derived from observables.) For this reason, they take the fit between the model and the data as providing some confirmatory evidence for the empirical posits about protein structure used to construct the model. Although the model was constructed after the data were available, Monod and his colleagues can reasonably claim that their model correctly "predicts" these data, since they only used the postulates of the model and standard calculations of equilibrium chemistry (Worrall 1989).The ability of the MWC model to replicate the binding curves of actual allosteric proteins lends support to the model, since the fit did not arise from illicit tweaking of unconstrained parameters (Epstein and Forber 2013). When they consider the MWC model's ability to capture the phenomena, Monod and his colleagues highlight not the fit between the model and the data, but rather the physical interpretation of the model which shows how a few structural features of proteins can account for the observed allosteric behavior. In this way, it was not merely a redescription of the phenomenon in mathematical terms, but an explanation.

According to most philosophical accounts, mechanistic explanation involves decomposing a biological system into its component parts and showing how the components, with their given activities, are organized in such a way that they generate the overall behavior of the system (Machamer et al. 2000, Bechtel and Abrahasen 2005, Craver 2007). The interactions between the components are taken to be causal, while the relationship between the components and the system is constitutive. A mechanism explains by showing that the causal interactions among lower-level components are responsible for the observed system-level behavior. The MWC model of allostery, as presented in the 1965 paper, fits this description. It cites structural features of proteins and their causal interactions with various ligands to explain the observed sigmoidal saturation curve taken to be the hallmark of the allosteric effect.

The model attributes certain features to allosteric proteins: they are oligomeric (i.e., they have more than one subunit), they have two distinct binding sites for the substrate and allosteric ligand on each subunit, they are capable of adopting one of two conformations, and they maintain molecular symmetry such that the subunits in a single molecule are always in the same conformation. Using these structural posits, along with principles of equilibrium chemistry, the MWC model shows why a set of proteins with these features will produce the characteristic sigmoidal curve. At the molecular level, an individual protein is hypothesized to be interconverting between the R and T states, spending more time in the T state, since it is by definition more thermodynamically stable. However, when the protein molecule encounters the allosteric ligand in the R state, it binds the ligand. This binding event is stereospecific and occurs at only the allosteric binding site because of the structural and electronic complementarity between the ligand and the binding cavity on the protein. This binding event essentially traps the protein in the R state, preventing it from converting back to the T state, much like a nail stuck into a cog would prevent

the cog from turning. Because an allosteric protein maintains molecular symmetry, by trapping one of the monomers in the R state, the allosteric ligand also traps the other monomers in the R state, thereby priming them to bind the second ligand at its binding site. This mechanistic account explains why the binding curve of an allosteric protein, such as hemoglobin, is shifted to have higher saturation at lower concentrations, compared to a non-allosteric protein. The binding of one oxygen molecule traps the entire molecule in the R state, and because hemoglobin has four identical subunits, this means that the other three subunits are now trapped in a conformation with high affinity for oxygen-binding.

This description of the molecular mechanism of allostery is an abstract mechanism schema, since it blackboxes certain features of the mechanism. Mechanistic explanations fall along a continuum from concrete mechanisms, which include many molecular details, to abstract mechanisms, which use functional abstractions or otherwise blackbox certain lower-level features. Although Carl Craver has argued that mechanistic explanations that include more lower-level detail are more explanatory, other philosophers have acknowledged that mechanistic explanations can be abstract without sacrificing explanatory power.[19] Indeed, some have suggested that abstract mechanistic explanations can be useful in certain circumstances since they can provide a unifying account of phenomena (Levy 2014). The MWC model is relatively abstract. It does not specify the relationship between the protein and the allosteric ligand in any detail, other than to claim that it must be stereospecific. For instance, it does not show how the orientation of the ligand matters or how certain chemical functional groups on the ligand interact with others on the protein. Similarly,

[19] Craver (2007) and, especially, Kaplan and Craver (2011) seem to favor a "more details the better" conception of mechanistic explanation (but see also Kaplan and Craver (2020), in which they contest this characterization). Machamer et al. (2000) discusses abstract mechanistic explanations that are not incomplete. Levy (2014) presents a more developed account of abstract mechanistic explanation.

the model does not explain how the protein switches between the two conformations, nor does it explain how the protein maintains molecular symmetry during these conformational shifts. What the model does explain, however, is that a protein with just the limited number of features posited by the model will behave in a particular way to bring about the observed allosteric behavior.

Monod and his colleagues find the abstraction of the model to be one of its virtues. It explains the sigmoidal binding curve characteristic of allosteric proteins by making only a few structural assumptions about these proteins. Other than the symmetry requirement and the assumption of distinct stereospecific binding sites for the various ligands, the model is silent about the 3D structure of the allosteric protein:

> No particular assumption has been, or need be, made about the structure of the specific sites or about the structure of the protein, except that it is a symmetrically bonded oligomer, the symmetry of which is *conserved* when it undergoes a transition from one to another state. It is therefore a fairly stringent, even abstract model, since co-operative interactions are not only allowed but even required for any ligand endowed with differential affinity toward the two states of the protein. (Monod et al. 1965, 93, emphasis in original)

What strikes Monod and his colleagues as most useful about the model is the fact that such an abstract model can account for the observed allosteric effect without getting mired in the details of any specific protein. They would likely agree with the sentiment expressed by philosopher Arnon Levy: "Rather than highlighting the vices of omitting mechanistic detail, [they] emphasize the virtues of abstracting from it" (2014, 471). On their view, the abstract nature of the model is a benefit, since it allows them to explain the allosteric behavior of proteins in the absence of atomic details of protein structure.

In addition to being abstract, Monod and his colleagues present the MWC model as a how-plausibly model. Evidence for this claim comes from the full title of the paper: "On the Nature of Allosteric Transitions: A Plausible Model." Biologists often distinguish between "how-possibly" and "how-actually" models (Brandon 1990). How-possibly models satisfy the minimum

constraints placed on the explanation by the empirical data, but the mechanisms they posit are relatively unsupported by the evidence (Craver 2007). For instance, a how-possibly model may attribute activities to components that are only conjectured to exist. In contrast, how-actually models invoke only real components, activities, and the causal relationships that connect them to bring about the explanandum phenomenon: "They show how a mechanism works, not merely how it might work" (Ibid., 112). "How-plausibly" models fit between these extremes; they invoke fewer unknown components, activities, and organizational features and satisfy more constraints than how-possibly models. Monod and his colleagues intended for their model of allostery to be more than a how-possibly model, merely fitting the constraints imposed by the binding data. For this reason, they spent nearly half the paper presenting theoretical arguments and empirical evidence to support the structural posits of the MWC model.

At the time, there was limited structural data to confirm the postulates of the model. Regarding the claims about molecular symmetry and its conservation, Monod and his colleagues conceded that "next to nothing is known, from direct evidence, regarding this problem" (Ibid., 106). However, they argued that there were both good theoretical reasons and some compelling evidence from structural studies of hemoglobin to suggest that the postulates of the model were plausible. When Monod first began drafting the model sometime in 1963, he already had access to some empirical evidence that suggested that the subunit structure of allosteric proteins might be important to their function. His student Changeux (1963) offered an explanation for the "desensitization" effect that involved subunit structure. Through his experiments on threonine deaminase, he demonstrated that the enzyme lost its allosteric behavior during urea denaturation because it dissociated into monomers. Also, Gerhardt and Pardee (1963) had presented evidence at the 1963 Cold Spring Harbor Symposium that cited changes in the subunit structure of ATCase

to explain its allosteric activity. In their paper the following year, they presented a model of ATCase allostery, in which effector binding altered the subunit structure of the protein (Gerhardt and Pardee 1964).[20] Empirical results such as these provided some support for the postulate that allosteric proteins were oligomeric.

The claim that the subunit structure of allosteric proteins was symmetrical was an important innovation of the MWC model. There was less empirical evidence for this postulate of the model, but Monod was committed to it for largely theoretical reasons (and perhaps aesthetic reasons). More so than either Wyman or Changeux, he was committed to the idea that oligomeric proteins would be most likely to have symmetric subunit structure. He arrived at this conclusion, in part, from thermodynamic considerations.[21] In a letter to Wyman in September 1964, Monod considered the merits of the "general claim that the symmetric states of molecules are, in principle and all other things being equal, more stable than asymmetrical states." He concluded, "Although I cannot prove this statement with any precision, I remain convinced that there is some truth to it."[22] Monod's commitment to this general thermodynamic claim about protein symmetry would show up again in the published paper, but this time it came with a theoretical argument.

In the paper, Monod and his colleagues began from certain facts about oligomeric proteins that had been gleaned from experimental work: (1) these proteins were stable under a range of conditions (i.e., they did not readily dissociate into monomers), (2) the association between the

[20] This model was later discovered to be incorrect.
[21] In a commemorative issue of the *Journal of Molecular Biology* on the 50th anniversary of the publication of the 1963 review, Changeux recounts the origins of the 1965 MWC model paper. He attributes the symmetry considerations mostly to the 'personal reflections of Jacques Monod about the three-dimensional organization of proteins', as well as communication with Wyman, Francis Crick, and others. He claims, however, that he was not very involved in these discussions and was not too keen on that section of the published paper: "The final version of the manuscript included at the end a section about thermodynamic considerations on symmetry that I found far from my way of thinking" (Changeux 2013).
[22] Letter from Wyman to Monod dated January 10, 1963 (Fonds Monod).

subunits in these proteins was highly specific (i.e., monomers of a normally oligomeric protein would readily re-associate), and (3) the association between subunits did not involve covalent bonds but rather "a multiplicity of non-covalent bonds" (Monod et al. 1965, 106). They then considered two possible types of association between the subunits in these proteins: isologous and heterologous. Isologous association required that "any group which contributes to the binding in one [subunit] furnishes precisely the same contribution to the other [subunit]" (Ibid.). This reciprocal association results in a two-fold axis of rotational symmetry. Heterologous association, in contrast, has no element of symmetry. It occurs when binding site pairings between the subunits are unique. Heterologous associations lacking symmetry would likely give rise to long chains of polymers of different lengths. The properties of these polymers would be incompatible with the known properties of oligomers, such as their stability and specificity. The general properties of oligomeric proteins therefore suggested that these proteins were comprised of few subunits that formed a "closed structure where all the [subunits] use the same binding sets," which necessarily introduced elements of symmetry into their structure (Ibid., 108).[23] Through theoretical arguments such as this, Monod attempted to show that the symmetry constraint of the model was plausible.

Monod and his colleagues looked to hemoglobin as another avenue to provide evidence that the structural posits of the MWC model were plausible. Wyman had studied hemoglobin and its binding interactions long before his collaboration with Monod and had previously suggested that hemoglobin behaved similarly to certain regulatory enzymes (Wyman and Allen 1951, see

[23] Monod reaffirmed his commitment to these arguments in a 1968 address at a Nobel Symposium on biological symmetry. He poses the question: "Are there reasons to believe that these small aggregates [monomers arranged in quaternary structures] are constructed according to some symmetry rule?" His response: "A fairly strong general argument (based on considerations of finiteness, stability and self-assembly) can be made in response to…[this] question." (Monod 1968, 23) Interestingly, it seems Monod's reasoning and conclusion has withstood the test of time. In a review on protein structure and function, Goodsell and Olson (2000) state that "the majority of soluble and membrane-bound proteins in modern cells are symmetrical, oligomeric complexes" (105). Thus, according to the authors, "symmetry is the rule rather than the exception for proteins" (107).

also Brunori 1999). In the 1950s, he had also hypothesized that a connection between the cooperative binding of hemoglobin and changes in its subunit structure (Wyman 1948, 1951, 1963). By the time of their collaboration on the MWC model, Monod and Wyman both recognized that hemoglobin offered the best empirical support for the structural postulates of the model.

In the paper, they pointed to two sources of evidence from crystallographic studies of hemoglobin to show their abstract mechanism of allostery was, indeed, a plausible mechanism. First, they presented Perutz and colleagues (1960) structure of hemoglobin, which showed that its four subunits interacted via isologous associations and thus contained two elements of symmetry, as offering the clearest "direct experimental evidence" for the symmetry constraint of the model (Monod et al. 1965, 108). Second, they argued that Perutz's crystallographic work on hemoglobin, together with Kendrew's work on myoglobin, suggested that the allosteric properties of hemoglobin were related to changes in its subunit structure that occurred upon oxygen binding. Myoglobin, which had been crystalized in 1960 by Kendrew, is naturally a monomer, but otherwise it is similar to one subunit of tetrameric hemoglobin (Kendrew et al. 1960). Unlike hemoglobin, it is non-allosteric and has greater affinity for oxygen. Whereas the conformation of myoglobin did not change upon oxygen-binding, hemoglobin exhibited a marked difference in conformation between its oxygen-bound and free forms, suggesting that a change in subunit structure was responsible for the observed allosteric behavior.

Perutz's work provided additional support for the model. He had been able to obtain crystal structures for hemoglobin in its free state (reduced hemoglobin) as well as in its oxidized state (oxyhemoglobin) (Perutz et al. 1960; Muirhead and Perutz, 1963; Perutz et al. 1964). In the language of allostery, Perutz and his colleagues had isolated and crystalized hemoglobin in both the T (free) and R (oxygen-bound) state. Therefore, Monod and his colleagues could thereby point

to the structural differences between these two states as evidence in favor of the MWC model. The T and R states of hemoglobin showed that although there were significant differences in their quaternary structures, both maintained the symmetry of the tetramer. According to Monod and his colleagues, the crystal structures of hemoglobin provided "a virtually complete illustration of the model" (Monod et al. 1965, 112).

This evidence from protein crystallography, coupled with Monod's theoretical arguments about protein structure, played a central role in the explanation of allostery. It showed that the basic structural posits of the MWC model were plausible features of allosteric proteins. Therefore, this evidence allowed Monod and his colleagues to argue that the model was not a mere how-possibly model able to reproduce the sigmoidal binding curve, but instead was a how-plausibly model, citing only features of protein structure that were likely to obtain. With the publication of the MWC model in 1965, Monod and his colleagues thus made good on his suggestion from 1961: they demonstrated that the allosteric behavior of proteins could be explained via reference to a shared molecular mechanism.

2.5.3 The MWC Model and the Static View of Proteins

The MWC model of allostery showed that protein dynamics were relevant to their function. Specifically, allosteric proteins were hypothesized to undergo an allosteric transition between two, discrete 3D conformations, the R and T states. Nevertheless, I argue that the MWC account of allostery did not present a severe challenge to the prevailing static view of proteins. Although the model hypothesized that allosteric proteins could exist in one of two states, both these states exhibited all the features attributed to proteins under the static view. And even though a protein must be sufficiently flexible to transition between these two conformations, the MWC model

focused attention on the structural features of the two end-state conformations and not the dynamics required to transition between them. Like the two protein structures posited by Koshland's induced-fit model of catalysis, the two states of an allosteric protein were easily accommodated by the static view of proteins.

The MWC model backgrounds the motion that occurs during the allosteric transition and foregrounds the structural features of the two end-state structures. Although protein dynamics are inevitable if the protein is to undergo an allosteric transition, these dynamics are not the focus of the explanation of allostery. Because the model posits a preexisting equilibrium between the R and T states, it blackboxes the dynamic motion that enables the interconversion between states. Recent scientific commentators have noted the curious way in which the MWC model and other accounts that focus on protein structure manage to efface the protein dynamics required of allosteric proteins:

> Explanations of allosteric effects have typically addressed only an interconversion between principal conformations of proteins, paradoxically both highlighting and ignoring their fundamental flexibility. (Clarkson and Lee 2004, 12456)

Although an allosteric protein is dynamic, according to the MWC model, one does not need to know anything about that motion in order to explain allostery. Instead, the explanation focuses solely on the structural features and binding properties of the two static end-state structures.

The one place in the 1965 paper where Monod and his colleagues do explicitly consider the dynamic interconversion between the R and T states occurs in their discussion of symmetry. They describe various possible mechanisms to support their claim that any allosteric transition between two states is likely to be symmetry-preserving. They argue that changes in bond connectivity are likely to be the drivers of conformational changes. For example, they consider a model in which the T state is, by itself, unstable (Fig. 4).

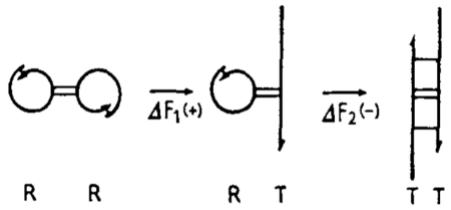

Figure 4 Symmetry-Preserving Allosteric Transition (Monod et al. 1965). Hypothetical symmetry-preserving model of dissociation between identical subunits in a dimeric protein. Republished with permission from Monod et al. (1965); permission conveyed through Copyright Clearance Center, Inc.

In this case, the RT transition state is also highly unstable. If this unstable, asymmetrical state were to occur spontaneously, then it would immediately transition to one of the stable, symmetrical states. It would either be stabilized in the TT conformation, which requires the breaking of bonds within the other subunit, leading it to also adopt the T conformation so that the two subunits could form stabilizing bonds. Or the subunit in the unstable T state would collapse back into the stable R state, leading the molecule back to the RR state. Monod and his colleagues include general hypothetical mechanisms of the allosteric transition, such as this, in order to support their argument that these transitions will be symmetry-preserving.

This discussion, however, also supports my contention that the protein dynamics required to transition between states is relatively unimportant to the MWC model. Monod and his colleagues' analysis of this transition show that the asymmetrical transition state is highly unstable and therefore highly transitory. According to their reasoning, any asymmetrical intermediate will be extremely short-lived, so any protein molecule is likely to spend nearly all of its time in either of the symmetrical end states. This description of the allosteric transition therefore justifies the model's focus on the end-state structures, rather than the intermediate. The dynamics required for the protein to shift between states are included in the model only as a posit that an allosteric protein can exist in two distinct conformations: two states "are reversibly accessible to allosteric oligomers" (Monod et al. 1965, 90). Even in this case where the discussion of protein dynamics is

most explicit, it still functions to reinforce the stability and discreteness of the two end-state structures. Most often, however, the author's discussion of the symmetry requirement focuses only on the end states, without any reference to the protein dynamics necessary to transition between states. On my view, the entire symmetry discussion suggests that Monod and his colleagues are largely wedded to the static view of protein structure, since it makes little sense to make claims about geometric symmetry for molecules without stable 3D conformations.

Further evidence for Monod's commitment to the static view of proteins comes from other papers and lectures he delivered in the mid-1960s. These sources show that Monod was committed to the static view of proteins as well as the structure-function rule, which held that the molecule structure of proteins and other biomolecules determined their functions (Sarkar 2008). He argues that the elucidation of protein structure, especially via x-ray crystallography, is the best way to advance knowledge of protein function more generally (Monod 1968). In his Nobel lecture in December 1965, he makes even more sweeping claims about the role of structural studies in molecular biology. He claims that "the ambition of molecular biology is to interpret the essential properties of organisms in terms of molecular structures" (Monod 1965, 207).[24] Although "that objective has already been achieved for DNA," Monod acknowledges that there was still much work to be done before the same could be said of proteins (Ibid.). Nevertheless, he claims the MWC account of allostery represents a step toward this goal of offering structural explanations for protein functions.

[24] This is a strikingly different conception of the role of structure in molecular biology than the one Martin Karplus—one of the proponents of the dynamic view of protein—would advocate in two decades: "The long-range goal of molecular approaches to biology is to describe living systems in terms of physics and chemistry….Great progress has been made over the last thirty years in applying the equations to chemical problems involving the structures and reactions of small molecules." (Karplus 1988, ix)

Although the MWC model of allostery implicitly posits a functional role for protein dynamics, the dynamics that enable transitions between two discrete conformational states are taken as background conditions of the model. Monod and his collaborators do not acknowledge any functional role for the fast timescale, dynamic motions of proteins—viz., the wigglings and jigglings of their atoms—that form the core of the later dynamic view of proteins. Thus, this account of allostery is strikingly different from the accounts of dynamic allostery that would begin to emerge in the 1980s. It requires that an allosteric protein adopt one of two "discrete states" (Monod et al. 1965, 116). And these two states are assumed to be compact, rigid, and static, and therefore accessible via x-ray crystallography. The role of molecular symmetry in the MWC model also assumes that the R and T states have stable structures, since without it, there could be no axis of symmetry. For these reasons, the MWC model of allostery, like Koshland's induced-fit model of catalysis, does not pose a significance challenge to the prevailing static view of proteins.

2.6 Conclusion

In this chapter, I have conducted a historical and philosophical analysis of some of the scientific work on proteins since the late nineteenth century. My primary focus has been accounts of protein structure and function during the heyday of classical molecular biology in the mid-twentieth century. I have shown how the static view of proteins, which held that proteins were compact, rigid, and static molecules, became the prevailing view during this time. Because of this view, protein dynamics were largely ignored, since, unlike protein structure, they were not thought to play a role in protein function. Against this backdrop, two different accounts of protein function were developed in the late 1950s and early 1960s that posited a role for protein dynamics:

Koshland's induced-fit model of enzyme catalysis and Monod and his colleagues' model of allostery.

In the first part of the chapter, I developed a historical account of the emergence and evolution of these two cases, and I argued that these accounts of protein function were easily accommodated by the static view of proteins. In both cases, the scientists claimed that a certain type of protein could adopt two different structures, but each protein structure had a discrete, fixed, and stable 3D structure, in accord with the static view. In the second part of the chapter, I looked more closely at the development of the MWC model of allostery, since this episode played a central role in the birth of structural biology. I presented an updated historical account of the evolution of the concept of allostery, first introduced by Monod in 1961 and then focused on the MWC model itself. While some recent scientific commentators have claimed that the MWC model is not explanatory, I argued that it fits philosophical accounts of mechanistic explanation. When it was published in 1965, the model was a how-plausibly model, since some of the structural posits were not yet confirmed empirically. Nevertheless, it was (and still is) an abstract explanation, revealing the common features of allosteric proteins and the molecular mechanism of allostery while omitting details. Therefore, I concluded that the MWC model is an explanation for allostery, namely, an abstract mechanistic explanation.

3.0 From Static to Dynamic: A Historical Account of the Emergence of the Dynamic View of Proteins

3.1 Introduction

In structural and molecular biology, proteins are one of the primary objects of study. Over the past century, as the importance of proteins in a myriad of biological processes has become clearer, the concept of the protein has evolved significantly. This chapter focuses on the most recent stage of this evolution, namely, the origin of the dynamic view of proteins in the second half of the twentieth century and the challenge it posed to the static, or structural, view of classic molecular biology. Although the origins of molecular biology have been the focus of much historical research, little has been done on this recent history (Morange 2018, Sarkar 2008). The rise of the dynamic view, or what Michel Morange (2020) has called the "new view" of proteins, raises two interesting historical puzzles. First, why and how did the dynamic view emerge in the 1970s and 1980s as an alternative to the dominant structural view? And, second, what explains the quarter-century time lag between the origin of the dynamic view and its eventual widespread acceptance in the 2000s and 2010s?

The answer to both of these questions highlights the importance of theory within this history. Although technological advances and the discovery of anomalies played a part, I argue that theoretical understanding of protein dynamics was the primary driver behind the emergence of this new view of proteins. The application of general thermodynamic principles to proteins fostered the dynamic view in the 1970s and 1980s, and commitment to these principles led a small cadre of scientists to seek out anomalous cases of protein dynamics that could not be explained by

the structural view. These researchers slowly accumulated empirical evidence for the dynamic view from the mid-1980s through the 1990s, and their empirical findings, along with additional evidence for protein dynamics resulting from new technologies, eventually convinced the majority of protein scientists. Explaining the uptake of the dynamic view therefore involves the discovery of anomalies aided by new technologies, but the commitment to treating proteins as small thermodynamic systems ultimately drove that search for anomalies. Hence, I argue that the emergence of the dynamic view of proteins as a competitor to the structural view was, at its core, theory-driven.

In what follows, I will present this history in four stages: the origins of the dynamic view of proteins in the 1970s and 1980s (Sec. 3.2), its slow reception (Sec. 3.3) and the role of true believers and their search for anomalies in the mid-1980s and 1990s (Sec. 3.4), and finally, the widespread acceptance of this new view of proteins in the 2000s and 2010s (Sec. 3.5).

3.2 Origins of the Dynamic View

The static view, which represented proteins as molecules with rigid and largely static structures, was the prevailing view of proteins leading into the 1970s. Closely tied to experimental work, this view "followed from the dominant role of high-resolution X-ray crystallography," which "led to an image of biomolecules with every atom fixed in place" (Brooks et al. 1988, 3). Advances in protein crystallography since the 1940s had produced a tranche of new structures of proteins. The future looked sufficiently bright that in 1970 a group of researchers started work on a repository for all this new structural data, which would become the Protein Data Bank in 1976 (Berman 2008). By 1980, the databank would include nearly 70 protein structures. As Alan

Cooper, a key player in the history of protein dynamics, recalled in a recent interview, the 1970s were "a time when the field was being dominated by protein crystallographers" who were producing, "for the very first time, images of protein molecules".[25] Despite the "immeasurable value" of x-ray crystallography, the focus on this technique caused many protein scientists to have "too rigid a view of the protein as it actually exists in solution," since it obscured the dynamic properties of proteins (Robson 1977, 577).

Against this backdrop, a small group of researchers, who were committed to treating proteins as thermodynamic systems, developed the dynamic view of proteins. They argued that the static view failed to accurately represent proteins, in part, because the structures obtained via x-ray crystallography abstracted away all the dynamic fluctuations of proteins in solution:

> The structures determined by X-ray diffraction of protein crystals in particular are inevitably biased toward a rigid picture because the molecules are "frozen" by the constraint of crystal packing and periodicity. (Robson 1977, 577)

Introduced as a corrective, the dynamic view thus aimed to present a more accurate picture of proteins based upon theoretical rather than experimental analysis.

Proponents of the dynamic view showed that if you applied the principles of statistical thermodynamics to proteins, then you would be led inexorably to the conclusion that they were constantly undergoing structural fluctuations (Cooper 1976). The individual protein molecule is a relatively small thermodynamic system, and the electrostatic interactions that maintain the 3D structure are relatively weak (Frauenfelder 1983). Therefore, constant collisions between the protein and solvent—i.e., "the continual Brownian-motion-like buffeting by solvent molecules"—would supply enough thermal energy to disrupt these weak interactions and cause the protein to be in constant motion (Cooper 1980, 489). "Such thermal pummeling" of the protein would

[25] Video interview by the author with Alan Cooper, 27 May 2019.

"induce fluctuations in thermodynamic properties" and cause significant fluctuations in structure (Ibid.). For individual protein molecules, these structural fluctuations would be random, and molecules would be rapidly interconverting among many energetically quasi-degenerate microstates.

The thermodynamic properties of the system that reveal that individual protein molecules will adopt different microstate conformations are left out of the static view of proteins. According to Cooper, this fact explains an apparent tension between the two views of proteins. The results of x-ray crystallography show "a compact structure in which the polypeptide chain is precisely folded to give a tightly interlocking, rigid molecule," while theoretical and experimental evidence for the dynamic view suggests a "fluid, dynamic structure for globular proteins involving rapid conformational fluctuations" (Cooper 1976, 2740). The tension is dissolved, according to Cooper, when one recognizes that the representation of protein structure obtained via x-ray crystallography describes the bulk properties of the population of protein molecules. The crystal structure obscures the dynamics of individual protein molecules by collapsing all the conformational substates into a single average structure—a structure that is unlikely to exist in solution. Because the energy distribution function is skewed toward higher energies, "it is highly improbable that at any instant of time even one individual protein molecule [in solution] has the average structure" (Karplus and McCammon 1986, 42).

By applying thermodynamic principles to proteins, the dynamic view sought to capture the structural fluctuations of individual molecules in solution rather than the average structure. It depicted an individual molecule as conducting a biased walk through conformation space, "wandering in a haphazard and non-periodic fashion amongst a multitude of possible conformational states," (Cooper and Dryden 1984, 107). This picture of protein molecules led

proponents of the dynamic view to reject a key assumption of the structural view. They denied that a population of protein molecules in solution would adopt a single discrete structure. As Gregorio Weber memorably describes it, the static, structural protein does not exist:

> The protein molecule model resulting from the X-ray crystallographic observations is a "platonic" protein, well removed in its perfection from the kicking and screaming "stochastic" molecule that we infer must exist in solution. (Weber 1975, 65)

The thermodynamic analysis of proteins demonstrated that individual protein molecules were in constant motion, and any population of protein molecules in solution exists in a multiplicity of conformational microstates.

The early proponents of this new view of proteins were all well-versed in thermodynamics. Hans Frauenfelder, Georgio Careri, and Alan Cooper were trained as physicists, but like many others at the time, switched to molecular and structural biology and biophysics. The most famous of this group, Frauenfelder completed his degree in physics in Switzerland before taking up a position in the Department of Physics at the University of Illinois Urbana-Champaign in the 1950s. Once there, he turned his attention to the physics of proteins, publishing influential theoretical and experimental work on protein thermodynamics over his forty-year career. Careri followed a similar trajectory. According to an obituary written by Frauenfelder, Careri first "establish[ed] [his] reputation in physics" and then "[saw] the light and switch[ed] to the life sciences" (Frauenfelder 2012, 3). He began his career studying the statistical physics of liquids, but later focused his attention on the physics of proteins. He also wrote a number of reviews advocating the dynamic view of proteins that "had considerable impact" (Ibid.). Cooper studied physics as an undergraduate at the University of Manchester in the 1960s, but switched to biophysics for graduate school after attending a lecture that "suddenly made [him] realize that there were

fascinating problems in biology that might by solvable by physical approaches."[26] He later conducted postdoctoral work at Oxford and Yale, before taking up a professorship at the University of Glasgow in 1976.

Although not trained in physics, the intellectual pedigrees of Weber and Martin Karplus also prepared them to study the thermodynamic aspects of proteins. After completing an MD at the University of Buenos Aires, Weber won a fellowship for graduate study at Cambridge, where he worked under the well-known enzymologist, Malcolm Dixon (Jameson 1998). He then pioneered fluorescence quenching techniques that are now widely used in the field of protein science. This technique proved well-suited to study the dynamics of proteins in solution. Consequently, Weber became one of the early advocates for the dynamic view of proteins. Karplus began his studies at Caltech in biology but ultimately transferred to chemistry to work with Linus Pauling (Karplus 2006). The theoretical background he developed during his time at Caltech helped him to build the first molecular dynamics (MD) simulations of proteins that would win him the Nobel prize in 2013. Despite their differences, these early adopters of the dynamic view all shared a strong background in thermodynamics.

Because of their professional training, these early proponents of the dynamic view were primed to think about proteins as thermodynamic systems. In a book-length review of protein dynamics for *Advances in Chemical Physics*, Karplus and his colleagues claimed that "to chemists and physicists," unlike others trained in biology, "it is self-evident that polymers and proteins undergo significant fluctuations at room temperature" (Brooks et al 1988, 3). This way of thinking, along with their theoretical training, enabled them to be capable champions of the dynamic view of proteins in the 1970s and 1980s. As proselytizers for the view, they saw that part of their role

[26] Cooper interview, 27 May 2019.

was to show the importance of dynamics to those who might not be at home around thermodynamic equations. In the introduction to a 1984 review, Cooper acknowledges that the thermodynamics and mathematics that are necessary in order to arrive at the dynamic view are not everyone's cup of tea:

> I will attempt…to approach the problem of protein dynamics and function much as a physicist or physical chemist might have done from basic principles, without having had our advantage of seeing the structures or playing with the models. This may seem bizarre, and will certainly not be to everyone's taste. It will involve some mathematics (ugh!) and some thermodynamics (even more ugh!)… (Cooper 1984, 181-2)

For scientists like Cooper, the goal was to share their insights from the application of physics to biology. They took it as their task to help others, especially those who were not theoretically inclined, see the inevitability of the structural fluctuations at the heart of the dynamic view.

Although there were certainly multiple motivations for defending the dynamic view, the primary justifications for these early defenders were theoretical. The limited empirical evidence then available was insufficient to account for their strong adherence.[27] Instead, I argue, it was their commitment to the thermodynamic principles that underwrote their commitment to the dynamic view, since they saw the dynamic view as an inevitable consequence of these principles. One only had to apply certain principles of thermodynamics to proteins to conclude that they must be dynamic. The thermodynamic principles invoked were simple and uncontroversial, and the thermodynamic analysis did not require attributing any special biological properties to proteins:

> These dynamic fluctuations are simply a consequence of the thermodynamic uncertainty of small systems and are not a particular property of protein molecules as such—even a lump of mild steel the size of an individual protein molecule would show similar fluctuations. (Cooper 1980, 490)

[27]At that time there was limited evidence that proteins were dynamic from H-D exchange experiments, fluorescence, and other techniques. For contemporaneous reviews of the available empirical evidence, see Debrunner and Frauenfelder (1982) and Englander and Kallenbach (1984).

In a recent interview, Cooper confirmed this theoretical commitment to the dynamic view. When I asked him how he knew in the 1980s that proteins were dynamic, he responded, "otherwise I'd have to abandon all the physics I'd ever learned."[28] For Cooper and other early proponents, the dynamic view followed directly from unassailable thermodynamic principles.

From this analysis, we see that the dynamic view was not proposed in the 1970s and 1980s because of the failure of the static view to explain empirical data, nor was it based on any new technological advances that revealed protein dynamics empirically.[29] In fact, many of the early advocates of the dynamic view *blamed* technological advances in crystallography for "creat[ing] the misconception…that the atoms in a protein are fixed in position" (Karplus and McCammon 1986, 42). Rather, the early adopters of the dynamic view developed and defended it primarily for theoretical reasons, and they were well-positioned to draw conclusions from treating proteins as small thermodynamic systems because of their training in physics and chemistry.

3.3 Slow Reception of the Dynamic View

3.3.1 Allostery: The First Link Between Dynamics and Function

In the 1970s, these early defenders of the dynamic view typically left open the question of whether the inevitable dynamic fluctuations in protein structure were relevant to protein function. They focused instead on correcting the overly rigid view from x-ray crystallography in order to

[28] Cooper interview, 27 May 2019.
[29] Historian of biology, Michel Morange (2020) has suggested that technological advances were the primary driver of this change, while structural biologist Vincent Hilser has suggested that it was the discovery of anomalies (Hilser et al. 2012, Motlagh et al. 2014).

more accurately represent proteins. It was not until the early 1980s that these scientists would begin to consistently consider the possibility that dynamic fluctuations were necessary for protein function.

In their early papers, defenders of the dynamic view stressed the inaccuracy of the static view of proteins. They demonstrated that protein dynamics were an inevitable consequence of thermodynamic principles and showed how the dynamic view was necessary in order to account for certain experimental observations.[30] However, there was little explicit discussion of how structural fluctuations were causally relevant to protein function.[31] For example, in a critical note in the *Proceedings of the National Academy of Sciences* published in 1976, Cooper presented the dynamic view of proteins as necessary in order to understand experimental results, such as fluorescence quenching and relaxation studies, and explain away the apparent paradox between the representation of proteins that emerged from these techniques and the static representation from x-ray crystallography (Cooper 1976). But he does not go so far as to claim that the dynamic fluctuations at the heart of this view are essential for the proper functioning of proteins. His only mention of protein function comes at the end of the note, and the connection he draws between dynamics and function is relatively weak: the "complete understanding of the nature and function of protein molecules will require knowledge not only of their mean properties, but also of their dynamic characteristics" (2741). The major goal of Cooper and others in these early years seems

[30] See, for example, Cooper (1976, 1980, 1984), Frauenfelder (1983), McCammon et al (1977).

[31] The question of functional relevance turned around the interconversions between conformational microstates that arose from the "wiggling and jiggling" of the atoms in protein molecules. Large translational motions, that approximated rigid body motion, were already documented at that time, for hemoglobin, myoglobin, and lysozyme (Cooper 1980). For example, the crystal structure of deoxymyoglobin showed that the heme groups, which bind oxygen, were buried in the interior of the protein, inaccessible to oxygen. Motion of the protein would therefore be required for it to bind oxygen.

to be to correct the misrepresentation of proteins as static, regardless of the role of dynamics in protein function.

By the 1980s, these defenders of the dynamic view become more emboldened and consequently more explicit in discussing the causal relevance of dynamics for protein function. For example, in his concluding remarks at the Ciba Foundation Symposium on "Mobility and Function in Proteins and Nucleic Acids" in March 1982, Frauenfelder claimed that one major goal of the meeting was to learn "if dynamic motion is essential for biological function" (Frauenfelder 1983, 329). Cooper posed similar questions about the relationship between protein dynamics and function at the conclusion of a crystallography meeting on dynamics the following November.[32] He began the final discussion by asking a series of questions:

> Do dynamics have function? Are there any aspects of protein function for which dynamics offers the only explanation? Can dynamic processes provide alternative explanations of apparently well-understood phenomena? (Cooper 1983, 36)

According to his abstract, he then proceeded to offer examples that aimed to answer these questions in the affirmative: he clearly believed that the dynamic properties of proteins played a causal role in carrying out their biological functions. However, Cooper's rhetorical strategy—he claims some of his conclusions may be "(possibly) outrageous"—suggests that in the fall of 1983 the connection between dynamics and protein function was still tenuous and not widely accepted outside the community of early adopters of the dynamic view.

By the time of this crystallography conference in the fall of 1983, Cooper was already working with his graduate student David Dryden to develop a model of dynamic allostery. Their formal model, published in July 1984, provided the first proof of principle that protein dynamics could be the cause of at least one protein function. Allostery, or the phenomenon whereby binding

[32] I thank Cooper for telling me about this conference and sharing the conference program with me.

at one site on a protein affects binding at a second site, is a property of proteins that has an important regulatory function (Changeux and Edelstein 2006). Since the 1960s when Jacques Monod first identified the phenomenon and developed the influential MWC model, explanations of allostery had focused on protein structure.[33] These explanations held that allosteric proteins could adopt one of two rigid structures, an active and an inactive conformation. The binding of an allosteric inhibitor at one site would induce a conformational change, shifting the protein into an inactive conformation that could not bind the substrate molecule at the active site. This structural change thus explained how binding at one site on a protein could affect binding events happening at a distant site. Cooper and Dryden's model of dynamic allostery was strikingly different. It demonstrated that changes in protein dynamics, rather than structure, could be the cause of allostery. Their model showed how the binding of an allosteric inhibitor could induce changes in the flexibility of the protein that could transmit a signal from one binding site to a distant binding site, without any changes to the average structure of the protein.

In interviews with Cooper and Dryden, they have both indicated that their motivation for developing the model was to demonstrate that protein functions could arise from changes in dynamics, rather than structure. There was good theoretical evidence as well as limited empirical evidence that proteins were dynamic, and they suspected that dynamics were relevant to function. Dryden recently reflected on their thinking at the time:

> We [knew]…that…protein dynamics is there. The question [was] can we make it do something? That was [Cooper's] question to me….what things can we do with dynamics? Can we do allostery with dynamics? That was the fundamental question that I guess we started with: can you do allostery without a conformational change, just through dynamics?[34]

---

[33] For early examples of structural explanations of allostery, see Monod and Jacob (1961), Monod et al. (1963, 1965), and Koshland et al. 1966. The highly influential MWC model developed in Monod et al. (1965) has over 7,400 citations and is highly cited today (Web of Science accessed 10 December 2020).
[34] Video interview by the author with David Dryden, 24 June 2019.

Dryden recalled Cooper saying to him that allostery would be a good topic, since it "had been dominated by structural ideas" and "the contribution of entropy to allostery [had] been ignored."[35] According to Cooper, their model of dynamic allostery aimed to address the question, "What functions can be explained in terms of dynamics, or at least equally well in terms of dynamics as might be explained in terms of static conformational changes?"[36] It offered a way to show that inevitable dynamic motions could provide sufficient energetic coupling between two sites on a protein to produce an allosteric effect.

With the model, Cooper and Dryden sought to use basic principles from statistical thermodynamics to demonstrate that protein dynamics could be the cause of this biological function. Throughout the paper, they highlight how their model uses only uncontroversial physics—what they call "textbook statistical thermodynamics"—and assumes no special biological properties of proteins (1984, 104). The "dynamic phenomena are not unique to biological macromolecules," they claim, but instead are "simply a manifestation of heat energy" (Ibid.). In other words, proteins just happen to be small enough that the stochastic fluctuations in thermal energy are likely to have a significant effect on their structure. They also provide interpretations of the mathematical representations included in the model to make them accessible to those outside biophysics.[37] Yet, although their model demonstrated for the first time that the

[35] Ibid.
[36] Cooper interview, 27 May 2019.
[37] Cooper's papers, including Cooper and Dryden (1984), are typically written with clear physical interpretations of the mathematics and rationales for any assumptions. According to Cooper, he was often "trying to rationalize things for a biochemistry audience"—an audience that was unlikely to be trained in physics and thermodynamics. He claimed that the physical picture was often more important for understanding a phenomenon than the mathematical description: "For me, the mathematical description is necessary sometimes, although not always. But I've got to have a physical picture, something I can [use to] explain to myself as well as others to support [the mathematical description], and I guess that's the way I write papers on whatever subject." Cooper interview, 24 May 2019.

protein dynamics are relevant to protein function, it was met with a collective shrug by most protein scientists.

3.3.2 Reasons for the Slow Reception

The reception of this paper, along with another important paper in protein dynamics, can serve as a rough proxy for the slow uptake of the dynamic view of proteins (Fig. 5).[38] Citation data for Cooper and Dryden's paper, as well as the early Molecular Dynamic (MD) simulation published by Karplus' group in 1977, show relatively little activity in the 1980s and 1990s (aside from an early spike in interest for the novel MD simulation method) but reveal a marked uptick beginning in the early 2000s, when interest in the dynamic view first began to break into mainstream structural and molecular biology. These "sleeper" papers raise the question: what was happening within the community of protein scientists and structural biologists from the mid-1980s that might explain this quiet reception?[39] In what follows, I discuss four different factors that help explain the lack of traction of the dynamic view within the wider community of structural biologists and protein scientists during this time.

---

[38] I include McCammon et al. (1977) here in order to show that the citation trend is unlikely to be an artifact of the Cooper and Dryden (1984) paper. The results from Karplus' lab reported in McCammon et al. (1977) provide some of the earliest evidence for protein dynamics, although they do not link these dynamics to function. They use a new computational method: molecular dynamic (MD) simulation. Karplus, who would go on to win the 2013 Nobel prize in chemistry for this work, was already well-established at Harvard and much better known than Cooper at the University of Glasgow. Karplus' paper was published in *Nature*, whereas Cooper's was published in the *European Biophysics Journal*. Despite these different pedigrees, the citation data for both papers reveals the same trend (excluding the initial spike in interest in Karplus' use of the new method): both see their citations remain relatively flat through the 1980s and 90s, but then begin to increase in the 2000s.

[39] See Gringas (2010) and Ke et al. (2015) for recent characterizations of sleeper, or sleeping beauty, papers in the philosophical and scientific literatures, respectively.

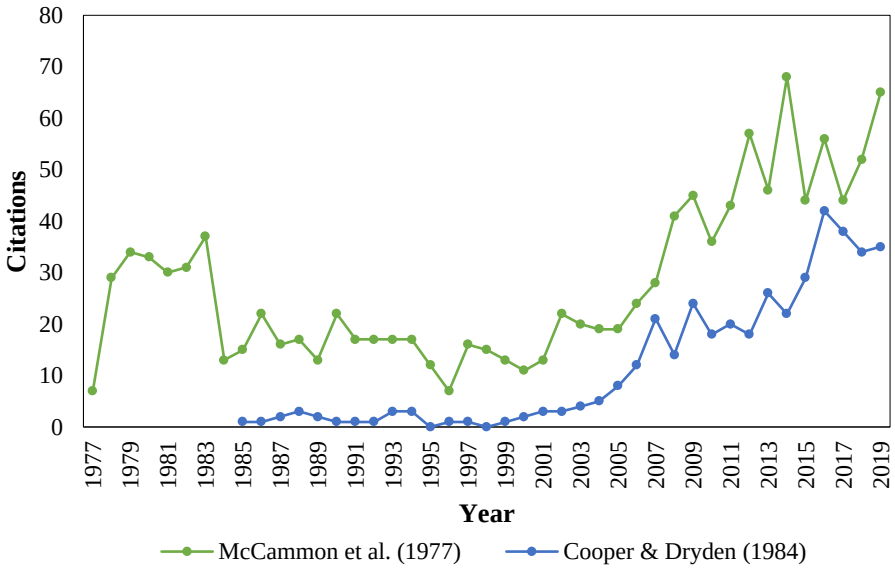

Figure 5 Citation Data for Dynamic View Papers. Citations for Cooper and Dryden (1984) and McCammon et al. (1977), 1977-2019. Citation data obtained from Web of Science. Accessed on 10 December 2020.

With the publication of Cooper and Dryden's model of dynamic allostery in 1984, there were two possible explanations for this protein function: it could be caused by changes in structure or changes in dynamics. Most scientists at the time were committed to the static view of proteins, and they saw the debate about the cause of protein functions as a relative significance debate (Beatty 1997). When they acknowledged Cooper and Dryden's dynamic account at all, they tended to downplay its significance. They admitted that the model of dynamic allostery might obtain empirically but suggested this was likely to be rare. For example, Russell Petter and his colleagues constructed a synthetic allosteric compound, and its allosteric behavior, they claimed, was fully explained by structural changes. In a footnote, they cite Cooper and Dryden (1984), noting that "alternative mechanisms have been advanced to account for cooperative behavior which do not require conformational changes," but then stress that such dynamic changes are not significant in their allosteric system (Petter et al. 1990, 3861, fn 20). Scientists, such as Petter, could thus pay lip-service to the dynamic view while at the same time dismissing it as largely irrelevant to their

structural studies. This limited acknowledgement of dynamic allostery functioned as a way for them to flag their awareness of the new literature on protein dynamics and dynamic allostery while continuing to research and explain allostery using the classic structural accounts.

Scientists adopting this strategy would often acknowledge that proteins are composed of atoms that are in constant motion but contend that these dynamic fluctuations are minor and could be abstracted away. Cooper and Dryden's (1984) model of dynamic allostery without structural change, they would argue, was an extreme case unlikely to obtain in many biological systems. Compared to changes in structure, changes in dynamics, were insignificant. Responding to Cooper's earlier work on protein dynamics, Thomas Creighton adopted this strategy:

> There is thus extensive evidence for varying degrees of flexibility in the folded states of proteins. This has often been considered to be inconsistent with the static models derived by crystallography. However, it is a general thermodynamic requirement that a small system with the thermal properties of proteins have transient fluctuations involving relatively large energies, even at thermodynamic equilibrium (Cooper 1976). What is not yet clear is the magnitude of the conformational changes which are produced by this thermal energy. (Creighton 1978, 247)

In this passage, Creighton acknowledges protein dynamics, but he later dismisses them as rare or insignificant, citing simulation evidence (e.g., "high energy barriers appeared to prevent the interior atoms from wandering far from their positions observed crystallographically") and x-ray evidence (e.g., "in no instance has substantial change of the conformation of a globular structural unit been observed") (Ibid., 247-8). Creighton, Petter, and other scientists who adopted this strategy were thus quite similar to the arch-adaptationist described by Gould and Lewontin (1979): "You acknowledge the rival, but circumscribe its domain of action so narrowly that it cannot have any importance in the affairs of nature" and then "congratulate yourself for being such undogmatic and ecumenical chap" (151). Regardless of the specifics of their arguments, these researchers

acknowledged the new evidence for the dynamic view but at the same time defended the status quo, continuing to use static structures to draw inferences about protein function.

Another factor that helps explain the slow uptake of the dynamic view arises from the fact that protein dynamics occur across a vast range of spatiotemporal scales and are empirically accessible by different methods. Philosophers of physics describe this problem as the "tyranny of scales." It arises because many systems "manifest radically different, dominant behaviors at different length scales" (Batterman 2013, 255). Although scientists have developed ways to explain behavior at different scales, "it is often hard to relate these scale-based models to each other" (Ibid.).

This problem played out in studying protein dynamics in the 1980s and 1990s since much research "center[ed] on the detailed application of certain physical and chemical techniques and on the analysis of data obtained from their application to specific systems" (Richards 1983, 2). The vast difference in timescales of different protein motions exacerbates this problem. The dynamic fluctuations at the heart of the dynamic view are often incredibly fast, occurring at the pico- to nanosecond timescale, whereas the large translational motions that were cited in structural explanations of allostery were typically in the millisecond to second timescale. Experimental methods focus on only one set of dynamic properties at a given timescale; thus, the dynamics observed via one technique could not be seen by the other. The landmark MD simulation of the globular protein BPTI presented by Karplus' group in 1977 is a case in point (McCammon et al. 1977). Because the simulation modeled the dynamic fluctuations of the protein for 8.8 ps, only those motions on the picosecond scale (e.g., oscillations of individual atoms, rotation of side chains) could be captured by this method (Case 2000). This MD simulation would be unable to resolve the large, 13° rotation of the hemoglobin subunits that Max Perutz (1970) discovered in

his crystallographic studies of allostery in hemoglobin. Because these different methods gave empirical access to dynamic features of proteins at different scales, scientists researching the "same" phenomenon—viz., allostery—could work largely independently, with surprisingly little productive crosstalk.

Obtaining evidence for distinct dynamic properties required completely different models and theoretical assumptions as well as methods, and it was a nontrivial matter to integrate the results at different scales. Those endorsing the dynamic view of proteins recognized this problem: the use of different techniques to study proteins was inevitable, since "there exists no single method which covers the whole range of interest from picoseconds to seconds" (Demchenko 1986, 125). However, the majority of researchers who operated under the dominant static view of proteins overlooked it, since their method of choice, x-ray crystallography, rendered invisible the rapid local dynamics at the core of the dynamic view. The challenge of integrating these different methods, combined with the methodological entrenchment of x-ray crystallographers, further explains why the dynamic view of proteins received such little traction in the 1980s and 1990s.

A third factor relates to the perception of theoretical work by the larger community of structural biologists and protein scientists and the related problem of publishing such work. The early proponents of the dynamic view were largely persuaded of its merits based upon theoretical considerations, and much of their work cited theoretical conclusions that resulted from applying thermodynamic theory to proteins. They often found it difficult to publish theoretical work in high impact journals. In an autobiographical essay, Karplus described his difficulties in publishing such a paper in 1971:

> We submitted the paper, which received excellent reviews, but came back with a rejection letter stating that because there was no experimental evidence to support our results, it was not certain that the conclusions were correct. This was my first experience with *Nature* and

> with the difficulty of publishing theoretical results related to biology, particularly in high impact journals. (Karplus 2006, 29)

According to Karplus, theoreticians are still in this Catch-22 today: "if theory agrees with experiment it is not interesting because the result is already known, whereas if one is making a prediction, then it is not publishable because there is no evidence that the prediction is correct" (Ibid., 29).

This same problem beset Cooper and Dryden when they were choosing where to publish their model of dynamic allostery in the spring of 1984. When asked why they published the paper in the *European Biophysics Journal,* Cooper recalled an experience similar to Karplus' that he had earlier in his career and his desire to avoid a repeat experience with the model of dynamic allostery paper:

> I was probably conscious of the fact that it's going to be a fairly controversial idea. If I tried to go to one of the more highly cited *Nature* or *Science* or wherever I'd have brought out a bit of a battle to get it published. And I didn't want to get involved in that kind of argument. The 1976 paper that came out in *PNAS* [*Proceedings of the National Academy of Sciences*] had previously been rejected by *Nature* for…what I seem to recall were fatuitous reasons—just that the referee didn't understand the argument.[40]

Dryden remembered a corroborating fact about their publication decision: "Alan [Cooper] knew the editor of the *European Biophysics Journal*…. So that's why Alan chose that, because he knew that the editor wouldn't throw it out on its ear straight away."[41] Thus, even though their goal was to influence the molecular and structural biologists who were wedded to the static view of proteins, they ultimately published in a place where it was likely to be read only by those already interested in biophysics. The general bias against theoretical work, along with the fact that these papers were "a little bit controversial" since they were "getting away from the static picture" and "trying to

[40] Cooper interview, 24 May 2019.
[41] Dryden interview, 24 June 2019.

overcome some of these ideas from the majority," led to the marginalization of many early papers that championed the dynamic view.[42] And this, in turn, slowed the uptake of the dynamic view by the wider community of protein scientists.

Relatedly, explanations of protein behavior that invoke statistical thermodynamics are very different from those that cite changes in structural features. This contrast is evident when we compare Cooper and Dryden's model of dynamic allostery to the classic structural models of allostery (Monod et al., 1965, Koshland et al. 1966). The former requires knowledge of statistical thermodynamics and mathematics, along with reasoning about thermodynamic equilibrium, whereas the latter treat proteins as molecular machines and require only mechanistic reasoning about structural, or conformational changes. Writing in a review in 1980, Cooper acknowledged this difference and cited it as a reason why dynamic explanations of protein function were slow to be adopted:

> The concept that "conformational changes" are involved is widely accepted and is a comforting view since we might then picture the mechanism as similar to the relative motions of springs and levers and pulleys familiar from macroscopic machines. (Cooper 1980, 474)

Historian of molecular biology Michel Morange (2012) reiterates this point, specifically in the context of teaching. He claims that anyone who has taught biochemistry and structural biology knows that mechanistic explanations of protein behavior that cite structural changes are "understood much more rapidly" than thermodynamic explanations that cite changes in equilibrium (Ibid., 14). Advocates of the dynamic view used types of reasoning and explanation that differed from those used by most molecular and structural biologists, who were trained to infer function from structure, and this difference is another factor that accounts, in part, for the slow

[42] Cooper interview, 27 May 2019.

spread of the dynamic view. Taken together, these four factors help to explain why the dynamic view of proteins received little traction outside the small community of early adopters in the 1980s and 1990s.

3.4 True Believers and the Search for Anomalies

3.4.1 Early Dynamicist Researchers

The rise of the dynamic view depended on the small community of researchers who were studying protein dynamics empirically during this time. Like the earliest proponents of the dynamic view, they were already committed to treating proteins as small thermodynamic systems, and it was their theoretical commitment, I argue, that led them on their quest to discover anomalies that would ultimately undermine the static view of proteins.

Within this group of dynamicists, researchers typically fell into one of two camps. On the one hand, there were the true believers, who were proselytizers for the dynamic view. In their publications, they explicitly argued that their findings demonstrated the inadequacy of the static view and refuted previous explanations of protein function that cited only structure. For example, Gregory Reinhart, a true believer *par excellence*, titled one of his papers on allostery, "The Failure of a Two-State Model," arguing that the evidence for dynamic allostery in a particular enzyme system showed the inadequacy of allosteric models that cited structural changes between an active and inactive conformation (Johnson and Reinhart 1997). True believers often used multiple experimental systems and different techniques in their quest to discover anomalous cases that could not be accommodated by the static view. On the other hand, there were the workhorse

dynamicists. These researchers were also committed to the dynamic view of proteins, but they did not actively promulgate the view nor did they present their findings as refutations of structural explanations. Often, they worked on one experimental system, and they were methodologically entrenched within a particular technique, content to engage primarily with others using similar techniques or working on the same system.

Among the scientists working on dynamics in the 1980s and 1990s, the true believers bore most of the responsibility for promoting the dynamic approach to studying proteins. They were proselytizers for the view with deep-seated commitments to dynamics arising from application of the principles of thermodynamics. Reinhart is one of the most prominent scientists who fits this description. He completed his doctorate in Biochemistry at the University of Wisconsin Madison in 1979. By the early 1980s, he had established his own lab at the University of Oklahoma, investigating the regulation of metabolic enzymes, such as phosphofructokinase (PFK). His theoretical commitment to dynamic rather than structural approaches to protein function was evident from the beginning of his career: he published two theoretical papers in the 1980s touting the superiority of thermodynamic approaches to allostery (Reinhart 1983, 1988).

In his empirical work, Reinhart discovered and characterized different cases of dynamic allostery. His lab conducted activity assays on PFK and other enzymes with various ligands to ascertain their role in regulating enzyme activity.[43] Their research demonstrated that the interaction between PFK and its many ligands included different allosteric couplings in which the binding of one ligand would affect the binding of a second or even third ligand. They identified cases in which the binding of a ligand did not result in significant structural changes in the protein, indicating that

[43] See, for example, Reinhart and Hartleip (1986, 1987), Reinhart et al. (1989), Johnson and Reinhart (1994a, 1994b, 1997), Braxton et al. (1996)

the allosteric linkage was caused by changes in protein dynamics.[44] Highlighting these empirical results, Reinhart argued that the MWC model of structural allostery was an inadequate framework for explaining complex enzyme-ligand interactions. As the lab continued to accumulate cases of dynamic allostery in PFK and other regulatory enzymes, Reinhart almost always framed the group's empirical findings as anomalies that could only be explained by the dynamic view.[45]

Other scientists working on protein dynamics during this time—the workhorse dynamicists—were not proselytizers for the dynamic view, but their research also demonstrated the importance of dynamics for protein function. Mikael Akke and his colleagues working on the calmodulin superfamily of calcium-binding proteins in the 1990s provide good examples of scientists of this sort. Akke began investigating the dynamic behavior of Calbindin D_{9k} upon Ca^{2+} binding during his graduate work split between the University of Lund in Sweden and the Scripps Research Institute in La Jolla, California. He then continued this work after joining the faculty at the University of Lund in 1998. His research focused primarily on this one experimental system, using one technique, namely, NMR.[46] It demonstrated that changes in protein dynamics

---

[44] For instance, in one early study, Reinhart showed that the inhibitory effect of MgATP, a known allosteric inhibitor of PFK, was attenuated by decreasing the pH (Reinhart 1985). These results indicated that H^+ functions as a third ligand and that H^+ binding is positively coupled to substrate binding.

[45] For examples of Reinhart and his colleagues presenting evidence for dynamic allostery as evidence against structural models, such as the MWC model, see Reinhart et al. (1989) and Johnson and Reinhart (1997).

[46] In their research during the 1990s, Akke and his colleagues match the description of workhorse dynamicists in two ways. First, they focus their attention on one experimental system—the calmodulin superfamily. Of the papers Akke published between 1990 and 1996, for example, nearly 80% (17/19), studied the dynamic-driven allosteric behavior of proteins in the calmodulin superfamily. Second, their work on protein dynamics relies largely on only one technique, namely, NMR. In fact, much of their early work in the 1990s aimed at developing and perfecting the use of NMR as a probe of protein dynamics (Skelton et al 1992, 1993; Akke et al. 1993, 1994). Both of these features of Akke and his collaborators help to explain why they were able to continue to publish evidence of dynamic allostery throughout the 1990s without much engagement at all with structural accounts of allostery. The research communities in which they considered themselves members were ones that were already committed to the study of protein dynamics in a particular protein family using a specific experimental technique. At that time, Akke and his collaborators were already committed to studying protein dynamics using NMR, and their work focused on demonstrating the causal relevance and importance of these dynamics in generating allosteric behavior within the calmodulin superfamily of proteins.

contributed to allostery in calmodulin, and he cited his findings as validation of Cooper and Dryden's model of dynamic allostery (Akke et al. 1991, 1993). But he did not stress the difficulties his findings posed for the static view and structural accounts of allostery. In fact, unlike Reinhart, he did not cite previous structural models of allostery in his papers at all. His discussion engaged more with scientists working on calmodulin than others advocating the dynamic view of proteins. Rather than focusing on the negative import of his findings for structural models, Akke's primary goal was to show that "entropic effects associated with the protein dynamics play an important role" in the production of allosteric coupling in calmodulin (Akke et al. 1991, 173). Like the true believers, Akke and other workhorse dynamicists furnished empirical evidence for the dynamic view through their research in the 1980s and 1990s, but they did so without framing their results as a refutation of structural explanations.

3.4.2 The Thermodynamic Framework

Despite their different approaches, both the true believers and workhorse dynamicists had strong theoretical commitments to dynamics, and their research produced empirical results that supported dynamic rather than structural explanations of protein function. A key reason for their success in discovering dynamic anomalies was their commitment to a thermodynamic framework.[47] Whereas previous structural explanations of protein function were qualitative, dynamic explanations were quantitative. Moreover, the thermodynamic framework provided a method for researchers to distinguish structural and dynamic contributions to function.

[47] Cooper (1976, 1980) and Reinhart (1983, 1988) lay out this thermodynamic framework and apply it to allostery.

In order to analyze allostery, for instance, researchers used the Gibbs free energy equation: $\Delta\Delta G = \Delta\Delta H - T\Delta\Delta S$. The binding of an allosteric inhibitor to a protein could induce changes in dynamics, which would be captured by the entropy term (ΔS), or structural changes, which would be reflected in the enthalpy term (ΔH). Taken together, the changes in these two values would determine the sign and magnitude of the allosteric coupling between two ligands. For positive allostery, $\Delta\Delta G$ is negative because the binding of the allosteric ligand makes the binding of the second ligand more favorable. If $\Delta\Delta G$ is positive, then the allosteric ligand is a negative effector since it makes the binding of the second ligand less favorable. This framework provided a way to quantify the energetic coupling between two ligands at distinct binding sites. It also enabled researchers to distinguish the enthalpic and entropic contributions to coupling so that they could determine whether dynamic or structural changes in the protein were responsible for allostery. That is, they could use the Gibbs equation to identify which component was larger and therefore responsible for determining the sign of the coupling free energy.

Cooper and Dryden's (1984) model used this framework to demonstrate that dynamic allostery is entropy-driven:

> We have shown that dynamically mediated cooperativity should be entropy driven: that is, binding of a second ligand is made thermodynamically more favorable because of a less negative $\Delta\Delta S$. (Ibid., 108-9)

In their paper, they presented two models of dynamic allostery, that is, two different processes that could lead to sizeable coupling energies driven by changes in entropy. First, they showed how ligand binding could affect vibrational normal modes in the protein—without any concomitant structural changes—which would in turn influence the binding of a second ligand at a distant site. Second, they developed a model to show how ligand-binding might lead to an allosteric response

by changing the distribution of conformational substates in a population of protein molecules in solution, without changing the average structure.

In this second case, they sought to model what would later come to be called the "conformational entropy" of the protein (Igumenova et al. 2006, Frederick et al. 2007). The conformational entropy can be thought of as a measure of the number conformational microstates that are accessible to the protein molecules at a given temperature. The number of conformational microstates available tracks the flexibility of the protein molecule: a relatively rigid protein will have few thermally accessible microstate conformations, and most will be similar to the mean structure, whereas a relatively flexible or dynamic protein will be able to sample many different conformational microstates, with some of these having structures that are significantly displaced from the mean.

Cooper and Dryden's (1984) model of dynamic allostery showed how changes in conformational entropy could be the driving force behind an allosteric response in relatively flexible proteins. In this case, individual protein molecules would be able to sample a large number of conformational microstates in solution. Then, as these protein molecules bound the allosteric ligand, they would stiffen, losing dynamic flexibility, thereby becoming more rigid. This first binding event, like most ligand-binding processes, would be enthalpically favorable ($\Delta H_1 < 0$), but there would be a significant unfavorable loss of conformational entropy ($\Delta S_1 << 0$). The ΔG_1 for the first binding event might then be either positive or negative, depending on the protein-ligand system. However, what matters for allosteric coupling is the difference between the free energy of this first binding event (ΔG_1) and that of the second binding event (ΔG_2). In the case of dynamic allostery described by the Cooper and Dryden model, the second ligand binds the protein to yield a similarly favorable ΔH_2, but now the protein molecules are already relatively stiff, with access

to only a limited number of conformational microstates, all of which are relatively close to the mean structure. In a sense, then, the protein is trapped in conformational microstates with high binding affinity for the second ligand. Binding of the second ligand will therefore yield a much smaller decrease in conformational entropy ($\Delta S_2 << \Delta S_1$) because the majority of the entropic cost of binding the ligand was already paid by the first ligand. Because the enthalpy of binding the first and second ligand was roughly equivalent ($\Delta\Delta H \approx 0$) but the entropic cost of binding greatly decreased ($\Delta\Delta S >> 0$), the coupling free energy in this case is likely to be favorable ($\Delta\Delta G < 0$), indicating the two binding sites are positively coupled.

As Cooper and Dryden's model illustrates, the thermodynamic framework proved useful for those committed to the dynamic view of proteins who wanted to demonstrate the functional importance of protein dynamics. In the case of allostery, it provided a way to distinguish dynamic from structural allostery, and it revealed when and how dynamic allostery was likely to occur. Those scientists seeking to discover cases of dynamic allostery should look for flexible proteins in which ligand-binding induced large changes in entropy. Moreover, the thermodynamic analysis of allostery also showed why structural accounts of allostery, such as the MWC model, would necessarily fail to be able to account for dynamic allostery, since they only considered the enthalpic contribution to free energy.

3.4.3 The Search for Anomalies

Aided by this thermodynamic framework, the true believers embarked on their quest in the 1980s to find dynamic "anomalies" to demonstrate the importance of protein dynamics and reveal the inadequacy of the static view of proteins. These were anomalies in name only—at least to the true believers who went out searching for them—because these scientists were already

theoretically committed to the thermodynamic framework and the dynamic view of proteins. The true believers thus went out in search of empirical cases that could confirm Cooper and Dryden's model of dynamic allostery and help convince others to endorse the dynamic view. Because of the thermodynamic framework, they knew exactly what they were looking for and where to look in advance, namely, cases in which the entropic contribution swamped the enthalpic contribution. Already committed to dynamic accounts of allostery for theoretical reasons, they were particularly keen to find empirical evidence that would confirm it.

The overarching goal of discovering dynamic anomalies greatly influenced the true believers' methodological choices. Their desire to accumulate enough instances of dynamic allostery in order to argue for the replacement of the structural approach with the dynamic one affected their choice of both experimental techniques and experimental systems. These scientists typically avoided using x-ray crystallography, which was the dominant technique for structural characterization at that time, and instead opted for experimental techniques that would be better able to track changes in dynamics. By using microcalorimetry, fluorescence and Raman spectroscopy, activity assays and the like, these researchers were able to experimentally obtain data from which they could then derive the individual contributions of entropy ($\Delta\Delta S$) and enthalpy ($\Delta\Delta H$) to the free energy of allosteric coupling ($\Delta\Delta G$). When coupled with the thermodynamic framework, these techniques enabled them to identify cases of dynamic allostery. Moreover, because they wanted to amass as many anomalous cases as possible, the true believers would often use these techniques to study multiple experimental systems. Driven by their theoretical commitments to protein dynamics, the true believers spent the late 1980s and 1990s building the case for the dynamic approach, using all the resources at their disposal to accumulate empirical evidence that supported dynamic rather than structural accounts of allostery.

Let us consider Reinhart in greater detail, since his research program illustrates this approach to the dynamic view of proteins. The publications from his group in the 1980s and 1990s show that Reinhart and his collaborators were seeking anomalous cases of dynamic allostery that could bolster the case for the dynamic view of proteins. Focused on investigating the dynamic behavior of allosteric enzymes, his research used a variety of different experimental techniques and studied different enzyme-ligand systems. Most of his research utilized enzyme activity assays, including kinetic assays, which could be used to obtain constants from which $\Delta\Delta H$ and $\Delta\Delta S$ could be derived.[48] However, during this time, his group also expanded their studies to include fluorescence polarization and quenching assays—experimental techniques that could yield more specific information about the structural fluctuations involved in protein dynamics.[49] All these techniques were chosen because they would be the most likely to yield information about protein dynamics. As for their object of study, the preferred experimental system of the Reinhart lab was the metabolic enzyme PFK. Although the majority of their work involved this enzyme, they were able to study many different enzyme-ligand systems using it, since PFK is embedded in a complex regulatory network. Hence, even though they focused primarily on one regulatory enzyme, they were able to study its different allosteric responses to a variety of ligands, including MgADP, MgATP, fructose-6-phosphate, H^+, and phosphenolpyruvate. In this way, they were able to obtain multiple different examples of allosteric coupling, in which the entropy component played an important role in determining the sign and magnitude of that coupling.[50]

---

[48] For examples, see Reinhart et al (1989), Johnson and Reinhart (1994a), Braxton et al. (1996), and Johnson and Reinhart (1997).
[49] See Reinhart and Hartleip (1987) and Johnson and Reinhart (1994a, 1994b, 1997).
[50] See Reinhart (1985), Reinhart and Hartleip (1987), Johnson and Reinhart (1994a, 1994b, 1997), and Braxton et al. (1996).

One of the main aims of their research program was to produce sufficient evidence for dynamic allostery to show that it was a common feature of allosteric systems. They wanted this evidence in order to convince those who thought changes in protein dynamics were functionally insignificant compared to structural changes—i.e., those on the opposite side of the relative significance debate. Their results with different PFK-ligand systems provided some evidence that dynamic allostery was not rare. To further bolster their view, Reinhart and his group expanded their program to study two additional regulatory enzymes: isocitrate dehydrogenase (Reinhart et al. 1989) and carbamoyl phosphate synthetase (Braxton et al. 1996). The primary motivation for including these additional enzymes in their studies was to show that dynamic allostery was not peculiar to PFK. For all three enzymes, Reinhart and his colleagues demonstrated that the entropy component was necessary to explain the magnitude of the observed allosteric effect. They offered this as suggestive evidence that dynamic allostery might be common, at least among regulatory enzymes.[51] All these methodological choices—i.e., which enzymes to study and which techniques to employ—reveal Reinhart's desire to search for cases of allostery that would support the dynamic account that emerged from thermodynamic analysis.

Reinhart's group also fully exploited the thermodynamic framework to help them discover cases of dynamic allostery. The majority of the group's allostery publications from the mid-1980s onward discussed the importance of entropy to specific allosteric systems. They emphasized both the importance of entropy in giving rise to allosteric coupling and the need to include the entropy component in explanations of allostery. They wanted to show that entropic changes, such as

[51] Although they would like to make this generalizing induction from these two cases, the authors caution against doing so before more evidence confirms this finding: "It is of course inappropriate, without further investigation, to attribute all regulatory enzymes the characteristics reported herein for phosofructokinase and isocitrase dehydrogenase. In those enzymes to which these properties do pertain, however, an additional role for protein dynamics is suggested." (Reinhart et al. 1989, 4035).

changes in conformational entropy, were important contributors to the free energy of allosteric coupling. In some cases, Reinhart and his colleagues were able to identify allosteric systems in which the change in entropy upon ligand binding was sufficiently large to overcome the enthalpy and determine the sign of the coupling free energy. The Gibbs free energy equation ($\Delta\Delta G = \Delta\Delta H - T\Delta\Delta S$) shows that this will occur when the absolute value of the entropy component is greater than that of the enthalpy component: $|T\Delta\Delta S| > |\Delta\Delta H|$. Because the enthalpy and entropy components oppose each other, the sign of the coupling free energy in these cases will flip from positive to negative. These cases of dynamic allostery provide the most compelling evidence for the dynamic view, since structural accounts—if they were even possible to give—would predict negative cooperativity between the two sites when the system, in fact, exhibits positive cooperativity. Thus, the failure to consider the entropic component to allosteric coupling energy on systems such as this leads to complete explanatory failure. Reinhart and colleagues (1989) described these cases as ones in which "the coupling entropy plays an important role in establishing both the nature and the magnitude of the allosteric response" (4032). The change in entropy controls the "nature" of the allosteric response, on their view, since it determines whether a certain ligand acts as an allosteric activator or inhibitor. Throughout the 1990s, Reinhart and his colleagues continued to stress that the entropy component could reverse the expected coupling interaction in these paradigm cases of dynamic allostery and could thereby determine the "nature of the allosteric effect" (Braxton et al. 1996, 11922, see also Johnson and Reinhart 1997).

Even in cases in which the entropy component is not greater than the enthalpy component, Reinhart and his colleagues were still quick to point out the effects of entropy in order to support the dynamic view of allostery. In these cases, the entropy component is not large enough to overcome the enthalpy component. It thus cannot determine the "nature" of the allosteric

interaction, but these researchers emphasized that the change in entropy would still affect the magnitude of allosteric coupling. They adopted this strategy in their discussion of the observed allosteric coupling in PFK:

> Net coupling energies responsible for PFK's allosteric responses are always smaller in absolute value than the corresponding enthalpy, indicating that enthalpy-entropy compensation is occurring. (Johnson and Reinhart 1994a, 2636)

Here, the researchers failed to find exactly what they were searching for—viz., another case in which the changes in protein dynamics were the driver of allosteric coupling. In this particular case, they discovered that the entropic contributions to coupling free energy were insufficient to overcome the enthalpic contribution. Nevertheless, they still spin their findings as evidence of the importance of protein dynamics and the need for accounts of allostery to consider entropic changes. They made what amounts to little more than a mathematical point: the experimentally determined $\Delta\Delta G$ for the allosteric system is less than it would be if it were wholly determined by $\Delta\Delta H$, which according to the Gibbs free energy equation means that $\Delta\Delta S$ is greater than zero. Thus, as the researchers claimed, their results showed that changes in entropy, even in this system, affected the magnitude of allosteric coupling.

In other studies, Reinhart and his colleagues went even further in order to show that entropy was an important determinant of allosteric coupling and thereby ought to be included in any explanation of allosteric behavior. Guided by the thermodynamic framework, they exploited the fact that increases in temperature have a disproportionate effect on the entropic component; increasing the temperature increases the importance of the entropy component ($T\Delta\Delta S$) relative to the enthalpic component ($\Delta\Delta H$). In these experiments, Reinhart and his colleagues discovered that the allosteric systems they were studying exhibited the same behavior as the PFK case just described—the entropy and enthalpy components were opposing but the enthalpy term dominated.

However, by raising the temperature from 25°C to 37°C in one case (Braxton et al. 1996) and to 40°C in another (Johnson and Reinhart 1997), they were able to increase the magnitude of the entropy component until it surpassed that of the enthalpy component. By raising the temperature, the researchers demonstrated that they could flip the sign of the coupling free energy, causing a given allosteric ligand to switch its function, from an activator to an inhibitor. This "temperature-induced inversion" of the allosteric effect showed that in systems in which the entropy and enthalpy components are of opposite sign, changes in the relative magnitudes of these components could cause the "very nature of the allosteric response" to be reversed (Johnson and Reinhart 1997, 12820).

With these variable temperature experiments, Reinhart and his colleagues exploited the thermodynamic relationships to highlight the importance of entropy in determining the nature of allosteric coupling. However, the results of these experiments themselves were unsurprising. They simply showed how the experimental manipulation of temperature affects the entropy of the system, which was already known from the Gibbs free energy equation. Moreover, the researchers gave no reason to think that these elevated temperatures were physiologically relevant to the allosteric systems they were characterizing. In light of these facts, it seems the primary motivation for these variable temperature experiments was to transform less compelling evidence for the dynamic view into more compelling evidence. At room temperature, entropy only decreased the magnitude of allosteric activation of PFK by MgATP, which was still controlled by the enthalpy term. But, at higher temperatures, the compensatory effect of entropy was enhanced: entropy now controlled the nature of the allosteric coupling as well as the magnitude, thereby transforming MgATP from an activator to an inhibitor (Johnson and Reinhart 1997). With this latter result, the

researchers were better able to drive home their message that the entropic contribution to allosteric coupling could be quite significant and thus ought not be neglected.

Reinhart's research program in the 1980s and 1990s clearly shows how scientists already convinced by theoretical arguments for the dynamic view of proteins set out to confirm it empirically. His research program was built upon the premise that protein dynamics were an important driver of protein function, and his work on allostery aimed to provide empirical proof of it. Rejecting structural models of allostery that considered only the enthalpic contribution, Reinhart showed that entropic changes could play a major role in allosteric coupling. Guided by the thermodynamic framework, he knew where to look and how to identify protein-ligand systems that exhibited allosteric coupling driven by changes in dynamics rather than structure. Indeed, this energetic analysis of allostery even showed him how to manipulate allosteric systems to increase the entropic contribution to coupling free energy. Throughout his publications during this time, he cited his findings as evidence that the static view of proteins and structural models of allostery needed to be replaced by the dynamic view and dynamic explanations of allostery that considered both enthalpic as well as entropic contributions to allosteric coupling.

Cooper and Dryden were also members of the community of dynamicists whose research centered around the pursuit of empirical evidence for the dynamic view. When Dryden joined Cooper's lab in October 1983, he and Cooper immediately started thinking about topics for a theoretical paper to demonstrate the functional importance of protein dynamics. In their discussions, Cooper's previous suggestion of allostery without conformational change came to mind. In 1980, Cooper had considered the possibility of dynamic rather than structural changes transmitting a signal from one site to another site on a protein in a relatively obscure paper in *Science Progress*. At that time, he concluded that "it [was] difficult at this stage to predict whether

dynamic changes can contribute significantly to allosteric effects" (Cooper 1980, 495). In the fall of 1983, Cooper and Dryden decided to "put some numbers on this very basic idea" and develop a formal model of dynamic allostery.[52] The goal was to produce a mathematical model that would enable them to estimate the potential magnitudes of the allosteric effects that arise from dynamics alone, which would in turn help determine the plausibility of allostery driven by dynamic changes in the absence of conformational change.

Dryden thus began his graduate studies doing theoretical work, developing a formal model of dynamic allostery. His contribution to the model published in 1984 focused on allostery caused by changes in vibrational energy. His model showed that allosteric coupling could occur in the absence of any structural changes in the protein if the binding of the allosteric ligand induced changes in its vibrational spectrum. Dryden modeled two cases in which changes in low frequency normal modes could produce allosteric coupling. In the first case, ligand binding causes many "small frequency shifts resulting from an overall stiffening of the protein structure" (Cooper and Dryden 1984, 106). The shift in many individual vibrational modes from lower to higher frequency makes the protein less dynamic, since higher frequency modes are less thermally excitable. This in turn facilitates binding of the second ligand, since the system has already paid most of the entropic cost associated with ligand binding. In the second case, ligand binding suppresses "thermally excited collective modes" that are "strongly coupled to ligand binding sites" (Ibid.). If these global vibrational modes—the so-called "breathing modes"—were suppressed by ligand binding, it would lead to positive cooperativity between the binding sites. Here, binding of the ligand again shifts the frequency of the collective modes to higher, less thermally-excitable

[52] Cooper interview, 27 May 2019.

frequencies, but unlike the first case, the shift in frequency disrupts the collective mode that formed the energetic linkage between the two binding sites.

Alongside this general theoretical treatment, Dryden also developed a more concrete model of dynamic allostery. In a recent interview, he confessed that initially he "didn't want to do a general theory" of dynamic allostery but instead "wanted to get a specific physical model that would prove that binding of one ligand would change the vibrational spectrum of this physical object and enhance the binding of the second ligand."[53] Hence, he developed the "scissors model." The model was meant to represent a highly idealized enzyme that could serve as a mechanistic example of how changes in low frequency vibrational modes could lead to allosteric coupling without any change in the average protein structure. In a recent interview, Dryden described the generation of the model:

> I came up with a little sketch…of the scissors model…. If you imagine a pair of scissors and that they can sort of vibrate…as one side opens, the other closes. And imagine if a ligand goes in [and] is jammed in the jaws of the scissors, then the scissors can no longer close…And so the second ligand would fit on the other side in between the handles.[54]

Starting from this idea, he developed a simple theoretical model of the enzyme that he could use to evaluate the potential allosteric coupling between the two binding sites. He approached the problem like a physicist building a toy model:

> I'm going to make a theoretical pair of scissors. [In] the model, there's going to be two rods, and they'll be joined at a rubber pin. And they'll have a certain length, and the rods won't have any mass. All the mass of the protein was going to be in point masses at the ends of the four ends of the system… And then I thought, "Well, I'd be able to solve the equations of motion for that physical model, and that would give me the vibrational modes."[55]

---

[53] Dryden interview, 24 June 2019.
[54] Ibid.
[55] Ibid.

According to Dryden, Cooper ultimately decided not to include the scissors model in the 1984 publication, likely because he "thought it was too detailed" and "wouldn't be taken seriously," since it was not supported by any known empirical systems.[56] But Dryden continued to work on the model and his analysis of its vibrational spectrum, and he included this work in his dissertation (Dryden 1988).

For Dryden, the scissors model provided a possible mechanism that showed how changes in vibrational normal modes, rather than a conformational change, could cause allosteric coupling between two sites:

> One can imagine the free enzyme vibrating at low frequency with a large amplitude due to the highly excited states, encountering the ligand which binds without changing the mean conformation, as measured by the angle θ, but which restricts the amplitude of the scissor motion severely. (Dryden 1988, 66)

Unlike the general theoretical treatment of this topic included in the 1984 paper, the scissors model offered better guidance on where to look for empirical confirmation of dynamic allostery. This feature of the model was important to Dryden since the primary goal for "the rest of [his] PhD then was…trying to prove it…trying to prove the scissors model" by obtaining data "that could show changes in low frequency vibrational modes in proteins upon ligand binding."[57]

The arc of Dryden's doctoral research in Cooper's lab thus exemplifies the approach of the true believers in the 1980s and 1990s. He began with a strong theoretical commitment to the dynamic view of proteins, developing a theoretical model that demonstrated the functional importance of protein dynamics. Then, guided by this model, he went out in search of empirical

[56] Ibid.
[57] Ibid.

confirmation for it. As Dryden summed it up recently, his research was "theory first, then experiments."[58]

However, when he turned from the theoretical model to empirical research he ran into problems. From empirical and simulation work published at that time, he learned that proteins were likely to have hundreds or even thousands of low frequency vibrational modes. Thus, the likelihood that he would be able to identify the shift in low frequency modes that he had calculated for the scissors model was slim, since the background vibrational spectra for the rest of the protein would likely swamp the changes of interest. In order to confirm the model, he realized he would need to try to identify a simpler allosteric system.

While scouring the literature, Dryden came across two papers by Julius Rebek, a synthetic organic chemist at the University of Pittsburgh who had been developing synthetic allosteric compounds (Rebek 1984, Rebek et al. 1985). The crown ether Rebek's group had synthesized was much smaller than any polypeptide molecules and therefore would have many fewer vibrational modes. Dryden thought he might have better luck confirming the change in vibrational spectra with this compound, so he and Cooper wrote to Rebek, who graciously sent them about 10mg of the compound. Rebek's group had NMR and x-ray crystallographic data for the compound at different temperatures, but they had not done Raman spectroscopy to analyze changes in the vibrational spectra upon ligand binding. Dryden aimed to obtain this data to confirm the scissors model. After much troubleshooting, he was able to identify a shift in the low frequency region of the spectra. However, the initial spectra he obtained were quite broad, and the spectrometer broke before he had a chance to refine his technique to produce sharper peaks. Therefore, the results were

[58] Ibid.

never published, although Dryden did include them in his dissertation on "Functional Consequences of Protein Dynamics" (Dryden 1988).

This account of Dryden's theoretical and empirical research, coupled with the description of Reinhart's research program and publications, provides a glimpse of what was happening within the small community of true believers in the 1980s and 1990s. They were early adopters and promoters of the dynamic view at a time when most mainstream structural and molecular biologists were still committed to the static view of proteins and were focused on developing structural explanations of protein functions. Committed to the dynamic view for theoretical reasons, the true believers set out to find empirical examples that could confirm their view while disconfirming the static view. My accounts of both Dryden and Reinhart show their tenacity and the lengths they would go in their quest to find evidence in support of their theoretical commitments. When studying proteins proved too challenging, Dryden searched for and found a synthetic compound in order to confirm his model, and Reinhart went so far as to manipulate his experimental setup, raising the temperature, in order to increase the importance of protein dynamics in allosteric coupling.

Because they knew where to look and were determined to find evidence in support of the dynamic view, these researchers slowly but surely discovered empirical examples of protein functions driven by changes in dynamics throughout the late 1980s and 1990s.[59] The discovery of "anomalous" cases of dynamic allostery was therefore driven by the theoretical commitments of these true believers as well as the workhorse dynamicists. It was their belief in the application of the principles of thermodynamics to proteins and the analysis of proteins as small thermodynamic

[59] For empirical examples of dynamic allostery in a variety of systems discovered during this time, see Cooper et al. (1989), Akke et al. (1993), Braxton et al. (1994), and Johnson and Reinhart (1994a, 1994b, 1997).

systems that led them to adopt techniques that could identify the dynamic features of proteins, in general, and also led them to use and develop analytical methods that could distinguish the structural and dynamic contributions to allosteric coupling, specifically. While the true believers would be more likely to frame their findings as anomalies that could not be explained by structural models, both sets of dynamicists were accumulating observations that were difficult, if not impossible, for the dominant structural accounts of proteins to explain. The accumulation of dynamic anomalies by the early 2000s proved sufficient to convince protein scientists, including crystallographers and others committed to the static view, to take protein dynamics seriously. By that time, the dynamic behavior of proteins could no longer be written off as inconsequential "noise," but instead had to be acknowledged as an important driver of protein function.

3.5 Conclusions: Widespread Acceptance of the Dynamic View

By the mid-2000s, the mounting empirical evidence for protein dynamics, coupled with the clarion calls of the true believers to replace structural models with dynamic ones, finally caused the larger community of protein scientists and structural biologists to take note. Through the work of scientists, such as Cooper, Dryden, Reinhart, Akke, and others, many examples of protein dynamics had been discovered—too many, in fact, for them to be ignored or written off as rare. Building on this foundational work from the 1980s and 1990s, scientists such as Vincent Hilser at Johns Hopkins and Ruth Nussinov at the National Cancer Institute developed well-funded research programs around protein dynamics. Influential representatives of a new wave of true believers, they continued to promulgate dynamic accounts of protein function in the 2010s, helping the

dynamic view achieve widespread acceptance by the broader community of protein scientists and structural biologists.[60]

This history of the dynamic view of proteins differs from the standard Kuhnian account of anomaly-driven theory change, an account which has been endorsed by some of the scientists involved.[61] In this case, I have argued that it was the theoretical commitment to the dynamic view of proteins that drove the discovery of anomalies. Cooper and Dryden's model, along with the thermodynamic framework more generally, provided guidance to the dynamicists, by shedding light on where to look for cases of dynamic protein functions. Although these researchers slowly accumulated anomalies that chipped away at the dominant static view, these anomalies did not emerge during the course of Kuhnian "normal science" (Kuhn 1962). Instead, the causal arrow is reversed. What my historical analysis has shown is that these empirical examples of dynamic protein function were only discovered and catalogued in the 1980s and 1990s because a certain subset of scientists applied the principles of thermodynamics to proteins. Thus, a deep theoretical commitment to protein dynamics lies at the heart of this history, providing the impetus for a larger theoretical and explanatory shift from the static to the dynamic view of proteins.

[60] In the span of four years, Hilser published high-impact reviews on dynamic allostery in *Science* (Hilser 2010), *Nature* (Motlagh et al. 2014), and the *Annual Review of Biophysics* (Hilser et al. 2012). During this time, Nussinov and her colleagues also published multiple reviews, perspectives, and opinion pieces advancing dynamic accounts of allostery (Tsai et al. 2009, Tsai and Nussinov 2014, Nussinov and Tsai 2015, Lui and Nussinov 2016).

[61] Hilser implicitly endorses a Kuhnian account of the development of the dynamic view of proteins when calls the switch from the static to dynamic view a "paradigm shift" and cites the discovery of anomalies as the driver behind this change (Hilser et al. 2012, Motlagh et al. 2014).

4.0 Protein Concepts, Representations, and Their Epistemic Relations

4.1 Introduction

Within the molecular life sciences, many disciplines, including molecular and structural biology, biochemistry, biophysics, enzymology, and molecular genetics, count proteins as among their objects of study. Representations of proteins figure in explanations, such as mechanistic explanations, either as part of the explanans (e.g., enzymes are cited in explanations of metabolism and antibodies are cited in explanations of immunity) or as the explanandum (e.g., studies of enzyme catalysis or protein-ligand binding seek to explain the behavior of proteins). Proteins also serve as sites of intervention in both laboratory techniques (e.g., PCR and CRISPR use proteins to replicate or edit DNA sequences) and medical treatments (e.g., HIV enzymes and other proteins function as drug targets for antiretrovirals). Over the past century, as the importance of proteins in a myriad of biological processes has become clearer, the concept of the protein has evolved significantly.

This chapter focuses on the two most recent concepts of the protein that have emerged since the birth of molecular biology: the structural concept (protein_S) and the dynamic concept (protein_D).[62] The former was the dominant protein concept of classical molecular biology in the 1950s and 1960s, whereas the latter first emerged in the 1970s and 1980s. In previous chapters, I have considered the historical relationship between these two concepts, but in this chapter, I turn my attention to the epistemic relations between them. Specifically, I aim to characterize the

[62] This notation is my own.

epistemic relationships that obtain between these two concepts and their associated representations. A major innovation of my philosophical approach to this topic is the identification of two levels of analysis. I argue that we must distinguish the epistemic relationship that obtains between the two protein concepts from the epistemic relationships that obtain among their associated representations. Although no philosophers have undertaken a similarly systematic study of protein science,[63] my analysis draws on two bodies of work by philosophers of biology. It is guided, on the one hand, by philosophers who have scrutinized the relationship between successive concepts of the gene in the twentieth century, and on the other hand, by the work of Sandra Mitchell, who has considered the epistemic relationships that obtain between different models and representations of proteins.

In this chapter, I argue that conceptual replacement, rather than reduction or integration, is occurring in protein science. The conceptual core of the structural concept includes beliefs about proteins that privilege structure over dynamics that are rejected by the dynamic concept. The inferential role of the dynamic concept replaces those beliefs with ones that recognize structure and dynamics as co-determinants of protein function. On my view, this revision in the inferential role of the concept is too drastic to count as reduction and instead constitutes conceptual replacement. At the level of representations, however, I argue that dynamic and structural representations of proteins are typically related via abstraction. Because protein dynamics and structure are largely partitionable features of proteins, the inclusion or exclusion of one of them in a representation has little effect on the other. Therefore, scientists can often transition between structural and dynamic representations of proteins via processes of abstraction or de-idealization.[64]

---

[63] Throughout the chapter, I will use the terms "protein science" and "structural and molecular biology" as shorthand to refer to those areas within the molecular life sciences that consider proteins to be one of their objects of study.
[64] I will clarify what I mean by these terms (e.g., abstraction and de-idealization) below, in Sec. 4.4.

In what follows, I begin my analysis by characterizing the two concepts of the protein. Following Philip Kitcher's (1977, 1982) and Ingo Brigandt's (2010) accounts of conceptual change, I distinguish the structural and dynamic concepts of the protein by characterizing their references, inferential roles (or meanings), and their associated epistemic goals in Section 2. I show how the differing beliefs about proteins associated with each concept guide scientists toward different uses of the concept in inference, explanation, and scientific practice. In Section 3, I consider the applicability of certain philosophical accounts that aim to understand scientific change, and I conclude that neither reduction nor the Kuhnian account accurately characterizes the relationship between protein concepts. In Section 4, I develop my own account of the relationships between protein concepts and representations and attempt to draw certain philosophical morals from my analysis of concepts and representations in protein science.

4.2 Protein Concepts and Representations

In this section, I first present an account of concepts and conceptual change (Sec. 4.2.1). I then use this account to identify and distinguish the key semantic features of both the structural (Sec. 4.2.2) and dynamic protein concepts (Sec 4.2.3). I also discuss some of the methods and representations associated with both concepts.

4.2.1 Concepts, Representations, and Conceptual Change

I endorse an account of concepts and conceptual change in science, which is largely drawn from Kitcher (1978, 1982) and extended by Brigandt (2010). These philosophers of biology

developed their accounts to make sense of scientific change and demonstrated their usefulness by applying them to the history of the gene concept. Their accounts agree that concepts have at least two components of semantic content: reference and something akin to meaning. Both commit to an extensionalist account of reference. Thus, conceptual change can be understood as "shifts in referential relations between words and the world" (Kitcher 1982, 339). They use different vocabularies to cash out the meaning of a concept: Kitcher prefers "reference potentials" while Brigandt uses "inferential roles," but these terms are quite similar. On both accounts, the meaning of a concept is more than its theoretical definition.

Kitcher develops the notion of "reference potential" to help understand the use and changes in use of scientific concepts over time. He claims that the reference potential for a scientific term "is a compendium of the ways in which the referents of tokens of the term are fixed for members of the community" (1982, 340). On this view, the referent of a token is fixed by the explanation for the speaker's production of that token. Similar to other causal theories of reference, explanations of this sort trace the reference through a chain of events back to an "initiating event," in which someone was "either in causal contact with an entity or intend[ed] to refer to whatever satisfy[ed] a description" (Ibid., 345). According to Kitcher, the reference of a scientific term can be fixed by more than one initiating event. Whereas other semantic theories of reference require the reference to be fixed once and for all by the original baptism of the term, Kitcher's account liberalizes the idea of reference-fixing to better accord with the usage of scientific terms. He argues that the connections between scientific concepts and the world "are frequently renewed and extended" (Ibid.). Therefore, the reference potential of a scientific term—i.e., the collection of initiating events for the production of a token of the term—is likely to include many distinct initiating events. Kitcher argues that scientific terms "frequently have heterogeneous reference

potentials" because scientists find it useful to initiate their tokens by different events (Ibid.). He cites the production of chemical substances via different methods as an example, arguing that scientists in different disciplines can employ different modes of production to fix reference.[65]

Brigandt (2010) discusses the meaning of a concept in terms of its "inferential role" rather than reference potential.[66] The inferential role of a concept, according to Brigandt, is "a subset of the beliefs that scientists have about a term's referent" (Ibid., 22). Borrowing from inferential, or conceptual, role semantics, his account aims to highlight how the content of a concept "figures in inference, or more broadly how it figures in reasoning, perception, and action" (Ibid.). More precisely, he claims that the inferential role of a scientific concept consists of the "inferences and explanations in which the term occurs" that "account for the term's successful use in scientific practice" (Ibid.).

This focus on the use of a concept in scientific practice leads Brigandt to move beyond Kitcher's original formulation. He posits a third component of semantic content for scientific concepts, which he calls "the *epistemic goal pursued by a concept's use*" (Ibid., 23, emphasis in original). With this third component, Brigandt establishes a method to judge the rationality of conceptual change in science. For central terms in biology, such as the gene, he claims that we can often assign general epistemic goals that motivate their use by a community of scientists at a given time. We can then evaluate various concepts as to how well they enable their users to achieve these

---

[65] According to Kitcher, scientific concepts with heterogenous reference potentials are theory-laden, since the scientific community must have some theoretical description that enables them to link tokens with distinct modes of production to the same referent.

[66] Brigandt (2010) claims that his "inferential role" is similar to Kitcher's "reference potential": "*Inferential role* is my term for what has been called a term's meaning (sense, intension), and is fairly similar to what Kitcher dubs reference potential" (22, emphasis in original). I am skeptical of this claim, since Kitcher does not seem to endorse a theory of meaning, but instead focuses on concept usage. For my purposes, I will rely primarily on Brigandt's account and his discussion of a concept's meaning, or inferential role.

epistemic goals. Importantly, he argues that his innovation enables us to judge the rationality of conceptual change in science, even when the reference and inferential role of a concept changes.[67]

Both Kitcher and Brigandt develop their accounts to make sense of conceptual change in science, in part to respond to the challenge of incommensurability posed by Kuhn and Feyerabend, and also to clarify the debate about gene concepts. Conceptual change, on both their accounts, involves more than a change in definition. It also requires a change in the reference potential or inferential role of the concept. Conceptual change occurs as part of the development of science when new discoveries and hypotheses alter the set of initiating events in the reference potential (e.g., through the removal of inadequate modes of reference or the acquisition of new ones), or when a concept's inferential role expands as the concept begins to figure in new inferences or explanations. Both accounts also acknowledge that the referent for the concept can change along with the meaning (as almost everyone agrees happened with the gene concept), but this change in reference need not lead to incommensurability (Kitcher 1978) nor make conceptual change in science irrational (Brigandt 2010).

Furthermore, Brigandt (2010) argues that because we have access to a more fine-grained analysis of conceptual change, we can largely sidestep the issue of concept individuation. He claims that

> labeling two contexts in which a term is used as expressing the "same concept" or two "different concepts" is not of much philosophical relevance. Instead, what is important is to lay out how exactly two uses of a term differ regarding the various semantic properties. (Ibid., 25)

I agree. I will follow the scientific actors when they talk of a "new" dynamic concept of proteins, but whenever I use the terms "dynamic concept" ("protein$_D$") or "structural concept" ("protein$_S$")

[67] He further argues that his account can show the rationality of conceptual change even in cases in which the epistemic goals change. However, I will not discuss this further since it is not be relevant to the change in protein concepts.

this should be taken as shorthand for referring to the concept—with its referent, inferential role, and epistemic goal—as it is or was used in a particular community of scientists at a particular time. Little hangs on whether we construe the dynamic concept as a new concept or just a modification of the earlier structural concept, since the differences in their reference, inferential roles, and epistemic goals will serve to distinguish them.

For my analysis of protein concepts, I must append a brief note on the relationship between concepts and representations. Neither Kitcher nor Brigandt discusses the latter in their accounts of concepts, but other philosophers of biology have posited an explicit link between gene concepts and representations. Griffiths and Stotz (2013), for instance, begin and end their book-long exploration of gene concepts with the following claim: "The concept of the gene is therefore best conceived as a set of contextually activated representations" (8, see also 221). Their identification of concepts with sets of representations is compatible with the account of concepts and conceptual change I have endorsed, so long as we include linguistic and mathematical representations, as well as visual representations or models of proteins. A given concept will have associated visual representations, but will also include linguistic representations, expressing explanations, inferences, and hypotheses. We can think of the relationship between concepts and representations as mereological: a given concept will include a characteristic set of representations. Hence, the concept of DNA will support various representations, such as a linear sequence of ATGC base pairs in a DNA primer, Crick's famous Central Dogma diagram, Franklin's x-ray diffraction pattern, as well as the two-story model of DNA at the Virginia Science Museum. And two different scientific concepts will support different sets of representations. Conceptual change will also be reflected in the different representations used within the inferential role of a particular concept. The reason for highlighting this distinction will become clearer in Section 4, where I will argue

that the epistemic relation that obtains between protein concepts is quite different than the one that holds between different representations of proteins.

4.2.2 The Structural Concept and Its Representations

4.2.2.1 The Concept

The structural protein concept (protein$_S$) emerged from classical molecular biology in the 1950s and 1960s. According to the theoretical definition of protein$_S$, a protein was a rigid and static molecule that was composed of densely-packed, highly-charged amino acid chains (Fruton 1999, Tanford and Reynolds 2001).[68] Moreover, every protein molecule was thought to fold into a unique stable 3D conformation such that "each molecule, of a given kind, is identical in shape to all others" (Cooper 1980, 474). The referent of the concept, which was established via different methods and initiating events, was an individual molecule of this sort. The inferential role, or meaning, of protein$_S$ included what Sarkar has called the "structure-function rule," which held that "the behavior of biological macromolecules can be explained from their structure as determined by techniques such as crystallography" (Sarkar 2008, 60-61). A review of protein structure from 1956 applies this structure-function rule to proteins:

> The biological activities of proteins such as catalysis and antigenicity are dependent to a large degree on the complete structural integrity of the molecule and the organized participation of the more-or-less complete peptide structure. (Anfinsen and Redfield 1956, 72)

[68] The concept of protein includes many more features within its inferential role. Throughout this chapter, however, I will focus only on those aspects of the structural and dynamic concept's inferential roles that distinguish them from each other.

Classical molecular biology focused on protein structure precisely because of this hypothesized connection to function.

The commitment to explaining protein function by citing the 3D structure of the protein and the molecules with which it interacts was seen as a solution to the explanatory problem posed by biological specificity. Since the late nineteenth century, organic chemists, such as Emil Fischer had demonstrated that enzymes were designed to interact with only certain substrates. For instance, Fischer's work demonstrated that a yeast enzyme would hydrolyze the D form of a sugar but not the L form, which differed only in the transposition of one hydrogen (H) and one hydroxyl group (OH) on a backbone carbon. To explain this specificity, Fischer introduced the lock-and-key model: "To use a metaphor, I would like to say that the enzyme and glucoside [substrate] have to fit together like lock and key in order to exert a chemical effect on each other."[69] From this and other work in biochemistry and immunology, classical molecular biology "inherited the proposed mechanism that the function or behavior of biological molecules is 'determined' by its structure" (Sarkar 2008, 60).

When Fischer introduced his model linking enzyme structure to its catalytic activity in the 1890s, the hypothesis that protein structure determines protein function had little direct empirical evidence. By the 1960s, however, things had changed. The elucidation of protein structures via crystallography by John Kendrew (myoglobin) and Max Perutz (hemoglobin) in the 1950s and 1960s, coupled with structural modeling work by Linus Pauling and others, bolstered the structural concept and "seemed to confirm the hypothesis that structure explains behavior" (Sarkar 2008, 60).[70]

---

[69] Fischer (1894, 2992) translated in Barnett and Lichtenthaler (2001, 377).
[70] See Green et al. (1954), Kendrew et al. (1958), Perutz et al. (1960), and Pauling et al. (1951).

4.2.2.2 Structural Methods and the Epistemic Goals of Proteins

The structural concept of the protein, coupled with the structure-function rule, formed "the theoretical core of molecular biology" (Sarkar 2008, 60). But the inferential role of the concept includes more than just the structure-function rule. The concept was also embedded within the new research tradition of classical molecular biology, with its methodology and experimental practices. It provided a research heuristic, directing scientists interested in protein function to study their 3D structures, since "what was important in a protein was its shape" (Morange 2006, 515). The concept also laid out the foundation for adequate molecular explanations of protein behavior. To explain binding, catalysis, allostery, or other protein functions, one needed to cite the structural features of the protein in order to show how the molecular surface of the protein complemented the surface of the substrate or ligand. The structural concept of the protein was thus closely tied to a particular set of research methods and explanatory strategies.

The primary epistemic goal of proteins for the community of molecular biologists was to explain protein behavior and function, both for the sake of basic scientific understanding as well as intervention. Because of the commitment to the structure-function rule, this goal was to be achieved by elucidating the 3D structure of biological macromolecules, such as proteins. For this reason, techniques such as x-ray crystallography, which could reveal "the rigidly defined structural relationships which characterize the native protein," were highly valued (Anfinsen and Redfield 1956, 80). Although structural features could be inferred from other methods (e.g., activity assays and various spectroscopic analysis), in the early days of molecular biology, the only technique able to determine the structure of entire proteins was x-ray crystallography. Nuclear magnetic resonance (NMR) was also developed for protein structural determination, with the first NMR protein structure published in 1985 by Kurt Wuthrich (Williamson et al. 1985). However, NMR

was only practical for relatively small proteins and polypeptides. X-ray crystallography, in contrast, had no in principle size limit. This difference alone was often considered to be "decisive in favor of x-ray diffraction" (Drenth 1994, vii). Thus, x-ray crystallography was typically taken to be the gold standard for protein structural determination, and the belief that the development of x-ray diffraction and structural refinement techniques would ultimately yield more and higher-resolution structures was widespread beginning in the 1960s.

This optimism and excitement surrounding advances in x-ray crystallography and the steady publication of new protein x-ray crystal structures stemmed from the belief that these 3D protein structures would serve as the basis for explanations of protein functions and cellular processes. Early textbooks, such as the *Principles of Protein Structure*, as well as the research articles they summarized, provide examples of scientists citing the structural features of the enzyme active site and the location of certain chemical groups on the substrate to explain "the structural basis of protein mechanism, action, and function" (Schulz and Schimmel 1979, 233). The structural concept was thus deeply wedded to a research tradition dominated by x-ray crystallography and focused on structural determination, as well as an explanatory strategy that cited structure as the mechanistic basis for protein behavior and function. The ability for scientists to infer mechanistic features of enzyme catalysis or ligand binding from 3D protein structures further reinforced their commitment to the structure-function rule and the research strategies that developed around the structural concept.

4.2.2.3 Structural Representations

The information obtained via x-ray crystallography and other techniques led researchers to develop a characteristic set of modeling practices to represent protein structure in accordance with the structural concept of the protein (de Chadarevian 2018). X-ray diffraction patterns are the direct

output of an x-ray crystallographic study. For proteins, these patterns were complex and nearly impossible to interpret until Perutz developed a method using heavy atoms that enabled researchers to "solve the complex X-ray pattern of a protein crystal and produce a model of the structure of the molecule" (Kendrew 1961, 104). Using Perutz's method, researchers were able to convert the diffraction pattern into an electron density map, which then allowed them to overlay the polypeptide backbone structure of the protein in order to visualize the 3D shape of the molecule. If the resolution of the x-ray diffraction pattern was sufficiently high or additional information were available from other techniques (e.g., sequence information), crystallographers could add information about the position and orientation of amino acid side chains to these backbone structures.

In the 1950s and 1960s, Kendrew and his colleagues working on myoglobin produced some of the earliest 3D representations of proteins developed from the results of x-ray crystallography. In 1957, they designed the first backbone model, deemed the "sausage model" (Fig. 6). This relatively low-resolution structural representation offered a "good general picture of the layout of the molecule" (Kendrew 1961, 106). Kendrew then developed an atomic model of myoglobin, which represented amino acid sidechains along the backbone structure. This "forest of rods" model, which was constructed from six-foot-tall steel rods, used the rods to fix the atoms of the protein molecule in space, and the connectivity of bonded atoms was depicted via wire connections. Architect and illustrator Irving Geis was commissioned in 1961 by *Scientific American* to create a 2D drawing of this model, without the rods and other mechanical parts, for publication in the magazine. His work as a molecular illustrator, which began in 1961, "deeply influenced the conventions for depicting and viewing intricate protein structures" (Ibid., p. 148).

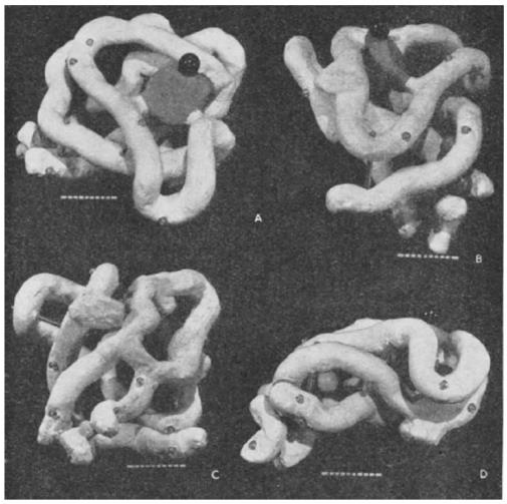

Figure 6 Model of Myoglobin (Kendrew et al. 1958). First model of myglobin to appear in print. Polypeptide backbone depicted in white, and the heme group depicted as a gray disk. Republished with permission from Kendrew et al. (1958); permission conveyed through Copyright Clearance Center, Inc.

Within a few decades, many types of representation became standardized while new representations associated with the structural concept were also developed, and this set of representations formed an important part of the inferential role of the concept. Schultz and Schirmer (1979) surveyed the various types of representations of proteins in the *Principles of Protein Structure*. The authors highlight the variety of 3D and 2D models and identify the virtues of each. Two decades after Kendrew built the first "forest of rods" model, Schulz and Schirmer claimed that these wire models still offered "the most accurate three-dimensional representation of a protein" (135). Backbone wire models, in contrast, were easy to build and portable, but only provided a "chain fold representation of low accuracy" (Ibid.) (Fig. 7A). Plastic push-fit models, the bane of organic chemistry students to this day, could also be used to represent the structural features of proteins. Colored balls represent atoms with different properties, but the "balls and sticks are large with respect to scale," compared to wire; thus, a push-fit model "lacks transparency" (Ibid., 137). Space filling models are even less transparent. According to the authors, these models were "very difficult to build" (in 1979) and did not depict the interior of the molecule.

However, they illustrated surface properties very well, and hence were used mainly "to depict substrate and effector sites" (Ibid., 139) (Fig. 7B).[71]

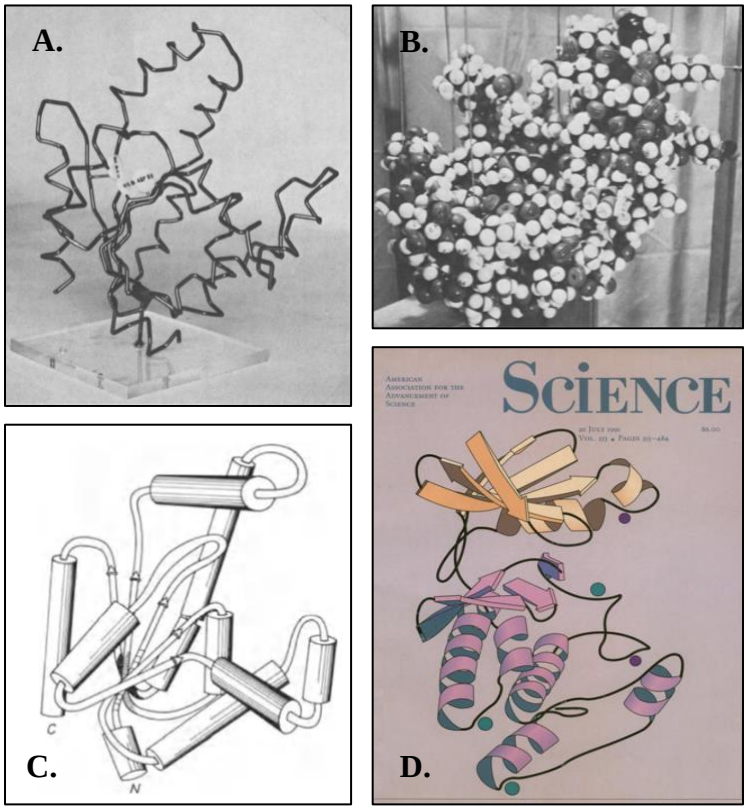

Figure 7 Static Representations of Proteins. (A) Photograph of 3D backbone wire model of adenylate kinase. (B) Photograph of 3D space-filling model of ribonuclease S. (C) 2D carton model of adenylate kinase. (A-C) Republished with permission from Schulz and Shirmir (1979); permission conveyed through Copyright Clearance Center, Inc. (D) Cover of *Science* magazine showing 2D structural cartoon, or "ribbon diagram," of catalytic core of eukaryotic kinases. *Science*, July 26, 1991, Vol. 253, Is. 5018. Reprinted with permission from AAAS.

The structural properties of proteins gleaned from x-ray crystallography and other techniques could also be represented in 2D models. These could be drawings, such as Geis' drawing of myoglobin, with its "introduction of small distortions and exaggerations to resolve

[71] The ability to reveal surface features would have been quite important, since in addition to the structure-function rule, classic molecular biology was also committed to the "molecular-shape rule" (i.e., "structures…could be characterized entirely by molecular size and, especially, external shape") as well as the "lock-and-key fit rule" (i.e., "molecules interact only when there is a lock-and-key fit between two molecular surfaces") (Sarkar 2008, 61). Space-filling models of proteins would be useful for assessing the shape and fit of proteins and their substrates and ligands.

overlaps and create an understandable image" (de Chadarevian 2002, 148), or they could be "exact line drawings" done by hand or with a computer (Schulz and Schirmer 1979, 142). In addition, "structure cartoon" models could represent the "essential parts" of the protein structure, usually by representing alpha-helices as cylinders and beta-sheets as arrows (Ibid., 143; Fig. 7C). They depict the structure but "relieve the reader of unnecessary details" (Ibid.). Although more sophisticated computer-generated models have replaced the hand-drawn 2D models of the 1970s and 1980s, similar structural and cartoon models of protein structure have remained common (Fig. 7D).

Despite their differences, all these representations of proteins are involved in the inferential role of the structural concept. Both Kendrew's "forest of rods" model from the 1950s and the cartoon structural model from the 1990s represent the fact that a protein, according to the structural concept, had a single stable structure that enabled it to carry out its characteristic function. Modern computing has altered the production of protein representations, with cheap and computationally efficient visualization programs, such as PyMOL, Jmol, and VMD.[72] But even these programs, which let a user display and modify structures deposited in the Protein Data Bank, represent the protein as having a fixed, stable structure.[73] Reasoning with these characteristic representations, as well as the structure-function rule, the scientists committed to the structural concept of the protein developed a productive research program in the second half of the twentieth century that

[72] The Protein Data Bank offers a lengthy list of the available protein visualization software on its website: https://www.rcsb.org/pages/thirdparty/molecular_graphics.

[73] These programs let users manipulate the way the structure is represented, for instance, by switching between ball-and-stick representation of atoms and bonds and a cartoon structure with rods and arrows, and they also let users rotate and flip the molecule. However, because they are generated from the PDB files deposited in the Protein Data Bank, each structure has fixed 3D coordinates for each of the atoms in the molecule. Thus, although the user can manipulate the structure, for example, by turning it to see the molecule from different angles, these computer-generated 3D models represent the structure itself as rigid and static.

centered on determining protein structures and using those structures in explanations of protein function.[74]

4.2.3 The Dynamic Concept and Its Representations

4.2.3.1 The Concept

The dynamic concept of the protein (protein$_D$) emerged in the 1970s as a critique of the structural concept. The community of scientists who championed protein$_D$ argued that the structural concept figured in faulty inferences and explanations. They aimed to drastically revise or replace the defective structural concept with a new concept, associated with new representations, hypotheses, and explanatory commitments. Writing in 1986, Martin Karplus and Andrew McCammon, two champions of protein$_D$, identified a fundamental problem with the structural concept of the protein:

> The study of how proteins serve the needs of a living organism is a curious case in which a method that yielded dramatic advances also led to a misconception. The method is X-ray crystallography…. The intrinsic beauty and the remarkable detail of the structures obtained from X-ray crystallography resulted in the view that proteins are rigid. This created the misconception, namely that the atoms in a protein are fixed in position. (Karplus and McCammon 1986, 42)

The structural concept, they claimed, had also led scientists to endorse a problematic explanatory strategy: "most attempts to explain protein function…have been based on the static structure," even though there are many "phenomena that cannot be explained by the static view" (Ibid.).

---

[74] The concept, with its associated representations, explanations, and methods, was successful enough to prompt Michel Morange, a molecular biologist-cum-historian, to claim that "for the molecular biologists of yesterday and today a protein is sufficiently understood only when its three-dimensional structure has been determined by x-ray diffraction" (Morange 1998 [1994], 215).

The dynamic concept of the protein arose from a commitment to treating proteins as small thermodynamic systems. Similar to the structural concept, one potential referent for the dynamic protein concept was an individual folded amino acid chain. (We will see below that the reference potential expanded with the dynamic concept so that it could also refer to a population of such molecules.) However, the theoretical definition of protein$_D$ privileged a thermodynamic rather than structural description of these molecules in solution. The individual protein molecule is a small system, and the electrostatic interactions that maintain the 3D structure are relatively weak. The constant collisions between the protein and the solvent—i.e., "the continual Brownian-motion-like buffeting by solvent molecules"—would supply enough thermal energy to cause the protein to be in constant motion (Cooper 1980, 489). Hence, according to the dynamic concept, a protein was "a fluid and flexible system" (Cooper 1976, 2740), which was "constantly changing the details of its conformation" (Karplus and McCammon 1986, 42).

On this view, an individual protein molecule would undergo constant structural fluctuations, rapidly interconverting among many quasi-degenerate microstates. This theoretical description of a protein molecule clearly distinguishes it from protein$_S$. The latter included the belief that all protein molecules of the same type would adopt the same 3D conformation. Indeed, it is this belief that enabled users of protein$_S$ to talk about *the* structure of a particular type of protein molecule. The dynamic concept denies that a population of protein molecules in solution would adopt a single discrete structure. As Gregorio Weber memorably describes it, the protein of the structural concept does not exist:

> The protein molecule model resulting from the X-ray crystallographic observations is a "platonic" protein, well removed in its perfection from the kicking and screaming "stochastic" molecule that we infer must exist in solution. (Weber 1975, 65)

In other words, individual protein molecules conduct a biased walk through conformation space, "wandering in a haphazard and non-periodic fashion amongst a multitude of possible conformational states" (Cooper and Dryden 1984, 107). Thus, the inferential role of the dynamic concept includes the belief that a population of protein molecules in solution exists in a multiplicity of different structures, or conformational microstates.

This set of theoretical commitments leads to a new potential referent for protein_D. Although the individual molecule will not have a stable structure in solution, according to the dynamic concept, the population, or ensemble, of molecules at equilibrium will be stably distributed across a range of microstate conformations. A probability distribution determines the likelihood of finding an individual protein molecule in any one of these microstates. The unit of function, in this case, is the ensemble of molecules, rather than any one molecule, and explanations of protein behaviors (e.g., catalysis or allostery) cite changes in the ensemble distribution, rather than changes in structure. If we are to preserve the notion that a protein has a function, then it makes sense to say that the ensemble of molecules is the protein.[75] We can interpret this as an expansion of the reference potential for protein_D (when compared to protein_S): depending upon the context, the reference for the token "protein" could be an ensemble of polypeptide molecules or a single dynamic molecule.

Adoption of the dynamic concept also leads to other changes in the inferential role of the concept, in particular the role it plays in explanations. The structure-function rule of classical

[75] This claim may seem more provocative than it actually is. Although no scientists (of whom I am aware) have advocated this change in the reference of the concept of protein, it falls out naturally if we take ensemble explanations seriously (see Chapter 5 for a detailed discussion of these explanations). If one wants to describe the unit that carries out a function, such as allostery, it is an ensemble of protein molecules rather than a single molecule. For example, the functional unit of hemoglobin responsible for transporting oxygen in red blood cells is not a given molecule in one or two structural states, but rather it is an entire population of hemoglobin molecules in an ensemble of microstate conformations. The change in reference here simply forces us to attend to the fact that a single molecule with a given structure cannot perform a function, since it can only adopt one structure at a time.

molecular biology has been jettisoned, and in its place, the dynamic concept includes the hypothesis that explaining function requires attending to dynamic properties of proteins as well as their structure. Many cases have been discovered that simply cannot be explained by the structure-function rule.[76] One particular striking case is the class of intrinsically disordered proteins that have allosteric, catalytic, or other functions. These proteins do not have a stable 3D structure, so their functions cannot arise from structural complementarity; thus, explanations that treat protein molecules as molecular machines have proven inadequate. The explanatory strategies associated with the dynamic concept require researchers to attend to both structural and dynamic contributions to function.

To avoid a potential confusion about the relationship between protein$_S$ and protein$_D$, a terminological clarification is in order. Although advocates of the dynamic concept have christened it "dynamic" so as to contrast with the earlier "structural" concept, this rhetorical choice does not mean that all structural features of protein molecules have been purged from the inferential role of the dynamic concept. Rather, the dynamic concept, and its associated inferences, explanations, and hypotheses, posits that protein molecules have important structural *and* dynamic features. This is evidenced by the fact that proponents of the dynamic concept often refer to the dynamic properties of the protein as "structural fluctuations," as well as their discussion of structural or conformational microstates of protein molecules in an ensemble. In an important way, protein$_D$ expands the meaning of the protein and its inferential role to include dynamic features, but it is more than just an expansion, since this concept also loses the theoretical core of protein$_S$—viz., the belief that proteins had stable 3D structures that could explain their functions. In other words, the dynamic concept displaces the privileged role of protein structure. With the dynamic

[76] See Motlagh et al. (2014) for some examples.

concept, researchers have had to develop new methods to investigate proteins, as well as new explanatory strategies to explain protein behavior and function.

4.2.3.2 Dynamic Methods and Epistemic Goals of Protein$_D$

Unlike the scientists who endorsed the structural concept, the community of scientists who adopted the dynamic concept did not use a single, dominant method. Instead, these scientists developed a diverse set of methods and experimental strategies to study protein dynamics, as well as the relationship between dynamics and structure. However, the advocates of protein$_D$ had the same epistemic goal as those who used protein$_S$: they aimed to use the concept to explain protein behavior and function. On their view, protein$_S$, with its associated hypotheses and explanatory strategies, was inadequate for this goal, since they believed that protein dynamics were important drivers of protein function.

Scientists adopting the dynamic concept of proteins have two general strategies with which to characterize these dynamic fluctuations. The first strategy characterizes the thermodynamic parameters of proteins to shed light on their dynamic properties. The standard thermodynamic approach uses the Gibbs equation, which describes the free energy (ΔG) of a given conformational state of the protein: $\Delta G = \Delta H - T\Delta S$. Changes in free energy are used to explain protein behaviors, such as binding. While the enthalpic contribution (ΔH) to free energy is thought to arise from changes in structural features of the protein, the entropic contribution (ΔS) is identified with changes in the internal dynamics of the protein. Although the focus of this analysis is on the dynamics of the system, the thermodynamic framework reinforces the fact that the behavior of proteins is governed by both structure *and* dynamics by identify them as the two components of free energy—ΔH and ΔS.

Early theoretical and experimental work sought to show that changes in dynamics alone could be the driver of protein functions, by causing changes in conformational entropy (Cooper and Dryden 1984). The conformational entropy of a protein is a measure of the number discrete conformational states that are accessible to it at a certain temperature. The number and relative accessibility of conformational states tracks the dynamic flexibility of the protein: a relatively rigid protein will have few thermally accessible conformations, whereas a relatively flexible or dynamic protein will be able to sample many different conformational states. Protein interactions that decrease the flexibility of the protein decrease the conformational entropy, making the interaction entropically unfavorable. Although thermodynamic approaches alone do not provide information about the set of thermodynamically accessible structures, they can provide information about the magnitude and relative importance of dynamics in protein interactions. For example, researchers can measure the heat capacity for a protein-ligand interaction via microcalorimetry, and from this value, they can distinguish the enthalpic and entropic contributions to free energy (Sturtevant 1977). The signs and relative magnitudes of these contributions reveal the extent to which protein dynamics, rather than structure, is responsible for particular protein behaviors.[77]

The thermodynamic approach cannot reveal much "about the structural character of the fluctuations nor about their timescale," but the second general strategy of investigating protein dynamics can (Brooks et al. 1988, p. 226). It uses diverse methods to characterize the range and diversity of structures available to the protein and to learn about the kinetics of these motions and the timescales in which they occur. The earliest experimental methods used to investigate the dynamic protein measured changes in spectroscopic properties or chemical reactivities of certain

[77] For example, see the discussion of the thermodynamic framework of the dynamic view, as well as the case of agonism-antagonism switching described in Chapter 3. This was one of the cases of anomalous dynamic behavior that could not be explained within the static view.

groups (Cooper 1980). Methods such as fluorescence quenching, or hydrogen-deuterium exchange, can identify regions of a protein that are more flexible than others, and both techniques can provide information about the timescale of protein dynamics. Some early x-ray crystallographic studies also provided evidence of protein dynamics. For instance, Frauenfelder used temperature-dependent x-ray diffraction to trap less probable, higher-energy conformational substates of myoglobin (Frauenfelder et al. 1979). By determining the 3D structure of these states, Frauenfelder and his colleagues revealed a subset of the ensemble of structures accessible to the protein in solution. Molecular dynamic (MD) simulations, which were pioneered in the 1970s, provide information about the timescale of various protein motions. They begin from known x-ray crystal structures of proteins and use Newtonian equations of motion to simulate the motions of the atoms over time (Brooks et al. 1988). Methods such as these were further able to characterize the heterogeneity of structures that a protein molecule could adopt.

Since the early days of the dynamic concept, technological advancements have afforded researchers new means to study protein motions. New empirical techniques that can provide direct, atomic-level evidence of protein dynamics have been developed. The most important among these techniques is protein NMR. Since the first protein structures were solved via NMR in the 1980s, the technique has been perfected to study larger proteins and polypeptides. Because NMR determines the structure of the protein in solution, it can provide information about the range of structures a protein can adopt in solution. Other techniques, such as small angle x-ray crystallography (SAXS), can be used to obtain similar information. In addition to these empirical methods, improvements in MD simulations, including vastly increased computing power, have enabled researchers to run longer simulations, which have offered dynamic information about slower protein motions. Other computational approaches have been developed to generate the

predicted ensemble of possible structures in solution, based upon the protein structure and thermodynamic and empirical data about the flexible regions of a protein (Liu et al. 2006). Both the empirical and computational advances provide evidence for internal protein motions and support the contention that a population of protein molecules inevitably exists as a multiplicity of structures, thereby reinforcing this posit, which plays an important part in the inferential role of the dynamic concept.

4.2.3.3 Dynamic Representations

Developing representations and models that comport with the dynamic concept of the protein has proven more challenging than developing structural representations, since most scientific representations occur in static media, such as journal articles, reviews, and textbooks. Nevertheless, the community of researchers who endorse $protein_D$ have developed conventions for representing the dynamic information gleaned from empirical studies and computational approaches. There are two general representational strategies, with significant overlap. The first attempts to represent the dynamic properties of a protein within a structural model of a *single* protein molecule, while the second strategy presents an *ensemble* of thermodynamically accessible protein structures in order to represent the structural fluctuations of the protein as well as the heterogeneity of a population of molecules in solution.[78]

The first strategy represents the protein as a superposition of structures, highlighting the motion within a single molecule. Models of this sort often begin from the average structure of the protein, as determined by x-ray crystallography, and then add information to represent the

[78] This distinction is only one of degree, since most representations can be construed either as a superposition of structures for a single molecule or as an ensemble of structures in a population. Both these features, however, form the core of the inferential role of the dynamic protein concept.

structural fluctuations caused by protein dynamics. These fluctuations are most often measured as deviations from the average structure in the root mean squared position (RMS) of atoms. The superposition strategy takes the average static structure and overlays its dynamic properties. Frauenfelder and colleagues (1979) offer an early example of this strategy for myoglobin (Fig. 8A). In this study, the researchers crystallized the protein in different conformational substates. They calculated the difference in RMS position of the backbone atoms from the average structure for all the crystalized substates. They then added a shaded region to the average structure to represent the atomic displacements that occur as the protein undergoes structural fluctuations in solution.

Karplus and McCammon (1986) produce a similar model of the dynamic properties of a molecule of myoglobin from a different method.[79] In this model, they have superimposed structural representations of the protein as it appeared over five-picosecond intervals during an MD simulation. The mismatch between the structures aims to represent the rapid motion of the protein backbone. A special issue of *Science* magazine on protein dynamics offers a more recent example of the superposition approach to modeling the dynamic concept of the protein (Fig. 8C). Here, the dynamics of calmodulin are represented as a superposition of computer-generated structures interpolated between two empirically solved structures of the protein. This model aims to represent the hypothesized structural fluctuations a single protein molecule undergoes in order to bind its substrate.

[79] Model not shown. See Karplus and McCammon (1986), "Motions of a Protein," p.43.

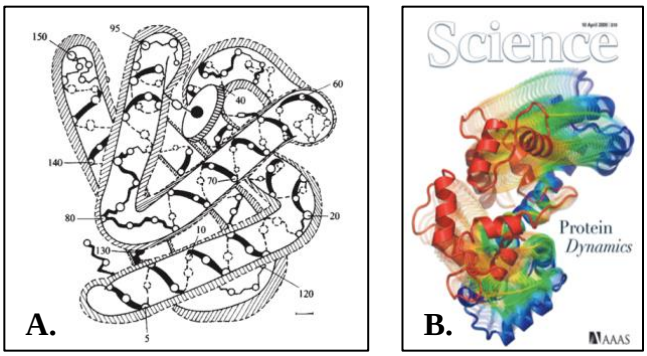

Figure 8 Dynamic Representation of Proteins: Superposition of Structures. (A) Flexibility of myoglobin. Solid lines represent the static structure, circles represent the backbone carbons, and the shaded region represents the extent of deviation from the static structure obtained by different substate conformations of the protein, with 99% probability. A 2Å scale bar is also shown. Republished with permission from Frauenfelder et al. (1979); permission conveyed through Copyright Clearance Center, Inc. (B) Dynamic structural fluctuations in calmodulin. Computer-generated structures interpolated between calcium-bound structure (blue) and structure bound to both calcium and one of the protein's regulatory targets (red). *Science*, April 10, 2009, Vol. 324, Is. 5924. Reprinted with permission from AAAS.

Shifting the focus from a single molecule to the population of protein molecules, the second strategy represents protein dynamics as an ensemble of distinct structures. The structures in the ensemble can be generated empirically or computationally. Figure 9 depicts an example of each type. Figure 9A shows a set 23 distinct structures, each obtained empirically via NMR. The superposition of these structures reveals significant heterogeneity in the population due to the flexibility of the N-terminus of the protein. Figure 9B depicts an ensemble of protein structures generated using the COREX algorithm, which takes into account the thermodynamic and structural data of the protein. Each microstate shows certain structural deviations from the average structure. The user of these ensemble representations can read off the dynamic properties of an individual molecule, for instance, identifying more flexible regions from less flexible ones. However, compared to the first strategy, these models better represent a key posit of the dynamic concept, namely, that a protein is not a unitary entity with a single structure but rather a structurally heterogeneous population of molecules.

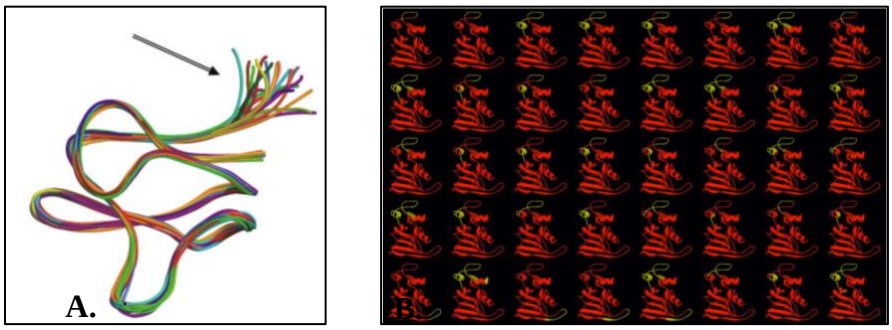

Figure 9 Dynamic Representation of Proteins: Ensemble Representation. (A) Acetyl coenzyme A carboxylase conformations. Colored lines show carbon backbone structure for 23 different NMR runs on the protein. Arrow highlights the structural diversity at the N-terminus of the protein. Republished with permission from Gunasekaran et al. (2004); permission conveyed through Copyright Clearance Center, Inc. (B) Ensemble representation of generic protein. Forty most probable ensemble structures generated via COREX are shown, with folded regions (red) and unfolded regions (yellow) shown for each state. Republished with permission from Lui et al (2006); permission conveyed through Copyright Clearance Center, Inc.

All these representations of proteins are associated with the dynamic concept: they incorporate beliefs and hypotheses that are involved in the inferential role of protein$_D$ but not protein$_S$. For instance, they represent the dynamic structural fluctuations in a single molecule, or they represent the key tenet that proteins in solution will exist as an ensemble of microstate conformations. Moreover, these representations figure in explanations of protein behavior that violate the structure-function rule and instead reveal the importance of dynamic properties. Compared to users of protein$_S$, the community of scientists using protein$_D$ and reasoning with this set of representations reached very different conclusions about the relationship between protein structure, dynamics, and function.

4.3 Inadequate Characterizations of the Relationship Between Protein Concepts

The emergence and development of two distinct concepts of the proteins since the birth of classical molecular biology raises the question: what epistemic relationship holds between the

two? Although philosophers of biology have not addressed this particular question, they have spilled much ink discussing conceptual change in genetics. Beginning in the 1960s, philosophers of biology interested in molecular biology have attempted to characterize the relationship between successive genetic theories and their different concepts of the gene. Within this context, philosophers focused their attention on reduction, with the earliest arguments centered around Nagelian intertheoretic reduction and more recent debates turning to consider explanatory reduction in genetics. In my analysis of protein concepts, I will follow the path established by these earlier philosophical debates about concepts of the gene. I will consider whether either Nagelian or explanatory reduction can adequately characterize the relationship between the structural and dynamic concepts of the protein, and I will also consider the Kuhnian account of scientific change, since many scientists who endorse the dynamic concept have suggested that molecular and structural biology are undergoing a paradigm shift. Ultimately, I conclude that these accounts shed little light on the relationship between the structural and dynamic protein concept.[80]

4.3.1 Reduction

4.3.1.1 Nagelian Intertheoretic Reduction

Nagel (1961) presented reduction as a way to characterize the relationship between scientific theories. Intertheoretic reduction is of philosophical and scientific interest because it is a special type of explanation: it is "the explanation of a theory or set of experimental laws

[80] Because these concepts emerged at particular moments during the history of molecular and structural biology, there is an inevitable historical dimension to the question motivating this section. The Nagelian account of reduction and especially Kuhn's account of scientific revolutions both include a historical component, since they aim, in part, to characterize the development of science. Nevertheless, my analysis here will focus, for the most part, on the epistemic rather than historical relationships these accounts posit between successive theories, explanations, or paradigms.

established in one area of inquiry, by a theory usually though not invariably formulated for some other domain" (Nagel 1961, 338). By demonstrating that one theory, developed for one domain, reduces to another more fundamental theory, which applies to a larger domain, one explains the former theory by subsuming it under a more general theory. Following the deductive-nomological, or covering law, model of explanation, Nagelian reduction explains the reduced theory (T_f) by showing that it is logically derivable from the reducing theory (T_r) and thus any consequences of it were to be expected. T_r reduces T_f if all the laws of T_f can be logically derived from T_r. Nagel supplements this "condition of derivability" with a "condition of connectability" in order to account for reduction between two theories with differing theoretical terms. If the theory being reduced includes terms that do not appear in the reducing theory, reduction is still possible if one can establish "bridge laws" that connect the terms in T_f to something that appears in T_r. Most of the early debate sparked by Nagel's account of reduction—including debates about the possibility of reduction in genetics—focused on these formal conditions.[81]

Nagel distinguishes two types of intertheoretic reduction. Homogeneous reduction, which occurs when the two theories share all the relevant concepts, is typically domain-preserving. Nagel's favorite example of this sort was the reduction of Galilean mechanics by Newtonian mechanics.[82] It is domain-preserving, since the reducing theory expanded the domain covered by the reduced theory to include both celestial and terrestrial dynamics. Cases of homogeneous

[81] For a history of the reception of Nagelian intertheoretic reduction by philosophers of biology, specifically, see Sarkar (1998). For a more recent history of the reception by philosophers of science more generally, see Sarkar (2015).
[82] Nagel (1961): "The range of application of a macroscopic theory may be extended from one domain to another homogeneous with it in the features under study, so that substantially the same concepts are employed in formulating the laws in both domains. For example, Galileo's *Two New Sciences* was a contribution to the physics of free-falling terrestrial bodies, a discipline which in his day was considered to be distinct from the science of celestial motions. Galileo's laws were eventually absorbed into Newtonian mechanics and gravitational theory, which was formulated to cover both terrestrial and celestial motions. Although the two classes of motions are clearly distinct, no concepts are required for de scribing motions in one area that are not also employed in the other." (Nagel 1961, 339)

reduction, such as this, are "normal steps in the progressive expansion of scientific theory," as the theory and laws of one domain are shown to apply more broadly (Nagel 1961, 338). In contrast, inhomogeneous reduction occurs when the reduced theory uses theoretical terms that do not appear in the reducing theory. Reduction of this sort often involves an older theory that "deals with macroscopic phenomena" and a newer, reducing theory that "postulates a microscopic constitution for those macroscopic processes" (Ibid., 340). While homogenous reductions are relatively straightforward, according to Nagel, inhomogeneous reductions are more philosophically puzzling, since the reduced theory contains concepts not included within the reducing theory. His account was developed primarily to explain how intertheoretic reduction could occur in these cases in which "the subject matter of the primary science seems qualitatively discontinuous with the materials studied by the secondary science" (Ibid., 342).

The most famous example of a successful inhomogeneous reduction, developed by Nagel, is the reduction of thermodynamics to statistical mechanics. Nagel focused specifically on the relationship between the ideal gas law, or the Boyle-Charles' law, and the kinetic theory of gases. The thermodynamic law can be derived from the kinetic theory of matter, but only with the addition of certain "bridge laws" to connect terms from thermodynamic theory (e.g., temperature) to traits represented by theoretical terms in statistical mechanics (e.g., kinetic energy of gas molecules). Nagel (1961) endorsed pluralism with respect to the bridge laws that connect terms in the old and new theories: they could be logical connections, conventions, or empirical claims.[83] For example, the bridge law that identified "temperature" from the Boyle-Charles' law with the "mean kinetic energy of gas molecules" in the kinetic theory of gases was not a logical connection,

[83] Later modifications to Nagel's account, such as Schaffner (1967), drop this pluralism and instead require the relationship between terms to be a biconditional, expressing an identity relation.

according to Nagel, but instead was best construed as either a convention or an empirical fact, depending on the context.

Early discussions of Nagel's account within philosophy of biology aimed to assess its applicability to changes in genetics in the twentieth century, specifically focusing on the relationship between Mendelian genetics and classical molecular genetics.[84] According to Sarkar, "most of these disputes have been about *formal* issues, the 'logical' form of reduction" (1998, 17, emphasis in original). Those who claim that the formal conditions for Nagelian reduction have failed to obtain in this case have typically focused on one of two issues. They have argued either (1) that biology, or a particular biological discipline, does not have theories and laws, so the Nagelian account does not apply, or (2) that there is no sufficiently simple bridge law that can connect the concept of the Mendelian gene with the newer molecular gene concept.[85] Similar formal problems with applying Nagel's account to biology reappear in the protein case. On my view, theories are not the right unit of analysis, and any bridge law we develop to connect the structural and dynamic concepts of the protein will be unilluminating, since it will likely amount to little more than replacement of the one concept for the other. After briefly presenting my arguments, I then turn to a more substantive issue. I argue that Nagelian intertheoretic reduction is not occurring in this case since the potential reducing theory associated with the dynamic concept (i.e., statistical thermodynamics) is not more fundamental than the theories of molecular biology and chemistry associated with the structural concept.

The focus on theories as the unit of analysis suggests that Nagelian intertheoretic reduction cannot accurately capture the relationship between the structural and dynamic protein concepts.

---

[84] See, for example, Schaffner (1967), Hull (1972), and Kitcher (1984).
[85] Kitcher (1984), Hull (1972), and Dupre (1993) are among many examples of philosophers arguing for these two positions.

As we have seen, the inferential roles of the two protein concepts involve theoretical definitions and descriptions, but they also involve various methods, research and explanatory heuristics, and representational and modeling conventions. Nagel's formal account, which is couched in terms of the syntactic account of theories, assumes scientific theories are axiomatized sets of propositional statements. But this is not the way most philosophers or scientists think about the inferences, explanations, heuristics, and hypotheses associated with the two protein concepts. Even if we could axiomatize the research programs in which each concept is embedded, which is doubtful, it is highly unlikely that this would result in a cohesive, unitary theory.[86] However, without axiomatization, it is impossible to determine whether the logical derivation requirement holds. For reasons such as this, even Waters (1990), a defender of reduction in the genetics case, suggests that biological theories will typically fail to satisfy the formal conditions of Nagel's account.

There is a further problem for applying intertheoretic reduction to the protein case. Even if we were able to reconstruct unitary theories associated with each protein concept, before we could begin the derivation required for reduction, we would have to settle upon bridge laws to connect the two distinct protein concepts. This step was the focus of many who argued against reduction in genetics. Philosophers argued that there was no systematic mapping between the Mendelian gene and the molecular gene that would allow the establishment of a bridge law. And since there were no bridge laws forthcoming to connect these most central concepts, they concluded that reduction—at least according to the Nagelian account—had not happened and was unlikely to happen.[87] In the genetics case, one major difficulty occurred because of the change in reference

[86] This is similar to Kitcher's (1984) argument that classical molecular genetics lacks a unitary theory. Waters (1990) also considers whether the reduction of Mendelian genetics will require multiple theories from different biological domains, what he calls the "splintering argument" against reduction.
[87] See Kitcher (1984) and Dupre (1993) for examples of this argument focused on bridge laws.

for the gene concept. The Mendelian gene was an abstract gene, but later gene concepts, such as the gene of classical molecular biology and more recent concepts, seemed to refer to different bits of DNA, such that different DNA sequences would be part of the reference potential for different concepts (Griffith and Stotz 2013). The reference for protein concepts is less fraught, since both protein_S and protein_D include a single, folded polypeptide chain. Therefore, there is less chance that the token "protein" will refer to different entities when used by scientists endorsing different protein concepts. However, even in this case, the relationship fails to be one-to-one, since the reference potential of the dynamic protein also includes an ensemble of polypeptide molecules.

Another more problematic difficulty arises when trying to establish an appropriate bridge law connecting the protein concepts themselves: their inferential roles involve many incompatible, or even contradictory, posits. Cooper (1976) sums up this problem, presenting the two "conflicting views" of proteins:

> One, a compact structure in which the polypeptide chain is precisely folded to give a tightly interlocking, rigid molecule; the other, a "kicking and screaming stochastic molecule" (Weber 1975) in which fluctuations are frequent and dramatic. These fluctuations produce a seemingly fluid and flexible system. (Cooper 1976, 2741)

To connect these concepts via a bridge law, then, would require radical revisions in the meaning, or inferential role, of the structural concept. Both the privileged status of the 3D structure of the molecule and the belief that structure alone determines function would need to be jettisoned from the inferential role of the structural concept in order to connect it to the dynamic concept. Since these theoretical posits form the core of protein_S, justifying its use in most inferential and explanatory contexts, it seems that this revision in the concept would be tantamount to replacement of protein_S with protein_D. Similar to the previous philosophical debates about genetics, it seems

that the formal conditions of Nagelian intertheoretic reduction make it unable to illuminate the relationship between protein concepts.[88]

There is a more substantive issue unique to the protein case that reinforces these formal arguments against the applicability of Nagelian intertheoretic reduction. At first blush, one might assume there is a relevant similarity between this case and the canonical example of successful Nagelian reduction—viz., the reduction of thermodynamics by statistical mechanics—because of the role of thermodynamic theory in the dynamic view of proteins. Despite this similarity, it is mistaken to think that thermodynamics plays a similar role in both accounts. Indeed, I argue that it plays precisely the opposite role in the protein case.

According to Nagel (1961), inhomogeneous reduction typically occurs as part of the normal progression of science when a higher-level, macroscopic theory is reduced to a lower-level, microscopic theory. But in the case of protein concepts, the older, structural concept involves theory from molecular biology and chemistry that provide lower-level molecular detail about the 3D structure of the protein (microscopic), whereas the newer dynamic concept involves higher-level thermodynamic theory (macroscopic). In this case the candidate reducing theory describes the macro-level (e.g., thermodynamic systems), whereas the candidate reduced theory focuses on the micro-level (e.g., individual molecules). The historical progression is therefore the reverse of the one described by Nagel. Moreover, we might reasonably think that the candidate reducing theory, in this case, is not more fundamental than the reduced theory. For this reason, even if the dynamic concept and its associated theoretical framework can provide a unified view, framework,

[88] This result was to be expected, since as the name suggests, Nagelian intertheoretic reduction was meant to characterize an epistemic relation that could obtain between theories. My reason for including this discussion of Nagel is because of the importance it played in earlier debates about genetic theory and gene concepts. However, although I have not argued for the position here, I also suspect that intertheoretic reduction would not obtain if we were to analyze the theories associated with the structural and dynamic concepts.

or description for explaining protein behavior, as some scientific advocates claim (see, e.g., Motlagh et al. 2014), it does not seem to achieve this unity via reduction of the older structural framework, in accordance with Nagel (1961).[89]

Consideration of these formal and substantive issues have shown that Nagelian intertheoretic reduction is either inapplicable or unilluminating for understanding the change in protein concepts in the late twentieth century. Nevertheless, there is a sense in which the evolution in protein concepts embodies some of the spirit of Nagel's account. One goal of his account was to make sense of the progress of science. By making evident the ways old theories relate to newer ones, Nagelian intertheoretic reduction offers an explanation of scientific progress, while at the same time demonstrating that this progress is rational.[90] The dynamic concept, with its theoretical framework, hypotheses, and methods, does have the resources to explain why explanations that used the structural concept succeeded as well as they did. However, as I will argue in Section 4, there are better ways to explain this fact that do not construe the epistemic relationship between the concepts as one of reduction.

4.3.1.2 Explanatory Reduction

Within philosophy of biology, recent discussions of reduction have shifted away from intertheoretic reduction, and the debate about the relationship between Mendelian and molecular

---

[89] According to Sarkar, "part of the appeal of Nagel's analysis is that it seems to capture what the mechanical philosophy of the seventeenth century was about; what Maxwell, Clausius, and Boltzmann attempted to do with thermodynamics in the nineteenth century; and what Pauling and Crick promoted through the molecularization of biology in the twentieth century" (2015, 44). The shift in protein science from descriptions of proteins as molecular machines using the structural concept to the analysis of proteins as small thermodynamic systems using the dynamic concept seems to be the reverse of the process Nagel (1961) aimed to capture.

[90] Nagel's account is also part of the mid-twentieth century logical positivism project that tried to demonstrate the unity of science. By showing how old and new theories and theories in different domains are logically connected, Nagel's account offered support for the unity of science thesis. If most scientific theories are connected via reduction, then the possibility of developing one unified account of science (or rather one unified account of scientific theories and laws) was much more likely.

genetics, in part, prompted this shift. The failure of this case to fit the formal Nagelian account of reduction led to the emergence of an anti-reductionist consensus among philosophers of biology by the end of the twentieth century (Sterelny and Griffiths 1999). But many philosophers of biology, including those in agreement with this consensus, saw molecular genetics as "an obvious triumph for reductionistic research" (Griffiths and Stotz 2013, 57). For this reason, some philosophers came to reject Nagel's account of reduction, since it failed to apply "to many important cases that [were] (pre-systematically) accepted as reductions" (Sarkar 2015, 44). Instead, they developed an alternative—viz., explanatory reduction—that could account for the apparent reductive successes of molecular biology. If this shift from intertheoretic to explanatory reduction could clarify the relationship between older and newer approaches to genetics, then we might wonder if such a philosophical move could also make sense of the relationship between protein concepts and their associated explanatory strategies. In this case, however, I argue that the history of gene concepts and protein concepts diverge: the shift from protein_S to protein_D is not accurately characterized by explanatory reduction.

Although philosophers of biology have developed different accounts of explanatory reduction, there is general agreement about the key features. For instance, most would agree with the spirit of Sarkar's brief characterization:

> Reduction will be construed as the explanation of wholes by parts, that is, reductionist explanations are those in which the weight of a putative explanation is borne by properties of the parts alone. (Sarkar 2008, 68)

From this gloss, we can identify two features that are necessary for an explanation to be a reductive explanation. First, the system to be explained must be represented as having an "explicit hierarchical organization" (Sarkar 1998, 43). This abstract hierarchy need not be spatial. However, "strong reduction" occurs when it is instantiated in space (Ibid., 45). In this case, we will be able

to distinguish entities at different levels because lower-level entities will be "spatial parts of entities at higher levels of the hierarchy" (Ibid., 44). Brigandt and Love (2017) agree that reductive explanation in biology is part-whole explanation.[91] The levels in the hierarchy are typically discovered via decomposition and localization of a biological system into lower-level parts (Bechtel and Richardson 2010). This leads to the second condition: the explanation must cite only features of the system that occur below the level of the system itself (Kaiser 2015). In other words, the "explanatory factors" cited to account for the behavior of the biological system must "refer only to properties of entities at lower levels of the hierarchy" (Sarkar 1998, 47). This second condition ensures that reductive explanations maintain "an epistemically important unidirectional 'flow' of explanation from the lower to the higher level" (Brigandt and Love 2017, 26).[92]

At first blush, we have some reasons to think this type of reduction might illuminate the relationship between the structural and dynamic concepts of the protein. First, the molecularization of biology and related sciences are seen as great scientific achievements because molecular biology has the ability to describe and explain the molecular basis for macroscopic biological phenomena. Scientific progress has occurred by drilling down deeper to uncover molecular mechanisms and by developing mechanistic explanations of biological phenomena. Indeed, this is precisely what Griffiths and Stotz (2013) claim happened with molecular genetics. Because of this historical trajectory, one might suspect that the newer, dynamic concept of the protein marks an improvement over the earlier structural concept because it offers a lower-level description of

[91] Kaiser (2015) agrees with the abstract hierarchy criterion but does not think that reductive explanations in biology require spatial containment. Thus, she disagrees with Brigandt and Love (2017), arguing that not all reductive explanations are part-whole explanations. She contends that a reductive explanation can include entities that are not parts, so long as they occur at the lower-level, as required by the second condition. (See footnote 13 for further discussion of Kaiser's account of levels.)

[92] Sarkar (1998) agrees. He claims that reductive explanations "proceed in a definite direction, that is, intuitively from lower through higher levels toward a sink" (55).

proteins. Second, the single 3D structure of the protein, posited by the structural concept, is replaced by an ensemble of structures in the dynamic concept.[93] This proliferation of structures within the dynamic concept requires more lower-level molecular and structural detail, which might seem to fit the reductive trend in molecular biology.

Despite this initial plausibility, it would be a mistake to think that explanations of protein behavior that use the dynamic protein concept provide explanatory reductions of protein behavior previously explained using the structural concept. I develop two arguments for this claim. Both focus on explanations that use ensemble representations of the dynamic concept, since these are the ones most likely to be misconstrued as reductive explanations. First, I argue that the relationship between the structural microstates and the ensemble does not exhibit the part-whole character required for reductive explanation. And second, I contend that dynamic properties should not be construed as lower-level properties of proteins. There is no abstract hierarchy that shows that the structural properties of proteins are at a higher level than their dynamic properties.

First, consider a representation of the "same" protein using either the structural or dynamic concept. In the former, there is a single structure whereas in the latter there is an ensemble of microstate structures, representing the various conformations that the protein can assume in solution. In this transition from the structural to the dynamic concept, the unique structural representation (i.e., what was previously considered to be *the* protein) becomes only one of many structures in the ensemble. Thus, the older structural representation exists as an individual within

[93] According to the structural concept, a population of protein molecules would behave like a set of identical molecular machines, each carrying out the same task. Since all the molecules were identical in this way, it made sense to look for an explanation at the level of the individual protein. In contrast, according to the dynamic concept, the single molecular structure of $protein_S$ has given way to a multiplicity of microstate structures all carrying out the same task but in slightly different ways. Hence, the adoption of $protein_D$ requires attending to features of the population, as well as the structural differences between individual molecules, in order to explain protein behavior.

the new group of structures in the ensemble representation. It is therefore possible to construe the relationship between that single structural representation and the dynamic representation as a part-whole relationship, since the structure is a microstate within the ensemble. This part-whole relationship, however, is very different from the part-whole relationship that exists in mechanistic or reductive explanations. In reductive explanations, the part-whole relation arises via decomposition of the system of interest into its components (Bechtel and Richardson 2010, Craver 2007), and this decomposition is spatial, yielding distinct parts that are spatially contained within the whole. Moreover, because of the constitutive relation between the parts and the whole, the parts are necessarily smaller than the whole of which they are a part.

The part-whole relationship between a single microstate structure and the ensemble differs in both respects. For one, the 3D structural representation of the protein is not decomposed into different parts (e.g., it is not decomposed into different structural domains). Instead, the proliferation of structures required by the dynamic concept of the protein represents all the possible structures that the protein could assume in solution. It can be thought of as providing either snapshots of the temporal structural fluctuations of any given protein molecule, or a representation of the instantaneous equilibrium distribution of molecules across all microstates.[94] The shift from the structural to dynamic representation is therefore not accurately characterized as a decomposition of the structural protein into component parts. In addition, the putative parts in the dynamic representation are not spatially smaller than the protein represented according to the structural concept. In mechanistic explanation, the parts are spatially contained within the whole, but in the ensemble representation, the old structural representation reappears as a microstate

[94] Because of the ergodic hypothesis, the ensemble representation can be interpreted as either (1) the temporal evolution of a single molecule or (2) the instantaneous representation of the population of molecules (Hilser et al. 2006).

structure. It is therefore of equivalent size to any of the other microstate structures in the ensemble representations. In no sense is one of the microstate structures in the ensemble representation smaller than the single structural representation of the protein under protein$_S$. Therefore, we can conclude that the only part-whole relationship that exists between the structural and dynamic representations of the protein—viz., the microstate and ensemble—does not have the appropriate character to enable explanatory reduction.

Second, in addition to their part-whole character, reductionist explanations must only cite explanatory factors at the lower-level in order to explain the behavior of phenomena at higher levels.[95] However, the dynamic properties of proteins associated with protein$_D$, and the structural properties, which are at the conceptual core of protein$_S$, are not related hierarchically. There is no "independent criterion" (Sarkar 1998, 53), for establishing an abstract hierarchy, with structure as a higher-level property and dynamics as a lower-level one.[96] Thus, even though the shift from the structural concept to the dynamic concept can involve the addition of dynamic properties to structural representations of proteins, this dynamic information does not introduce a new, lower level of analysis, either spatially or abstractly. The addition of dynamic properties, for example, does not require the use of quantum-level theories to explain the atomic-level structural properties. Talk of "levels" in this case hardly makes sense. The thermodynamic framework invoked by those who endorse the dynamic concept also fails to establish the necessary hierarchy. The energetic

[95] One major problem for this understanding of reductionist explanation in this way is ambiguity in the concept of "levels". When construed as part-whole explanations, the levels are established locally via the part-whole relationship: the whole is at a higher level than the parts from which it is composed (Craver 2007). Kaiser (2015) develops an account that attempts to provide a more global account of "levels," since her account of reductive explanation permits non-parts in the explanans so long as they are also at a lower level. However, I am skeptical that it can offer an unambiguous account of levels in biology. See Woodward (2020) for further discussion of levels.

[96] Sakar (1998) requires that there must be some independent criterion that establishes the hierarchy. It is "independent" in that it "should not have been posited only for the sake of the explanation at hand" (53). However, the hierarchy could be established by a general research program or explanations of another phenomenon.

components related to structure (ΔH) and dynamics (ΔS) are additive components that jointly determine the free energy (ΔG) of proteins, which governs protein behavior: the structural and dynamic properties of proteins are co-dependent and not at different levels. There is thus no plausible construal of the structural properties of proteins as being at a higher level than their dynamic properties.

Molecular biologists adopted the structural concept first, representing the protein as a single, static structure, and only later discovered the functional importance of structural fluctuations, which led them to adopt a dynamic concept that explicitly represented the dynamic properties of the protein. Despite this historical narrative, which expresses contingent facts about progress in molecular biology, the shift from the structural to the dynamic concept does not indicate that structural properties in representations associated with the former occur at a higher level than the dynamic properties included in representations associated with the latter.

In sum, explanatory reduction fails to capture the relationship between protein structure and dynamics. The history of protein concepts, in this case, diverges from that of gene concepts. Although philosophers of biology largely rejected Nagel's account of intertheoretic reduction in the context of genetics, many suggested that the new science, molecular genetics, provided reductive explanations of phenomena that the older Mendelian science only described cytologically (Griffiths and Stotz 2013, Waters 1990). In contrast, I have argued that neither Nagelian reduction nor explanatory reduction can accurately characterize the relationship between the structural and dynamic concepts and their associated representations.

4.3.2 Kuhnian Paradigm Shifts

Kuhn developed an influential, non-reductive account of the development of science. His *Structure of Scientific Revolutions* (1962), which introduced the concepts of "scientific paradigms" and "paradigm shifts," was the most-cited book in the history and philosophy of science in the twentieth century and was read by generations of scientists (Kaiser 2012). His account focuses on revolutions in science that drastically reshaped theory and practice in different scientific fields. It is therefore unsurprising that scientists endorsing the dynamic concept of the protein have gestured toward Kuhn's account of scientific revolutions and claimed that understanding of proteins and protein behavior is undergoing a paradigm shift. Vincent Hilser, a structural biologist at Johns Hopkins and one of the most vocal champions of the dynamic protein concept, has argued that it is the foundation for a "new paradigm" that has replaced the "historically dominant paradigm" (Motlagh et al. 2014, 331). Nobel prize-winner Martin Karplus has also presented this change as a "paradigm shift" from an old view focused on structure to a "new view" that "emphasized the intrinsic dynamic nature of proteins" (Cui and Karplus 2008, 1304). In a "Perspectives" article in *Cell Structure*, two structural biologists even cited Kuhn before claiming that a "paradigm shift is underway" (Forman-Kay and Mittag 2013). These scientists typically support their assertions by showing how the dynamic concept, with its associated theoretical posits, hypotheses, and explanatory strategies, has overturned the structure-function rule and provided a new framework to characterize and explain protein function. For the scientists, the reference to paradigm shifts is likely a rhetorical flourish used to signal the novelty and importance of their view. Nevertheless, it is worth considering whether Kuhn's account can make sense of the relationship between the structural and dynamic protein concepts.

Kuhn's (1962) account of scientific revolutions aims to offer a historical account of theoretical and conceptual change in the sciences. According to Kuhn, as a scientific field develops, it establishes a paradigm, which is a set of core concepts, theories, methods, and other aspects of scientific practice accepted by the relevant community of scientists. The paradigm defines the relevant scientific problems and outlines the shape of the solutions according to a set of exemplars. During "normal science," the work of most scientists will be solving problems following these exemplars and further articulating the paradigm to show how it applies in different situations. Theoretical and conceptual change occurs when anomalous phenomena that cannot be explained by or incorporated into the paradigm accumulate. After the accumulation of sufficient anomalies, the paradigm breaks down, leading to a scientific revolution, in which the scientific community has not yet settled on a new set of guiding concepts, theories, methods, and practices. Eventually, a new paradigm arises that can make sense of most of the anomalies, and it is adopted by the scientific community. According to Kuhn, the shift from one paradigm to another is like a Gestalt shift: within the new paradigm, the scientist sees the world differently than she did before.

Whereas Nagel's account takes theories to be the primary unit relevant to the analysis of scientific change, Kuhn's account posits the scientific paradigm. The bigger difference between these two accounts—and the one that sparked more debate within philosophy of science—lies in the epistemic relation that connects the old and new theories or paradigms. For Nagel, the two theories are connected via the reduction relation, which is grounded in logical derivability. In contrast, Kuhn argues that the Gestalt-like paradigm shift is not wholly rational. According to Kuhn, the terms and theoretical concepts in the new theory, even identical terms, will be incommensurable, since they will refer to different things and have new meanings in the context

of the new paradigm.[97] Therefore, a theory or concept that gets replaced during a paradigm shift will have no necessary epistemic relationship to any new theories or concepts that emerge to describe and explain phenomena within the new paradigm.[98]

Despite the claims of some scientists, the Kuhnian model of scientific change does not accurately describe the relationship between the structural and dynamic concept of the protein. For one, if the introduction and adoption of the dynamic concept were part of a paradigm shift in molecular and structural biology, then the new dynamic concept and the old structural concept would be incommensurable. Scientists working within the new paradigm would not be able to make sense of those working in the old paradigm, since they would "practice their trades in different worlds" (Kuhn 2012 [1962], 149). While it is true that the structural concept and the dynamic concept are incompatible, since their inferential roles involve inferences and explanations that are at odds (e.g., one cannot simultaneously accept the claim that protein function can be explained by structure alone as well as the claim that structure and dynamics jointly determine function), it is false to say that these concepts are incommensurable.

Scientists who endorse one or the other concept can make sense of claims made using the other concept, despite the change in reference and inferential roles. For example, in a recent review of protein function, Hilser and colleagues (2012) describe how one function—allostery—is

---

[97] It is precisely this claim that motivates Kitcher (1978, 1982) to develop his alternative account of conceptual change that shows how concepts can acquire and lose modes of reference across theory change yet still be compared. Brigandt's (2010) account of conceptual change also seeks to show how changes in the meaning of concepts, as well as their referents, can be rational, when we consider the epistemic goals of the concept's use.

[98] As Ian Hacking points out, however, Kuhn did not deny the "very rationality of science," as it is sometimes claimed (Hacking 2012, xxxi). On Kuhn's account, under any paradigm, "theories should be accurate in their predictions, consistent, broad in scope, present phenomena in an orderly and coherent way, and be fruitful in suggesting new phenomena or relationships between phenomena" (Ibid.). But these epistemic values track the relationship between theories and the world and not successive theories. Thus, it is not inconsistent with Kuhn's view to claim, as I have, that his account provides for no necessary epistemic connection between theories, which is the focus of the current chapter.

explained using the structural concept and the structure-function rule and then again using the dynamic concept and the ensemble representation. He presents these as alternatives and ultimately advocates the latter explanation of allostery, but he argues for this preference by claiming that the latter strategy offers a more accurate and unified explanation of allostery that can account for more cases.[99] That is, Hilser argues for the superiority of the dynamic concept—he does not simply hope for his interlocutors to undergo a "conversion" to his view. The expectation, then, is that scientists in this discipline will be able to recognize, comprehend, and reason with both concepts.

There is another reason to think that Kuhn's account cannot illuminate the change in protein concepts. Because it casts the two paradigms as radically distinct, leading to different worldviews, it does not have the philosophical resources to explain why some aspects associated with the old concept—e.g., research problems, methodological strategies, and experimental techniques—have carried over into research and theorizing associated with the new dynamic concept of the protein.[100] On the Kuhnian account, the newer dynamic paradigm should replace the older structural one, and the characteristic research problems, methods, and explanatory strategies of the old paradigm should also be replaced. Within structural biology, however, this has not happened. The Kuhnian account is too blunt an instrument to adequately account for the change that is occurring in structural biology. Although I will develop an account of conceptual change and advocate for replacement of the structural concept in Section 4, my account will show why we should not expect conceptual replacement in this case to lead to a transition from one set

[99] According to Hilser, the dynamic concept is more accurate, since it represents dynamic structural fluctuations that are excluded from the inferential role of the structural concept. He also seems to adopt the position that more unified explanations are superior to less unified ones (see, e.g., Motlagh et al. 2014). Since the thermodynamic framework associated with the dynamic concept of the protein offers a unified account of more protein behaviors than structural-mechanistic framework does, he claims that the former is superior.

[100] Kuhn moderates this view somewhat in *The Essential Tension* (1977).

of methods, models and representations to another set, but will instead lead to an expansion of scientific practices that aim to characterize and explain protein function.

4.4 How to Make Sense of Changes in Protein Science

4.4.1 Two Levels of Analysis: Protein Concepts vs. Representations

In order to adequately characterize the conceptual change in molecular and structural biology, we must distinguish two levels of analysis—viz., concepts and representations. Scientific concepts are theoretical terms that have an inferential role (or meaning), reference, and epistemic goal as components of their semantic content (Brigandt 2010). As we have seen in Section 2, a concept will be compatible with many different representations. But any particular concept will be compatible with only a certain set of representations.[101] Change at the conceptual level will likely lead to changes at the representational level, since changes in the inferential role (or reference potential) of the concept will likely involve the development of new representations and exclude the use of certain old ones. There is no principled reason, however, to think that the epistemic relationship that holds between two different concepts is the same as the relationship that holds between the representations associated with each concept. Therefore, any adequate account of the changes in molecular and structural biology vis-à-vis proteins will need to explain the epistemic

[101] The focus of my analysis is on visual representations of proteins, such as the models of proteins discussed in Section 2 of this chapter, although it could be extended to include other representations, such as mathematical or linguistic representations. The main claim I will defend is that the dynamic concept will be able to accommodate more types of representations than the structural concept, since its inferential role includes beliefs about protein dynamics that are excluded from the inferential role of the structural concept.

relationship between the structural and dynamic views at both the level of concepts and the level of representations.

4.4.2 Protein Concepts and Conceptual Replacement

The epistemic relationship between the structural and dynamic protein concepts, I argue, is best understood as an example of conceptual replacement. On my view, the difference in the inferential roles of the concepts is too significant to count as a meaningful reduction. To support this claim, it will be worthwhile to review the types of semantic content that are involved in the inferential role of a concept before comparing the inferential roles of the structural and dynamic concepts.

According to Brigandt (2010), the inferential role of a concept extends beyond the theoretical definition, including other features necessary to explain how the concept is successfully used in "scientific inference, explanation, and discovery" (Ibid., 22). Brigandt's specification of the inferential role of the classical molecular gene is illuminating. He claims it

> includes the genotype–phenotype distinction, the idea that genetic loci are arranged in linkage groups, beliefs about the basic principles specifying the transmission of genes (linkage, recombination, segregation), and ideas about the relation between genotype and phenotype (such as dominance and recessiveness). (Ibid., 26)

This example shows the breadth of the inferential role of a concept and just how far it extends beyond the theoretical definition a scientific user of the term might give when asked. The inferential role of the concept, in this case, explains how it was used in scientific practice and theorizing.

As we have seen, the protein concept underwent a significant change in the late twentieth century. It was sufficiently drastic for certain scientists to begin to distinguish the old structural

concept of the protein from the new dynamic concept. At the heart of the old structural concept were two commitments: (1) the belief that every protein had a rigid and static 3D structure and (2) the belief that the protein structure alone determined protein function. Taken together, these two core beliefs established a heuristic for research and laid out a strategy to explain protein behavior and function. They also promoted the use of certain instruments and experimental techniques, such as x-ray crystallography, to elucidate protein structure. This conceptual core gave rise to a host of other more specific features that were involved in the use of the concept, such as the "lock-and-key model" of catalysis and the "molecular shape rule," that directed scientists to attend to the general properties of the surface of molecules (Sarkar 2008, 61).

The new dynamic concept of the protein rejects both of these commitments at the core of the structural concept. Instead, the inferential role of this new concept includes the belief that proteins are flexible and dynamic and are constantly undergoing structural fluctuations. Moreover, users of the dynamic concept are committed to the belief that dynamics, as well as structure, are an important determinant of protein behavior and function. Unlike the structural concept, which was largely wedded to one technology, the users of the dynamic concept promoted a diversity of experimental, theoretical, and computational strategies to probe the dynamic properties of proteins.[102] The explanatory strategies involved in the inferential role of the concept were similarly diverse, with some using thermodynamic analyses to distinguish structural and dynamic contributions to function, and others adopting computational approaches. Despite the diversity of methodological and explanatory strategies, the common theoretical core of the dynamic concept maintained that dynamics and structure were both causally relevant to protein function and denied the privileged status accorded to protein structure by the structural concept.

[102] Cooper (1980) provides and early review of these dynamic approaches.

This characterization of the inferential role of the dynamic concept serves as a reminder that the name of this new concept can be easily misconstrued. Although it emphasizes the major novelty (viz., dynamics), the name is potentially misleading. Since it is often contrasted with the old structural concept, one might assume that the explanations, inferences, and hypotheses involved in the inferential role of this new concept direct scientists to focus on protein dynamics *rather than* protein structure. However, this is not the case. The dynamic concept guides scientists *to attend to both* structure *and* dynamics, if they are interested in explaining protein function. Table 1 summarizes the key differences in the inferential roles of the protein concepts.

Table 1 Three Semantic Components of Protein Concepts

	Structural Concept	**Dynamic Concept**
Reference	– single folded polypeptide chain	– single folded polypeptide chain – ensemble of such polypeptide chains
Inferential Potential	– rigid and static 3D structure – structure alone determines function (structure-function rule) – mechanistic explanations consider proteins to be molecular machines	– flexible and dynamic 3D structure – structure and dynamics determine function – thermodynamic analysis treats proteins as thermodynamic systems
Epistemic Goal	– explain and predict protein behaviors, in part, to enable interventions	– explain and predict protein behaviors, in part, to enable interventions

This side-by-side comparison of the two concepts supports my contention that the relationship between the old structural concept and the new dynamic one is best interpreted as a case of conceptual replacement rather than reduction. The core theoretical posits of the two concepts, which guide their use in inferences, explanations, and research, are either incompatible or contradictory. For instance, the structure-function rule of the old concept, which holds that structure alone determines function, is contradicted by the new concept's explanatory rule, which

holds that structure and dynamics co-determine protein function. The inferential role, or meaning, of the concept has changed significantly, which is sufficient to identify the conceptual change and distinguish the usage of two concepts (Brigandt 2010). Unlike in the case of gene concepts, however, there is no difficulty in establishing the appropriate connection between one concept and the other, since their referents have remained largely unchanged. [103] Nevertheless, this is not a case of reduction in the spirit of Nagel, since the change in the inferential roles of the concept do not allow one to connect the concepts while preserving their meaning in each respective theoretical context. In other words, if one were to attempt to correct the inferential role of protein_S so that it could be appropriately linked to protein_D, it would require replacing the inferential role of the old concept with that of the new.[104] And this, I suggest, would not be reduction in any meaningful sense. Instead, we ought to think of it as the replacement of the old structural concept of the protein with the new dynamic one.

However, we can say something further about the epistemic relationship between the two protein concepts. It is not at all the case that the new dynamic concept is incommensurable with the old structural concept. If we compare the theoretical posits that form the core of their inferential roles, we see that the posits and rules that guide the use of the dynamic concept are more permissive than those of the structural concept. Where the structural concept focused only one cause of protein function (structure), the dynamic concept focuses on two (structure and dynamics). Similarly, the structural concept was embedded in research contexts that focused primarily on elucidating the 3D structure of proteins, while the dynamic concept guides researchers to many methods that

[103] As Table 1 shows, the referent of the dynamic concept has expanded to also include an ensemble of polypeptide molecules.
[104] Schaffner (1967) modified Nagel's original formulation, arguing that what is actually derivable from the reducing theory is not the reduced theory, but rather a strongly analogous corrected version of the reduced theory.

illuminate dynamic as well as structural features of proteins. Thus, in a sense, the replacement of protein_S with protein_D amounts to a quasi-unification. Advocates of the dynamic concept do not argue that structure is irrelevant to protein function. Quite the contrary. They recognize that the dynamic features of proteins previously excluded from the explanatory framework of the structural concept are instantiated within structures. Consider two examples from advocates of the dynamic concept:

> It is clear that *complete* understanding of the nature and function of protein molecules will require knowledge not only of their mean properties, but also of their dynamic characteristics, and that static descriptions of molecular structure are incomplete and may be misleading. (Cooper 1976, 2741, emphasis added)

> For a *complete* description of proteins, it is important, therefore, to know, in addition to the average structure, the form of the fluctuations that occur, to determine how they take place, and to evaluate their magnitudes and timescales. (Brooks et al. 1988, 2, emphasis added)

These endorsers of the dynamic concept interpret the dynamic concept as providing a conceptual unification, bringing together both structural and dynamic features of proteins. The new concept, which replaced the old, is thus more permissive and more inclusive, guiding scientists to expand the set of factors they consider causally relevant to protein function.

This fact, then, explains the puzzling feature of this particular conceptual change that the Kuhnian account could not—viz., why so much reasoning and research in structural and molecular biology remains the same after the replacement of the structural concept with the dynamic one. This occurs because the inferential role of the new dynamic concept is significantly expanded to include beliefs about protein dynamics and only jettisons those beliefs that afforded privileged status to protein structure. Even after conceptual replacement, use of the dynamic concept guides scientists to continue to investigate and draw conclusions from protein structure but also encourages them to recognize and pursue new investigations into protein dynamics will also be necessary to adequately explain protein function.

Philosopher Sandra Mitchell has also developed an account of the relationships between concepts and representations in protein science (Mitchell and Gronenborn 2017, Mitchell 2019). Therefore, it will be useful to distinguish my account of protein concepts and conceptual replacement from her account of scientific perspectives and integration. According to Mitchell's account of perspectival pluralism, scientific practice produces a variety of different models and representations that often must be integrated in order for them to be applied to particular problems and specific cases.[105] Pluralism arises because science relies on a variety of epistemic sources, and integration is then required because the representational outputs of these different sources are often not unifiable or reducible.

Mitchell (2019) presents the problem of protein structural determination as an example of scientific practice involving perspectival integration. Molecular and structural biologists want to be able to determine and/or predict the 3D structure of proteins. To do this, they rely on a variety of methods and approaches: physics, chemistry, and biology. According to Mitchell, each one of these approaches constitutes a perspective.[106] The physics perspective focuses on the thermodynamic properties of proteins and the Gibbs free energy of different conformations. It studies the protein *in silico*, attempting to model the thermodynamics of folding that ultimately lead the protein to adopt a given 3D structure. The chemistry perspective occurs *in vitro* and uses experimental techniques, such as x-ray crystallography or NMR, in order to determine the 3D structure of the isolated protein. Finally, the biology perspective looks at the protein *in vivo*, using techniques such as fluorescence that let scientists visualize the protein inside the cell. Each

[105] Although Mitchell has only recently framed her analysis in terms of "scientific perspectives," her early arguments for integrative pluralism contain similar arguments about the partiality of representation and the need for integration rather than unification. See, for example, Mitchell (2002), (2003), and (2012).
[106] Mitchell (2019) claims that, in this case, by using these three different scientific approaches, "one can crudely distinguish" the three different perspectives (185).

approach focuses on different features of proteins that are relevant to determining or predicting the 3D structure, and no single approach includes all the relevant features. Whereas the representational outputs of a given perspective may be able to address certain research questions about protein structure, according to Mitchell, other questions will require integrating information from multiple perspectives, since each perspective offers only a "partial grasp of the phenomenon" (Ibid., 188).

Mitchell characterizes a "scientific perspective" in much the same way that I, following the philosophical debates in genetics, have used the term "concept":

> Different scientific perspectives are characterized by different assumptions, methods, instruments of observation, experimental arrangements, concepts, categories, and representations, all of which are associated with specific pragmatic concerns and explanatory or predictive projects. (Mitchell 2019, 181)

Nevertheless, from her example from protein science, it is clear that Mitchell's perspectives are not the same as the concepts of the protein that are the focus of my analysis. The three perspectives cut across both concepts of the protein. In other words, scientists who endorse either concept can study proteins from any of the three perspectives. Take, for example, the group of scientists who have adopted the dynamic concept of the protein. Within this community, even as early as the 1980s, we find certain scientists using computational and theoretical thermodynamic approaches (physics perspective), others using *in vitro* approaches, such as x-ray and NMR analysis (chemistry perspective), and others still using techniques, such as fluorescence (biology perspective), to study protein behavior and function.[107] Despite investigating proteins from different perspectives, this community of scientists saw themselves as having a shared commitment to the dynamic concept,

[107] See, for example, Cooper and Dryden (1984) and McCammon et al. (1977) for physics approaches; Kendrew et al. (1958) and Perutz et al. (1960) for chemistry approaches; and Weber (1975) and Demchenko (1986) for more biological approaches.

as evidenced by their attendance at conferences, such as the 1982 Cibia Foundation symposium on "Mobility and Function in Proteins and Nucleic Acids" (Porter et al. 1983). Furthermore, scientists working within a particular perspective can disagree over the relevant concept of the protein. For example, although much early work in x-ray crystallography was carried out by scientists who endorsed the structural concept, Frauenfelder used this method in clever ways in the 1970s and 1980s to investigate the dynamic features of proteins, and he was an early champion of the dynamic concept (see, e.g., Frauenfelder et al. 1979).

Mitchell's scientific perspectives are clearly not at the same level of analysis as the concepts of the protein in my account. Our different units of analysis, in part, explain our different conclusions about protein science. Mitchell argues that perspectives cannot be globally unified or reduced but instead typically require integration.[108] In contrast, I have argued that a process of conceptual replacement, which yields quasi-unification, is occurring. Although Mitchell's analysis of perspectives is different from my analysis of concepts and conceptual change, I will argue in the next section that her analysis of scientific representations and the relationships between them is compatible with and complementary to my analysis of structural and dynamic representations of proteins.

4.4.3 Relating Representations of Proteins: Abstraction and De-idealization

The mere fact that conceptual change is leading to the replacement of the structural concept with the dynamic concept does not tell us much about the epistemic relationship that holds between the representations associated with each concept. The elimination of the structural concept does

[108] On Mitchell's view, successful reductions and unification happen in science, but only locally.

not entail the elimination of all the old structural representations of proteins. Instead, because the protein concepts are put to many uses, they can support many representations, (as evidenced by the many examples discussed in Section 2). Since the new dynamic concept has a more expansive inferential role, it supports more types of representations of proteins that are useful for explaining and intervening upon protein behavior.[109] In other words, it generates more successful inferences in the relevant scientific domains. I argue that most structural representations of proteins, associated with the old concept, are still included in the set of representations of the dynamic concept because these representations are related via abstraction. Any scientific representation of proteins will necessarily be partial, so the fact that these representations omit dynamic features of proteins does not make them incompatible with the dynamic concept. In what follows, I will develop the argument that structural and dynamic representations of proteins are often related via abstraction and idealization. My account, I argue, explains why the conceptual change in protein science has not eliminated old structural representations of proteins as inadequate and also why these structural representations will still be useful to scientists in certain contexts.

To begin, we must first briefly consider the partiality of scientific representations. Mitchell's work on this topic offers an excellent starting point for my analysis of protein representations.[110] She argues that all scientific representations and models are partial, representing certain aspects of the phenomenon of interest while omitting others. This partiality arises from a selection process, in which a researcher chooses which features of the phenomenon to include in

[109] It is not merely that the inferential role of the dynamic concept is more expansive than that of the structural concept that makes it a more useful scientific concept. Rather, my argument turns on the fact that the inferential role is larger in the *relevant* domain—i.e., in cases that are actually instantiated. The inferential role of the structural concept is, in a sense, oversimplified—it is too limited and includes false beliefs (e.g., structure alone determines function)—and for this reason it cannot support the dynamic representations that are needed to explain certain protein behaviors.

[110] My presentation of Mitchell's work here primarily draws upon Mitchell and Gronenborn (2017) and Mitchell (2019).

the representation and which to omit. In addition, some of this selection will be unintentional. For instance, a lack of knowledge or a constraint of a given representational medium might prevent the inclusion of a particular feature. However, according to Mitchell, this partiality is not a shortcoming. Scientific representations must be incomplete in this way if they are to be useful tools for reasoning and explanation. If they were complete rather than partial, they would represent all the features of the phenomenon in a one-to-one mapping relation between the model and the world.[111] Hence, they would no longer be practical scientific models but instead would be exact replicas of the phenomenon. Mitchell concludes that perspectival pluralism in science is an inevitable consequence of the nature of representation: scientific models and representations are partial, and scientists will require multiple models in order to represent different aspects of the same phenomenon.

As Mitchell (2019) shows, this argument is directly applicable to the case of protein representations: they are also inevitably partial. Consider, for example, early structural representations of proteins, such as Kendrew's models of myoglobin (see Fig. 6). His models represent only structural features of the protein, with the earliest models showing the position of the polypeptide backbone and later models including more detail regarding the location of the amino acid side chains. However, all dynamic properties of the protein are omitted from the representation, arguably for two different reasons. First, x-ray crystallography does not reveal information about protein dynamics, since it requires protein molecules to be removed from

[111] Mitchell (2019) also argues that this ideal of a "complete representation" of a phenomenon, such as a protein, is unattainable because our empirical and theoretical access might engage with different aspects of the phenomenon. For instance, the two most common techniques of protein structural determination, x-ray crystallography and NMR, interact with different aspects of the protein—i.e., the electron cloud and nucleus, respectively. The representations of protein structure from both these methods will often disagree in certain respects, so a unified model of protein structure that simply and straightforwardly combines both representations is impossible. Instead, Mitchell claims, more complex process of integration must occur.

solution and packed into a crystal. Thus, Kendrew inevitably excluded the dynamic properties of myoglobin because his experimental method did not give him access to the molecule's dynamics. But even if he had access to the dynamic properties of myoglobin, he likely would have excluded them from the representation, since molecular biologists at that time were committed to the structural concept. This concept of the protein involved the commitment to the structure-function rule, which maintained that structure alone determined protein function. The user of this concept would likely have deemed most dynamic properties of proteins to be irrelevant to protein function and therefore excluded them from the representation of the protein.[112]

Even though they were developed when the structural concept of the protein was dominant, structural representations of proteins like Kendrew's models of myoglobin are nevertheless compatible with the dynamic concept. All the dynamic features have been abstracted away from the protein. But no matter the reason for the omission, structural representations are not false because they lack dynamic properties. Instead, they are merely partial. The change in inferential role of the concept that occurred with the shift from protein_S to protein_D jettisoned the structure-function rule and the idea that proteins have stable, static 3D structures, but it did not deny that proteins have structure or claim that structure is irrelevant to protein function. Instead, the inferential role of the new concept involves the belief that dynamics as well as structure are relevant to protein function. Because all representations are invariably partial, those representations associated with the structural concept do not need to be interpreted as representing structure as the sole cause of protein function. Hence, they do not represent the central feature of

[112] I do not mean to attribute this view to Kendrew or any particular historical actor. However, it seems plausible that many structural and molecular biologists in the mid-twentieth century considered the fast timescale motions at the heart of the dynamic protein concept to be largely irrelevant to protein function (see, e.g., Monod (1968) for suggestive evidence of this point).

the structural concept that has been rejected by advocates of the dynamic concept.[113] They are therefore compatible with the dynamic concept, since they can be interpreted as partial representations of proteins with their dynamic features abstracted away. This fact explains why structural representations of proteins persist, despite the conceptual change from protein_S to protein_D.

The argument from the partiality of representations shows why the dynamic concept is compatible with both structural and dynamic representations. Analysis of those representations demonstrates that they are typically related via abstraction. That is, one can transition from a structural to a dynamic representation or vice-versa through the process of de-idealization or abstraction, respectively. Many of the representations of the dynamic protein arise rather straightforwardly from the addition of dynamic information to structural representations. Recall the superposition model of hemoglobin that Frauenfelder and colleagues developed to represent the dynamic structural fluctuations the protein undergoes in solution (see Fig. 8A). This representation overlays dynamic properties of the molecule onto the average structure, thereby unifying the structural and dynamic features in one more representationally complete model of hemoglobin. This process is typical of the superposition family of dynamic representations. Similarly, Hilser and colleagues' (2006) energy landscape ensemble representation of protein molecules in solution overlays thermodynamic information over the 3D structures of the microstate conformations (Fig. 10).

[113] An advocate of the dynamic concept could nevertheless argue that these structural representations are misleading, but it would be a mistake to say that they are incompatible with the contemporary understanding of proteins captured in the dynamic concept.

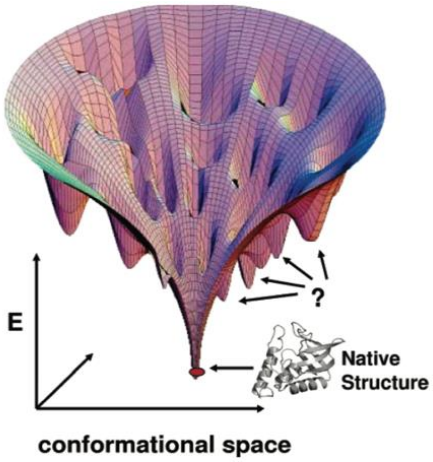

Figure 10 Energy Landscape Ensemble Representation (Hilser et al. 2006). Hypothetical energy funnel with local minima labeled with structures (only native structure shown). Republished with permission from Hilser et al. (2006); permission conveyed through Copyright Clearance Center, Inc.

The single 3D structure, associated with the structural concept, remains in the new dynamic representation, but other accessible structures are represented along with the energy landscape that conveys information about the accessibility and stability of these various structures. These examples show that many of the dynamic representations associated with protein$_D$ are more complete than previous structural representations, since no structural features that were previously represented are now omitted.[114]

Unification is possible in this case because representing dynamic features of proteins does not preclude representing structural features. Mitchell (2019) acknowledges that features represented in different models can be combined to "form a single, more complete model," if the features are "causally independent, partitionable into distinct subfeatures, or neatly mereologically nested" (180). In this case, I argue, the structural and dynamic features of proteins are largely

[114] Note that the representational "completeness" I invoke here is only relative: a representation of protein P associated with the dynamic concept that includes structural features A and B as well as dynamic features C and D will be more complete than a representation of P associated with the structural concept that includes only A and B, or only A, or only B.

partitionable. The relationship between these new more complete representations, which include structural and dynamic features, and the older structural representations is one of abstraction. To move from the dynamic representation back to a structural one, one need only abstract away the dynamic features from the representation to leave a static structure.

The relationship between the dynamic and structural representations is actually slightly more complicated. Although the abstraction relation captures the epistemic relationship between dynamic and structural representations, something is lost in the transition—viz., there is no way via abstraction to identify a unique 3D structure of the protein. That is, there is no epistemically principled way to privilege one particular 3D structure over any of the other structures included in the ensemble representation or in the superposition of structures representation.[115] In many cases of abstraction, certain features can be abstracted away because they are irrelevant to the feature one wants to represent in abstract or idealized form. Lockean abstraction of particulars works in this way (Rosen 2020). I can abstract away the material in which I find a triangle—for instance, whether it is drawn on a board or in the sand or made from three pieces of wood—to produce an abstract or idealized "triangle" precisely because the material from which it is made is irrelevant to its "triangle-ness." We perform similar abstractions when we judge certain features to be causally or explanatorily irrelevant (Sober 1984). For instance, in explaining why some balls fit through a certain hole but others do not, we can abstract away and ignore the color of the ball in our explanation, since we judge the color to be causally irrelevant. Indeed, everything but the diameter of the ball can be abstracted away in this way.[116]

---

[115] Because an x-ray crystal structure averages over all the structures in the crystal, it is rather like Quetelet's average man, so "it is highly improbable that at any instant in time one individual protein molecule has the average structure" (Karplus and McCammon 1986, 42).
[116] This example is from Sober (1984, p. 99).

But the structure and dynamics of a protein are not properties such that one can be abstracted away from the other in this way. The two properties are codependent. There is no sense in which dynamics could be said to be irrelevant to structure: the dynamic properties of a protein determine the range of structures it can adopt, and the structural features determine its range of dynamic flexibility. Although we can easily represent proteins in such a way that we reveal their structural features and hide their dynamic ones (indeed, this is the most common and simplest way to represent proteins), we cannot represent a singular structure of the protein as it would be without dynamics. There is no representation that would capture the abstract or idealized "pure structure" of the protein in a way analogous to the Lockean triangle, since no protein has a singular structure because of the inevitable dynamic properties.[117] Thus, although one can abstract away the dynamics to convert a representation of the structural fluctuations of a protein into a static, structure without any fluctuations, the process of abstraction alone cannot be used to identify any one structure as *the* structural representation of the protein.

My account of the relationship between structural and dynamic representations of proteins explains the persistence of structural representations despite conceptual replacement and also helps explain their usefulness in certain contexts. I agree with Mitchell that different scientific representations, which are inevitably partial, will be useful in different contexts. Because structural representations are related to dynamic representations via abstraction, it is not difficult to imagine contexts in which these simplified representations will be useful.[118] For example, structural

---

[117] Here, I am not simply arguing that structure is context dependent. Even when we fix the relevant functional context—e.g., in a particular solution, in a particular cell type, etc.—the structure is still dynamic and therefore better represented as a multiplicity (or ensemble) of structures rather than a single 3D structure. The structural variation within the ensemble will of course vary depending on the protein and the context.

[118] It is reasonable to think of structural representations as simpler than the dynamic representation from which it is derived, since the former represents primarily one relevant feature (structure) while the latter represents two (structure and dynamics).

representations are still productively used in pedagogical settings. Structural representations figure in explanations of protein behavior that show up in textbook accounts of allostery (e.g., Garrett and Grisholm 2011). Unlike accounts of allostery that use ensemble representations of the dynamic concept, which require thermodynamic and equilibrium thinking (e.g., Hilser et al. 2012), these structural accounts involve mechanistic reasoning, which is more in line with folk physics and therefore easier for students to grasp. One of the early proponents of the dynamic concept of the protein, recognized this difference and cited it as a reason why explanations that used the dynamic representations associated with the dynamic concept were slow to be adopted:

> The concept that "conformational changes" are involved is widely accepted and is a comforting view since we might then picture the mechanism as similar to the relative motions of springs and levers and pulleys familiar from macroscopic machines. (Cooper 1980, 474)

Morange (2012) reiterates this point, specifically in the context of teaching. He claims that "everyone who has taught" biochemistry and structural biology knows that mechanistic explanations of protein behavior, which invoke the structural representations of the protein, are "understood much more rapidly" than thermodynamic explanations that use dynamic representations (Ibid., 14). Structural representations of the protein abstract away their dynamic features, but they can still help students and laypeople gain an understanding of protein behavior. Furthermore, these representations can be used to develop successful scientific explanations of protein behavior insofar as the proteins being studied approximate the static structures posited in the representations.

In sum, I have argued that abstraction and de-idealization can often make sense of the relationship between structural and dynamic representations of proteins. My position is largely

compatible with Mitchell's (2019) account.[119] I have shown that the strategies of abstraction and de-idealization are essential to understanding the epistemic relationship between the representations associated with the two protein concepts. Generalizing from this case, I suspect that local unifications of representational outputs are common: the models and representations of the dynamic protein, which overlay dynamic features on structural representations, are a case in point.

4.4.4 Philosophical Insights and Morals

My analysis of concepts and representations of proteins has certain philosophical payoffs. For one, it provides a justification for the intuition that the dynamic concept of the protein provides a better way of explaining, predicting, and intervening on protein behavior than the previous structural concept. Relatedly, it also suggests that scientists should prefer more permissive and inclusive scientific concepts—i.e., ones that support more representations rather than fewer. And finally, the history of protein science highlights certain problems that can arise with the standard scientific strategy of beginning from a simpler representation—i.e., one that is more abstract or idealized—and later attempting to de-idealize to develop a more complete representation. In what follows, I will explore each of these three insights in turn.

We might intuitively think that the new dynamic concept of the protein marks an improvement over the old structural concept, in part because we now know the two beliefs that

[119] Our accounts are compatible because we focus on different scientific procedures for combining features from different representations: Mitchell (2019) focuses on understanding strategies of integration whereas I focus on abstraction. Similarly, she focuses her analysis on features that are not partitionable into subfeatures, whereas I have argued that structural and dynamic features largely fit this description and hence can be combined into a unified representation.

form the core of the structural concept's inferential role are false or misleading. It seems better to begin reasoning with a concept of the protein that from the start encourages scientists to attend to two causally relevant features of protein behavior (structure *and* dynamics) rather than just one (structure), especially when we have good empirical evidence for the prevalence and importance of that second causal feature (dynamics).[120]

Brigandt's (2010) account of conceptual change in science helps to confirm this intuition. Using his account, we can demonstrate that the change that occurs with the replacement of protein$_S$ for protein$_D$ is rational. Recall that, according to Brigandt, concepts have three components of semantic content: reference, inferential role, and epistemic goal. My analysis has already taken into account the first two components. However, it is this final component—"the epistemic goal pursued with the concept's use" (Ibid., 19)—that enables us to judge the rationality of changes in the concept's inferential role, or meaning. The primary epistemic goals that govern the use of the protein concept by molecular and structural biologists has remained the same since the heyday of the structural concept in the 1950s and 1960s. These scientists use the protein concept primarily because they aim to explain, predict, and potentially intervene upon protein function and behavior (see Table 1).[121] These same general epistemic goals hold for those who endorsed the structural concept as well as those who advocate using the dynamic concept. According to Brigandt's account, we can use the stability of the epistemic goals across conceptual change to judge the rationality of that change. If the new concept lets scientists better achieve the epistemic aims that

[120] The discovery of the ubiquity of functionally important intrinsically disordered (ID) proteins (i.e., proteins without stable 3D structure) as well as proteins with ID regions has highlighted the importance of dynamic properties of proteins (Oldfield and Dunker 2014).

[121] Brigandt (2010) argues that "epistemic goals can be assigned to those central concepts (at least in biology) that underwent conceptual change, such that this semantic change can be explained in these terms" (23).

motivated their use of the concept in the first place, then the change is warranted, thus making the conceptual change rational.

The shift from $protein_S$ to $protein_D$, I argue, fits this pattern: the new concept promotes the achievement of the epistemic goals that motivated its use. First, the dynamic concept, with its associated representations, enables scientists to explain more cases of protein function. As we saw in the previous section, the dynamic concept supports more diverse representations than the structural concept. Included among the set of representations associated with the dynamic concept (but not the structural one) are ensemble representations. Scientists have identified examples of protein behaviors, such as allostery and catalysis, that cannot be explained using structural representations of proteins but instead require explanations that use these ensemble representations (Hilser et al. 2012, Motlagh et al. 2014). The new concept thus improves the epistemic goal of explanation of protein behavior since it supports both the mechanistic explanations that invoke structural representations of proteins as well as these explanations that require dynamic representations of proteins.

Second, the dynamic concept also enables better predictions of protein behavior. In particular, the use of the thermodynamic framework, which identifies structural changes with enthalpy (ΔH) and dynamic changes with entropy (ΔS), enables scientists to predict when protein behaviors that have previously been explained using only structural representations will fail (Reinhart et al. 1989). These explanations abstract away dynamics and thus fail to consider the role of entropy. They are therefore likely to incorrectly predict protein behavior in cases in which there are large changes in entropy, or in cases in which the relative contribution of entropy increases, for instance when the temperature increases. By accounting for entropy, the dynamic concept marks an improvement over the structural concept in predicting protein behavior.

Finally, these improvements in the ability to explain and predict protein behavior, which occur when scientists use $protein_D$ rather than $protein_S$, further enhance scientists' ability to intervene. For example, antiretroviral drugs that bind to HIV proteins are more effective if they complement the structure of the viral protein but also take into account changes in the dynamic properties of the protein that occur upon binding the drug (Neal 2019). Thus, the dynamic concept provides these biomedical researchers a better starting point for considering the relevant causal factors. Having considered all three of the primary epistemic goals of the protein concept, we can conclude that this change in inferential role that accompanies the shift from the structural to the dynamic concept is rational, "since the revised inferential role [of the dynamic concept]…meets the epistemic goal to a higher degree than the previous inferential role [of the structural concept]" (Brigandt 2010, 24). My analysis confirms our intuitions and gives us good reasons to side with the scientific advocates of the dynamic concept of the protein.

We can also draw two tentative philosophical morals about scientific concepts and representations from this case study of rational change in protein science. Let us first consider what the protein case can tell us about what makes one concept better than another. All the features of the dynamic concept that make it a better tool for scientists who aim to explain, predict, or intervene on protein behavior are related to the fact that the inferential role of this protein concept is more inclusive than that of the structural concept. It guides the scientist to consider structure and dynamics as co-determinants of protein function whereas the inferential role of the structural concept led to exclusive focus on structure. From this, we might generalize that a concept that guides scientists' attention to more rather than fewer relevant causal factors and consequently supports the development of more diverse representations will prove a better starting point for

explanation, prediction, and intervention.[122] Although representations are necessarily partial, concepts are not. There are no in principle limits on what the inferential role of a concept can contain. Even though scientific users of more inclusive concepts will still have to engage in a selection process when they set out to represent the phenomena, these scientists will be more aware of the partiality of their representations and the features of the phenomena omitted.

This case can also tell us something about scientific representations. I argued in the previous section that a scientist could typically switch between a structural and dynamic representation via abstraction or de-idealization. The process of abstracting dynamics from dynamic representations that include both dynamic and structural features of proteins—e.g., Frauenfelder's superposition representation of myoglobin—is relatively straightforward. The reverse process of de-idealization is often not so simple. For example, developing a dynamic representation from a structural one may require the development of completely novel representations. Hilser's energy landscape ensemble representation is a case in point (see Fig. 10). Unlike abstracting the dynamic features from Frauenfelder's representation, the development of this dynamic representation from a structural one is a scientific achievement. At the representational level, abstraction is typically an easier process than de-idealization. Beginning from a more inclusive concept that guides scientists to consider more causally relevant features from the outset might prove an easier strategy than attempting to de-idealize from a partial representation.

Reasoning from the structural concept poses another risk: the scientist who develops an explanation of protein behavior using structural representations of the protein may think that if an

[122] Of course, this can be no more than a general heuristic. Much will likely depend on the particular context in which we want to explain or intervene

explanation of this sort can be given, then the dynamic features of the protein do not matter.[123] Advocates of the dynamic concept have long argued that this is not so (Cooper 1980, Reinhart et al. 1989). The dynamic concept tells us that there is no principled reason to privilege structural features in this way, thus even if a structural explanation can be given, it may still turn out that dynamics are the driver of a particular protein behavior. The general moral here, then, is that the strategy of beginning with the simplest representation and then de-idealizing as needed can pose an epistemic risk, but reasoning from a more expansive concept can help ameliorate this risk.

4.5 Conclusion

In this chapter, I have conducted a systematic analysis of protein concepts and representations in order to understand the conceptual change in protein science as well as the relationship among the diversity of representations of proteins. I have distinguished two concepts of the protein—protein$_S$ and protein$_D$—and have argued that former is being replaced by the latter. At the level of representations, I have argued that the relationship between structural and dynamic representations is best understood as one of abstraction and de-idealization. Because these two features of proteins are largely partitionable, it is possible to either combine them into more complete representations of proteins or to abstract one away leaving a more partial representation. By distinguishing the conceptual and the representation levels, my account explains why the dynamic concept, which is a more expansive concept supporting more representations of proteins,

[123] There are interesting parallels between this example and debates about neutral selection in evolutionary biology. Advocates of neutral selection have argued that we should not begin with adaptationist accounts and then only consider the possibility of neutral selection when those fail (Koonin 2016). Instead, we should consider the possibility of either neutral selection or adaptations from the beginning.

is a better conceptual tool. It further explains why replacement at the conceptual level does not lead to similar replacement of structural representations at the representational level.

Using Brigandt's (2010) account of conceptual change, I have shown that the shift from protein$_S$ to protein$_D$ is part of the rational development of protein science, since the use of this new concept enables scientists to better achieve their epistemic goals. I have also drawn two philosophical morals. This analysis, I suggest, demonstrates the benefits of reasoning with more inclusive scientific concepts, and it has also revealed certain epistemic risks associated with the reductive approach of beginning with the simplest representation and de-idealizing as needed. Although these philosophical morals are tentative since they have been drawn from this one case in protein science, they would be worth investigating in other areas of molecular and structural biology.

5.0 Ensemble Explanation of Protein Function: A Philosophical Account of a Novel Type of Explanation in the Biological Sciences

5.1 Introduction

Philosophers of science and biology have argued that mechanistic explanation is the dominant type of explanation in the biological sciences, especially in the molecular life sciences (Machamer et al. 2000, Craver 2007, Bechtel and Richardson 2010). On this view, scientists explain by decomposing biological systems into molecular components and then describing how these lower-level components interact to produce the higher-level behavior of the system. Indeed, uncovering the molecular mechanisms responsible for heredity, metabolism, and diseases, such as cancer, has been one of the central explanatory goals of the biological sciences since the middle of the twentieth century. This trend of molecularization in the life sciences has led to a presumption that advances in biological knowledge will come from mechanistic explanations that drill deeper, contain more details, and invoke more highly refined molecular structures.[124] Within molecular and structural biology, this trend has been particularly pronounced. Since the discovery of the DNA double helix, champions of molecular biology have claimed that advances in structural techniques would render the 3D structures of DNA, RNA, and proteins visible, and from these high-resolution structures we would then be able to determine the functions of these molecules and explain their role in various biological processes (Monod 1965).

[124] See Kaplan and Craver 2011, but see also Kaplan and Craver 2020, in which they moderate their commitment to 'the more details the better' account of mechanistic explanation.

Recently, however, molecular life scientists have begun to develop an alternative type of explanation—ensemble explanation—to understand and explain protein behavior. Unlike earlier mechanistic explanations, ensemble explanations do not cite structural changes in individual protein molecules as the cause of protein behaviors, such as catalysis, binding, and allostery. Proponents of these explanations reject the structural view that maintains that a protein has a largely stable and rigid 3D structure, and instead adopt a dynamic view that holds that proteins in solution are undergoing constant structural fluctuations. Whereas mechanistic explanations focus on changes in the average structure of a protein molecule, ensemble explanations consider changes in the thermodynamic properties of a population of protein molecules in solution. They use a novel ensemble representation of proteins to capture both their structural and dynamic properties.

In this chapter, I develop an account of ensemble explanation of protein behavior. These explanations deserve philosophical attention, since they are becoming increasingly common in structural and molecular biology as scientists recognize the importance of dynamic properties of proteins (Motlagh et al. 2014; Morange 2017, 2020). My account shows that scientists have not simply modified mechanistic explanations to include protein dynamics. Instead, they have developed ensemble explanations as a novel type of explanation to deal with the added complexity of protein dynamics.

In what follows, I first present my account of ensemble explanation of protein behavior (Sec. 5.2). To capture protein dynamics, these explanations represent a population of protein molecules as an ensemble of structurally distinct microstates. They cite changes in the equilibrium distribution of protein molecules across microstates, caused by perturbations, to explain protein functions. After developing my account, I then compare ensemble explanations to two different philosophical accounts of explanation: equilibrium explanation and causal explanation in

thermodynamics. I argue that these explanations in protein science do not fit the model of equilibrium explanations discussed in the philosophical literature (Sec. 5.3). The equilibria in ensemble explanations—i.e., the equilibrium distributions of microstates—do not play the right explanatory role. Instead, I argue that ensemble explanations are a species of causal explanation (Sec. 5.4), similar to but distinct from causal explanations in thermodynamics. Their major innovation, however, is their difference from mechanistic explanations and their focus on protein dynamics as well as structure to explain protein function.

5.2 Account of Ensemble Explanation

Ensemble explanations of protein behavior are a new type of explanation in structural and molecular biology. In a field dominated by mechanistic explanations, these explanations are strikingly different. Whereas mechanistic explanations of protein behavior focus primarily on the relationship between structure and function, ensemble explanations capture the role of dynamics as well as structure in bringing about protein function. To represent protein dynamics, ensemble explanations shift the primary unit of analysis from an individual molecule to a population or ensemble of molecules. Moreover, these new explanations include thermodynamic reasoning typically missing from mechanistic accounts.

To characterize this new type of explanation, I will focus on allostery, which is one of the simplest and most general protein behaviors with important regulatory functions (Cui and Karplus 2008). Allostery, as traditionally defined, is the process whereby binding of a ligand at one site on a protein affects binding of another molecule at a distal site. Beginning in the 1960s, scientists developed mechanistic explanations to account for this seeming action at a distance. These

mechanistic explanations of allostery cited a structural, or conformational, change in a protein molecule upon binding the first ligand that alters the second binding site, which in turn affects its affinity for the second ligand (Monod et al. 1963, 1965; Koshland et al. 1966). These explanations of allostery fit within the dominant static view of proteins, since an allosteric protein was assumed to have only two discrete structures or conformations.

Scientists have recently developed alternative ensemble explanations of the same phenomenon (Hilser et al. 2012, Motlagh et al. 2014). These scientists reject the static view of proteins and instead endorse the new, dynamic view, which stresses the fact that a protein molecule does not have a stable 3D structure in solution but instead undergoes constant motion (Cooper 1976, 1980). A population of protein molecules thus will not all adopt one or two discrete structures, as the mechanistic models assumed. Instead, individual molecules will be wandering through conformational space, and the population at any instant in time will include a multiplicity of heterogeneous structures. While mechanistic explanations used an idealized representation of the protein as a static molecule with stable structure, ensemble explanations were intentionally designed in order to capture the dynamic properties of protein molecules. In these explanations, a population of protein molecules is represented as an ensemble of structural microstates. The distribution of microstates is energetically-weighted, with more thermodynamically stable microstates more populated at any moment in time than less stable microstates. Allostery, according to this account, is not a property of an individual protein molecule but rather a property of the ensemble. An ensemble explanation of allostery therefore does not cite structural changes in a protein molecule but instead cites changes in the distribution of microstates that occur upon perturbation by the allosteric ligand.

Let us look more closely at the two major features of these explanations, beginning with the microstates. Each microstate within the ensemble has only three explanatorily relevant properties. First, each microstate has a thermodynamic stability, given by the Gibbs free energy, ΔG. For any microstate j, ΔGj is an individual property of the microstate, not dependent on the other microstates.[125] At a given temperature and pressure, the ΔG for the microstate is fixed. It will depend upon structural features of the protein, such as the interaction among amino acid sidechains in the hydrophobic core of the protein as well as the interaction between amino acid sidechains and the solvent (e.g., water). For example, a microstate structure with a hydrophobic region exposed to the solution will be less stable than another conformation with only hydrophilic residues exposed, all things being equal.

Second, each microstate will have binding affinities for various ligands. These will be described by the dissociation constants, K_D, which measure the rate at which the protein-ligand (PL) complex dissociates at equilibrium: $PL \rightleftharpoons P + L$. K_D is given by the following equation

$$K_D = \frac{[P][L]}{[PL]},$$

in which the protein-ligand complex [PL], free protein [P], and ligand [L] are concentrations in moles. The smaller the K_D for a particular ligand, the more strongly it binds the protein. To exhibit allostery, a protein must be capable of binding at least two ligands. In the most idealized case, the binding affinity for the two ligands is considered to be qualitative (Fig. 11). [126] Each domain in the

[125] The distinction I am drawing here does not map on to previous philosophical accounts of properties, such as Locke's primary and secondary properties or Aristotle's essential and accidental properties. Nor does it map on to the scientific description of intensive and extensive properties. Instead, it maps onto the distinction between individual and relational properties that Sarkar (2008) introduces in his analysis of frequency-dependent selection. In this case, what makes a property an individual property is the fact that it does not require reference to any other microstates or other properties of the ensemble.

[126] This figure depicts an example of heterotropic allostery, in which the allosteric ligand and the binding ligand are different. However, this account could also be used to explain homotropic allostery, in which the allosteric and binding

R (relaxed) state will, by definition, have a high affinity binding site for its ligand ($K_D \approx 0$), whereas in the T (tensed) state, the domain is assumed to have no affinity ($K_D \approx \infty$).[127] Like ΔG, the binding affinity of a microstate for a given ligand is an individual property, independent from the properties of the other microstates. It is fixed for a given protein-ligand complex, since it depends only on the electrostatic and steric features of the ligand and the binding domain.

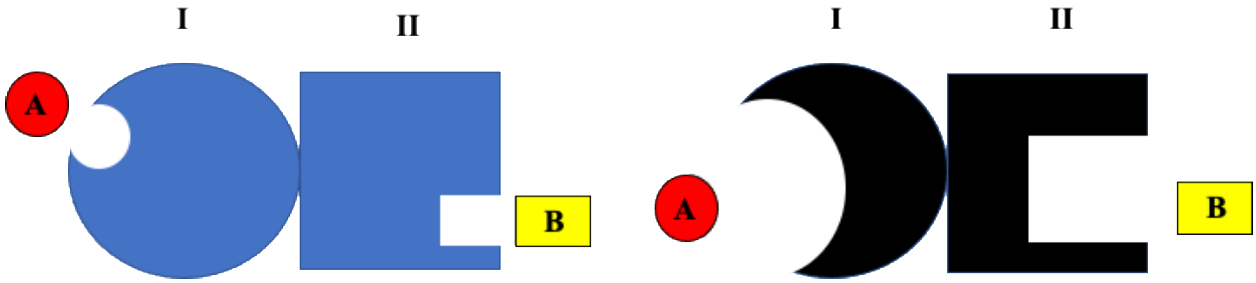

Figure 11 Two Domain Allosteric Protein. In the RR state (left, blue), domain I binds Ligand A (red circle) with high affinity and domain II binds Ligand B (yellow rectangle) with high affinity. In the TT state (left, black), each domain has no binding affinity for either ligand.

The third relevant explanatory property of each microstate is its probability, Pj. The probability of a microstate describes its frequency in the given ensemble. It depends on the relative thermodynamic stability of the microstate compared to the other microstates in the ensemble. For instance, a microstate with a lower free energy will be more stable than one with a higher free energy and thus will be populated by more protein molecules at any given time. Therefore, the probability of a given microstate Pj is not an individual property but rather a relational property. The probability for any microstate is only fixed within a given ensemble distribution. In other words, it depends on the state of the ensemble distribution. This dependence is evident from the mathematical representation:

---

ligands are the same (e.g., O_2 in hemoglobin binding). Using the heterotropic case as the exemplar makes it easier to distinguish the two ligands and their respective roles.

[127] These assumptions about K_D are examples of Galilean idealization (McMullin 1985). They are easily de-idealized using the mathematical formalism developed alongside this qualitative description of ensemble explanations.

$$P_j = \frac{S_j}{Q}.$$

The value of the probability Pj for a given microstate is determined by the statistical-weighted free energy of that microstate (Sj) and the partition function (Q). While the former depends only on the thermodynamic free energy of that microstate (ΔGj), the latter is the sum of all the weighted microstates in the ensemble:

$$Q = \sum_i S_i .$$

It thus depends on the thermodynamic properties of all the other microstates in the ensemble and describes the ensemble distribution. Changes in the free energies of other microstates in the ensemble will perturb the ensemble distribution, which will in turn affect the Pj for each microstate.

The ensemble itself is the other major feature of ensemble explanations. Imported from statistical mechanics, this representation provides a novel approach to explaining protein behavior. By representing the structural heterogeneity of protein molecules in solution, it enables these explanations to take into account the causal role of dynamics in bringing about protein functions. Using this representation requires two idealizing assumptions about the dynamics of protein molecules in solution. Most importantly, the interpretation of the ensemble as representing protein dynamics arises from the ergodic hypothesis of statistical thermodynamics (Frigg 2008). Applied to this particular case, the ergodic hypothesis tells us that the instantaneous ensemble distribution that shows the weighted distributions of molecules across microstates at any given moment mirrors the temporal evolution of an individual protein molecule wandering through conformation space over time. The assumption of ergodicity shows why we can interpret the ensemble as encoding the dynamic properties of a single protein molecule. Thus, any explanation that cites changes in the

ensemble to explain observed protein behavior will necessarily also include information about the dynamic properties of that particular type of protein. The second assumption is that a given population of protein molecules in solution, unless perturbed, will be at thermodynamic equilibrium. Although individual molecules will be constantly interconverting between microstate structures, the population of molecules as a whole will preserve the equilibrium distribution of microstates. [128]

Given these assumptions, an ensemble explanation can explain almost any protein behavior by citing a change in the equilibrium distribution of microstates that occurs upon perturbation of the system and then comparing the associated changes in the probabilities of microstates with various properties. To give an ensemble explanation of the allosteric response to a ligand, for example, one must identify the relevant microstates and determine their three explanatorily relevant properties—e.g., thermodynamic stabilities, ligand affinities, and probabilities. This information alone is sufficient to describe the preexisting ensemble equilibrium distribution (Fig. 12, left). The addition of the allosteric ligand then perturbs the preexisting equilibrium, by preferentially binding and stabilizing certain microstates. As depicted in the dotted arrows in Figure 12, binding of the allosteric ligand effects a change in the thermodynamic stability of each microstate, increasing the stability (lowering the ΔG) for all microstates it binds.

[128] Both of these assumptions—ergodicity and equilibrium—are background assumptions necessary for ensemble explanations. In other words, one has to accept that the biological systems of interest in this case are typically ergodic and that populations of protein molecules are in approximate equilibrium unless perturbed. An ensemble explanation itself does not explain why these assumptions are legitimate. It does not, for example, provide an explanation for why we can use the ergodic hypothesis or explain any other features of the foundations of statistical mechanics. Instead, it imports these background assumptions from physics, along with the ensemble representation, and uses them in the explanation of protein behavior.

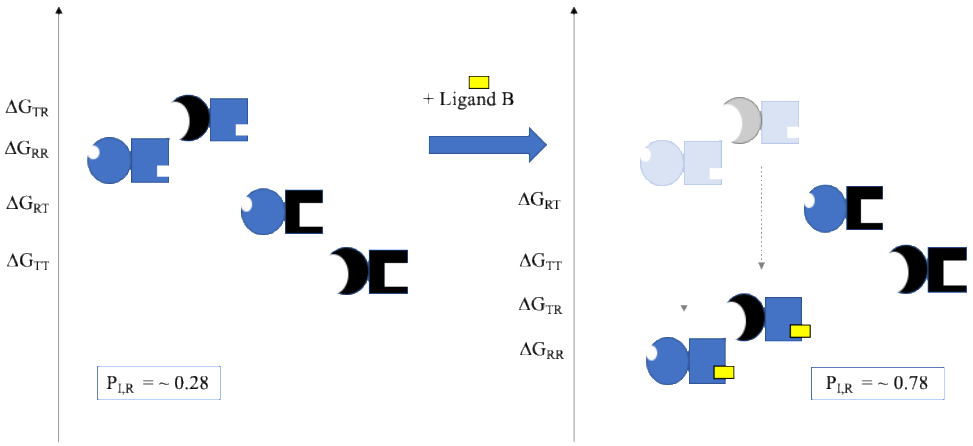

Figure 12 Example of Ensemble Allostery. Following Figure 11, domain I is represented by a circle and domain II by a square. Blue domains are in the R state, capable of binding ligand A or B, for domains I and II, respectively. Black domains are in the T state, incapable of binding either ligand. Left panel shows the preexisting equilibrium of the three functionally-defined microstates. The probability of the microstates that are capable of binding ligand A (i.e., domain I in the R state: RR and RT) is 0.28. Right panel shows the redistribution of the ensemble caused by the addition of ligand B, which binds and stabilizes microstates RR and TR. Ligand B is a positive allosteric effector, since the probability of microstates that can bind A (i.e., RR and RT) increases: $P_{I,R}$ = ~0.78. Adapted from Hilser et al. 2012, Fig. 3, p. 592.

The thermodynamic stabilization of those microstates, leads to changes in their probabilities, which in turn leads to a redistribution of all the microstates in the ensemble (Fig. 12, right).[129] To explain the observed allosteric effect, the new equilibrium is compared to the preexisting equilibrium. When the addition of the allosteric ligand leads to an ensemble distribution in which the probabilities of the subset of microstates capable of binding the second ligand is higher than before, this change explains why the ligand is a positive allosteric effector (i.e., why the allosteric ligand increased the protein's affinity for the second ligand). In an ensemble explanation of allostery, it is the change that occurs upon perturbation to the equilibrium distribution of microstates that is cited to explain the observed allosteric effect.

[129] The return to a new equilibrium after perturbation occurs because of the law of mass action, in accordance with Le Chatelier's principle (Chang 2003). That is, the binding of the allosteric ligand to a particular microstate creates a new species, removing it from the previous equilibrium. Consequently, the previous equilibrium shifts in the direction of the microstate structure that that can bind the allosteric ligand. The process that governs the change in equilibrium will be further discussed in Sec. 4.2 in the context of the interventionist framework of causal explanation.

A further clarification is in order. Ensemble explanations of allostery can be thought to involve answers to two distinct why-questions. The first is the one just discussed: why does a particular protein system exhibit changes in affinity for a second ligand after the introduction of an allosteric ligand? The answer given by the ensemble explanation in this case is that the allosteric ligand drove the ensemble equilibrium distribution to a new equilibrium that favored microstates with higher affinity for the second ligand.[130] But one might also ask a second question of this system, namely, why does the allosteric ligand perturb the equilibrium distribution of microstates? The answer to this question invokes laws of chemistry that show that ligand binding leads to the stabilization of the protein. In this case, the binding of the allosteric ligand to a particular microstate stabilizes that microstate—reducing its ΔG—and thereby altering one of its three explanatorily relevant properties. When we apply this chemical fact to the ensemble representation, we see that this change in the free energy of those microstates that bind the allosteric ligand leads to a change in their probability (Pj) in the ensemble. Since the Pj for a given microstate is a relational property, this intervention on a certain subset of microstates to alter their Pj leads to a change in this property for all the microstates, which in turn constitutes a new equilibrium distribution of microstates.

5.3 Equilibrium Explanations

One of the major innovations of ensemble explanations is the use of the ensemble to represent protein dynamics. Within the ensemble, the distribution of protein molecules across

[130] A similar explanation can be given to account for changes in protein behavior caused by other factors: e.g., a change in temperature or pH causes a change in the behavior of a protein system because it perturbs the equilibrium distribution of microstates.

microstates is assumed to be at thermodynamic equilibrium, and perturbations to the system lead to a redistribution of molecules across microstates and the establishment of a new ensemble equilibrium distribution. Because of the central role of the ensemble equilibrium, it is worth considering whether ensemble explanations are a species of equilibrium explanation. Sober (1983) first characterized equilibrium explanation as a type of explanation in evolutionary biology, but it has gained attention more recently as philosophers of science have debated whether it is a type of causal or noncausal explanation.[131] In what follows, I first present the canonical account of equilibrium explanation, focusing primarily on those features that distinguish it from paradigmatic cases of causal explanation (Sec. 5.3.1). I then argue that despite their invocation of equilibria, ensemble explanations fail to fit this pattern of explanation (Sec. 5.3.2).

5.3.1 Canonical Account of Equilibrium Explanation

In "Equilibrium Explanation," Sober (1983) characterizes what he considers to be a novel type of noncausal explanation. He generalizes from R.A. Fisher's (1931) explanation for why many species reproduce with a one-to-one sex ratio at birth. According to Fisher, a population will reach and maintain a one-to-one sex ratio because it is a stable equilibrium: if at any time the population diverges from this ratio, parental pairs who over-produce the minority sex will be favored by natural selection. The action of natural selection will maintain a one-to-one sex ratio and will return any population that happens to deviate from this ratio back to one-to-one. The explanation only holds for populations that satisfy certain empirical assumptions—i.e., the parents

[131] Sober (1983) claims equilibrium explanation is an example of noncausal explanation but more recently, certain philosophers have argued it is a variety of causal explanation, see, e.g., Woodward (2003) and Strevens (2008).

must mate randomly and invest equally in daughters and sons, reproductive strategies must be variable and heritable, and the sex ratio must never drift to fixation. But for those populations that satisfy (or approximately satisfy) these assumptions, the one-to-one sex ratio at birth can be explained by appeal to Fisher's equilibrium explanation. What is curious about this type of explanation, according to Sober, is that it explains without appeal to actual causes.[132] Fisher's account explains the one-to-one sex ratio of a population without specifying the actual causal trajectory that led the population to this sex ratio.

Strevens (2008) offers another example of an equilibrium explanation, using a mechanical example. He aims to explain why a ball placed on the inside lip of a basin rolls back and forth and ultimately comes to rest at the lowest point at the bottom of the basin.[133] He agrees with Sober's analysis that the best explanation of this phenomenon will not cite the initial position of the ball and trace its actual trajectory through the basin, but instead will cite the dissipation of motion through friction as the ball traverses the basin, the net force of gravity that pulls it down to the bottom, and the absence of any net force acting on the ball when it reaches the lowest point in the basin. Although Strevens and Sober disagree as to whether equilibrium explanations such as these are causal, there is a broad consensus that these explanations have certain features that make them different from paradigmatic cases of causal explanation.[134]

---

[132] In a bit more detail, Sober claims that this equilibrium explanation does not cite the actual cause but rather a disjunction of three possible causes: (1 and 2) if, at an earlier time, the population overproduced males (or females), then selection favoring females (or males) brought the sex ratio back to one-to-one, or (3) if the population was already at one-to-one at an earlier time, then selection against parental pairs that overproduced either males or females were selected against to maintain this ratio. Fisher's explanation, according to Sober, cites this disjunction of possible causes to explain the sex ratio in a particular population at a given time.

[133] Strevens (2008), p. 267-8.

[134] In this section, I will sidestep the issue of whether paradigm equilibrium explanations are causal or noncausal. All that is required for my analysis is the acknowledgement that equilibrium explanations have certain features that distinguish them from paradigmatic cases of causal explanation that cite lower-level causal trajectories. When I refer to "paradigmatic causal explanation" or use similar terminology, I mean to draw a contrast between an equilibrium explanation like Fisher's that does not cite the lower-level causal trajectory of a population through time and an explanation of the same phenomenon that does map out the causes and forces acting on the population through time

From these two examples, we can identify the key features of equilibrium explanations that distinguish them from paradigmatic causal explanation. First, equilibrium explanations are "deep" explanations. According to Strevens, they "strip away vast quantities of apparently relevant, large-scale causal detail" in order to "account for the phenomena in terms of a small number of abstract properties of the generating systems' causal dynamics" (Strevens 2008, 137).[135] Second, equilibrium explanations do not cite the actual causal trajectory that brought about the explanandum phenomenon; that is, they do not trace the causal evolution of the system. Instead, they present a "disjunction of possible causal scenarios" (Sober 1983, 204). Although the actual causal scenario is included as one of the disjuncts, the explanation does not specify which. For example, any ball placed on the lip of Strevens' basin will have a particular trajectory as it travels around and down the basin, but the equilibrium explanation for the ball's coming to rest at the bottom will not specify that path. It will instead cite the fact that the ball was placed somewhere on the lip of the basin along with certain properties of the basin and the relevant physics.

Third, paradigmatic cases of equilibrium explanation, such as the two described above, have a single, global equilibrium. For instance, according to Fisher's explanation no matter where a population starts—i.e., whether it produces more daughters or sons—it will eventually end in the same place, namely, at a one-to-one sex ratio. Less paradigmatic cases may lack a global equilibrium, but they will still have a wide basin of attraction. Equilibrium explanations require either global equilibria or equilibria with wide basins of attraction, since it is this structural feature

---

to show the production of the sex ratio at a later time. It seems, even most philosophers who claim equilibrium explanations are causal would accept this rather minimal distinction (see, e.g., Kuorikoski 2007, Sperry-Taylor 2019, Woodward 2003, 2018).

[135] The depth that is relevant to equilibrium explanations is Strevens' second axis of depth. (The first requires drilling down to fundamental physics.) According to Strevens (2008), explanations with depth along this second axis share "an unusual degree of generality" (137).

of the system that enables the explanation to omit information about the precise initial conditions and particular causal mechanism. The more local the equilibrium becomes the more causal detail must be included, making these explanations less distinct from standard causal explanations (Sober 1983).

Fourth, according to most philosophical accounts, equilibrium explanations work by first describing the structure and dynamics of a system that includes an attractor and then providing sufficient information about the initial conditions to show that they are likely to fall within the domain of attraction. Taken together, these features explain why a particular system is at or tends toward equilibrium. It is precisely these structural and dynamic features of the system that enable equilibrium explanations to omit so much causal detail and explain abstractly (Woodward 2018). For instance, Fisher's explanation of sex ratios at birth explains this phenomenon because it describes the structure of the population and the dynamics which demonstrate the one-to-one sex ratio to be a stable equilibrium. Any perturbation that disturbs the equilibrium ratio will be unstable, and the action of natural selection will work to return the population to a one-to-one sex ratio. It is this story about the structure of the system and the dynamics that render the actual causal trajectory of any given population, in this case, irrelevant to the explanation (Woodward 2018, Strevens 2008).

Fifth and finally, although they may be used to explain token events (e.g., why a particular population has a one-to-one sex ratio at a given time or why a ball placed at a particular location on the lip comes to rest at the bottom of the basin), equilibrium explanations, some have argued, are distinguished from paradigmatic causal explanations by their ability to explain patterns of behavior shared by different systems (Kuorikoski 2007). Fisher's explanation, for instance, explains why any population at any time will either be at or tending toward the equilibrium sex

ratio. The explanation picks out the abstract structure and dynamics of the system, thereby identifying the minimum relevant features necessary for the system to evolve to its equilibrium state. Hence, an equilibrium explanation will show why many seemingly different systems all evolve to the same state.

5.3.2 Ensemble Explanations are Not Equilibrium Explanations

Each of the five key features of equilibrium explanation just discussed can be thought of as distinguishing equilibrium explanation from paradigmatic cases of causal explanation. Equilibrium explanations are (1) deep explanations (2) that do not trace low level causal relationships. Moreover, they involve systems that (3) exhibit global rather than local equilibria, which (4) makes the precise initial conditions of the system irrelevant, and (5) they consequently better explain patterns rather than token events. Sober (1983) argues that paradigmatic cases of equilibrium and causal explanation are on opposite ends of a continuum, with equilibrium explanations losing their distinctive characteristics as the equilibrium becomes more and more local. As the dominant type of causal explanation in molecular biology, mechanistic explanations are on one side of this continuum. I compare these explanations to ensemble explanations of allostery to determine whether we ought to consider the latter to be on the opposite side of the continuum, as a species of equilibrium explanation. Although ensemble explanations are deeper than mechanistic explanations, they nevertheless resemble paradigmatic cases of causal explanation along most of the other dimensions. They are therefore sufficiently distinct from equilibrium explanations to merit their own account of explanation.

Let us first consider the depth of ensemble explanations. Strevens' description of explanatory depth includes two features that it is useful to distinguish. He claims explanations are

deep because they exhibit "an unusual degree of generality" and also because they are "so abstract" (2008, 137). However, as Darden (1996) has argued, in biological explanations, abstraction and generality need not overlap. With respect to mechanistic explanation, abstraction refers to the amount of detail included in the description of the mechanism, whereas generality describes the scope, or the domain over which the mechanism holds (Machamer et al. 2000). A mechanistic explanation could be quite abstract, blackboxing many causal details, but also be quite limited in scope or generality, applying to only a small domain. This distinction between abstraction and generality can illuminate the relationship between mechanistic and ensemble explanations vis-à-vis explanatory depth. On my view, both types of explanations can be pitched at similar degrees of abstraction, while ensemble explanations are more general explanations that apply to a larger domain of target phenomena.

If we consider explanatory depth in terms of abstraction, then ensemble explanations are not necessarily any deeper than mechanistic explanations of allostery. In other words, these explanations do not strip away more causal or structural details about allosteric proteins and their ligands than mechanistic explanations. Both explanations can either include or abstract away these details. Early mechanistic explanations of allostery cited a structural or conformational change in an individual protein molecule to explain the observed effect: binding of the allosteric ligand at one site caused a conformational change that affected binding of another ligand at a second site on the protein. In the 1960s, when this mechanism was first proposed, there was little knowledge about the 3D structure of proteins, so the relevant features of allosteric proteins were described abstractly, i.e., functionally (Monod et al. 1965, Koshland et al. 1966). These proteins were represented as capable of adopting two different structures, but the structures were not described structurally (and certainly not in atomic-level detail); rather, they were characterized functionally

by their affinity for particular ligands. The R (relaxed) state could bind the allosteric ligand with high affinity, while the T (tensed) state was unable to bind the ligand. Even today, this abstract explanation—what we might call the mechanism schema (Machamer et al. 2000)—is still taken to explain allostery in certain protein-ligand systems.[136]

Ensemble explanations of allostery can be similarly abstract. They represent a population of protein molecules as an ensemble of microstate structures, each with different functional properties (e.g., with different thermodynamic stabilities and binding affinities for various ligands). The addition of the allosteric ligand perturbs the system and leads to a new equilibrium distribution of microstates. The microstate structures in ensemble explanations can be characterized at a similar degree of abstraction as the R and T state structures in mechanistic explanations of allostery. Thus, it is not the case that ensemble explanations are deeper explanations, if we measure depth in terms of degrees of abstraction.[137] In fact, if anything, ensemble explanations include more causal detail, since they require both structural *and* dynamic information about protein microstates.[138]

However, when we consider explanatory depth in terms of generality, we see that ensemble explanations of allostery are more general because they have a wider scope than mechanistic explanations. The MWC model, which was the first mechanistic explanation of allostery, was

---

[136] Of course, the functional properties that make a binding site high affinity for a ligand in one conformation and low affinity for that same ligand in the other conformation are instantiated in particular structures. But one need not cite these structures in order to give an explanation of allostery according to the MWC or KNF models.

[137] Explanatory depth as a measure of abstraction does not seem to capture what was special about equilibrium explanations, since the paradigmatic type of causal explanation in structural biology, viz., mechanistic explanations, can have depth of this sort by blackboxing certain features by describing them functionally. Moreover, Strevens' notion of explanatory depth that tracks abstraction is not included in other accounts of explanatory depth, which focus instead on the generality of the explanation (e.g., Hitchcock and Woodward 2003).

[138] The representations of proteins and microstates that figure in mechanistic and ensemble explanations can include more or less information about protein structure, but any additional structural features beyond those already included in the explanation would be relevant only to a particular case and would not be part of a generic mechanistic or ensemble explanation of allostery.

proposed as a general mechanism schema that could explain the behavior of all allosteric proteins (Monod et al. 1965). However, it soon became apparent that this model could not account for the behavior of certain allosteric systems. For example, although it could account for cases of positive cooperativity, in which binding of the allosteric ligand increases the binding affinity for the second ligand, it could not explain negative cooperativity, in which the allosteric ligand decreases the affinity of the protein for the second ligand. Koshland and colleagues (1966) noted this shortcoming and developed a new mechanistic model to account for both positive and negative cooperativity. Much later, a whole new class of allosteric proteins were discovered that could not be explained with mechanistic models that cite a conformational change between two structures. These proteins were intrinsically disordered (or included intrinsically disordered regions) and consequently had no stable 3D structure, yet they still exhibited allosteric behavior (Wright and Dyson 1999, Hilser and Thompson 2007, Tompa 2011). Thus, we now know that mechanistic explanations of allostery that focus on structural changes have a limited scope.

Ensemble explanations, in contrast, are more general in scope. An ensemble explanation can be given for allosteric proteins that have primarily two structures (and therefore approximate the earlier mechanistic models of allostery) as well as those proteins that lack well-defined structure (and hence cannot be explained mechanistically). Proponents of ensemble explanations argue that these explanations, unlike mechanistic explanations, provide a "unifying framework that can be applied to all allosteric systems" (Hilser et al. 2012, 588). These explanations encode structural properties of proteins in the microstate representations, but they identify changes in the thermodynamic properties of the ensemble rather than any particular structural changes in the microstates to explain allostery. These thermodynamic properties, unlike any specific features of protein structure, apply to all protein ensembles. Thus, compared to the abstract mechanism

schemata of the MWC and KNF models, ensemble explanations have a more general scope. Whereas the MWC and KNF models are only applicable to a subset of cases of allostery, an ensemble explanation can be given for any case of allostery. Thus, in this sense, ensemble explanations are in fact deeper than mechanistic explanations of allostery.

The second unique feature of equilibrium explanations, according to Sober (1983), is the fact that they explain by citing a set of possible causes without specifying the actual cause. In fact, this is why he claims equilibrium explanations are not causal, namely, because "causation abhors an ineliminable disjunction" (206). An equilibrium explanation, such as Fisher's account of the one-to-one sex ratio, provides information about causes, but it does not specify which low-level causal trajectory obtained.[139] Neither mechanistic nor ensemble explanations of allostery match this description. Although they pick out different causes, both mechanistic and ensemble explanations of allostery purport to cite the actual causes of the phenomenon rather than a disjunction of possible causes. [140] Mechanistic explanations cite changes in protein structure as the cause of allostery whereas ensemble explanations cite changes in the equilibrium distribution of microstates. Providing an ensemble explanation in place of a mechanistic explanation of an allosteric system cannot be interpreted as explaining via higher-level causal structure rather than lower-level causes. The transition from mechanistic to ensemble explanations does not lead to the omission of causal detail nor does it replace an actual causal trajectory with a larger set of possible trajectories. It is therefore quite different from the relationship between an explanation that cites a

[139] In contemporary discussion, philosophers have largely dropped this positivist framing of equilibrium explanations in terms of possible versus actual causes and have instead focused more on how little information (e.g., about initial conditions) is required to explain the evolution of a system towards its equilibrium. I will discuss this contemporary framing below.

[140] Sec. 4 offers a more robust defense of the claim introduced here, namely, that ensemble explanations of protein function are causal.

specific lower-level causal trajectory, for example, to explain the sex ratio of a given population, and an equilibrium explanation of the same phenomenon. Ensemble explanations do not achieve their generality by omitting causes contained in mechanistic explanations. Instead, they simply pick out different causes as explanatorily relevant.

Equilibrium and ensemble explanation differ in a third respect: the relevant equilibria in equilibrium explanations are global whereas they are local in ensemble explanations. Moreover, the equilibria in ensemble explanations are not the target of explanation. That is, the explanation does not purport to explain why a set of protein microstates is at equilibrium unless perturbed. To elucidate the different roles of equilibria in these two types of explanations, let us consider again Fisher's one-to-one sex ratio. This ratio is a global equilibrium for the population. You can perturb the system however you like, for instance, by increasing the ratio of daughters to sons or vice versa, yet it will eventually return to the original one-to-one ratio.[141] Because this sex ratio is a stable global equilibrium, the explanation for any token case can leave out almost all the causal details of the specific population. We do not need to know anything about the initial conditions of the population or the perturbation itself (so long as the ratio never leads to fixation and various other conditions are met) in order to explain why a population has a one-to-one sex ratio. The equilibrium in Strevens' basin example is similar. So long as certain features of the system are fixed (e.g., gravity, friction, and the basin itself), we can put the ball at any place on the inside lip of the basin and rest assured that it will land in the bottom of the basin. The explanation in both cases functions by demonstrating that the structural and dynamic features of the system are such that it has a global equilibrium that will be stable so long as certain features of the system remain

[141] Of course, this is not quite true. You cannot perturb certain features of the system that are responsible for establishing the global equilibrium. For example, you cannot change parental investment nor can you eliminate natural selection.

fixed. Once this has been established, we can then cite the existence of the equilibrium to explain why a given system, with initial conditions within a certain range, is either at or tending toward equilibrium.

Ensemble explanations, in contrast, invoke only local equilibria, and the fact that the microstates adopt an equilibrium distribution is simply assumed. The equilibrium distribution is an idealization, derived from the laws of chemistry and physics, which is taken to be approximately true of all proteins in solution. Unlike equilibrium explanations, ensemble explanations do not aim to show how or why this equilibrium emerges from the structure and dynamics of the system—it is simply taken as given. There is a further difference related to the dynamics of the allosteric system. In the case of an equilibrium explanation like Fisher's, the internal dynamics of the system permit it to temporarily depart from equilibrium (e.g., because of drift, random mutation, migration, etc.) but also include a countervailing force (e.g., natural selection) that will bring the population back to equilibrium. This dynamic setup explains why any population will return to the equilibrium state. In ensemble explanations, however, the perturbation to the system causes the system to move to an entirely new equilibrium distribution. It is precisely this *change* in the equilibrium distribution upon perturbation that is cited to explain allostery. Since the dynamics that cause a population of molecules in solution to adopt an equilibrium distribution of microstates tell us nothing about the evolution of system when it is perturbed, these local equilibria are unable to explain the observed allosteric effect. Thus, compared to the equilibria in equilibrium explanations, those in ensemble explanations have a different and less central explanatory role.

The role of initial conditions is a fourth feature that distinguishes equilibrium explanations from ensemble explanations. Most philosophical accounts of equilibrium explanation conclude that part of what makes these explanations unique is the relative unimportance of the initial

conditions of the system. For instance, Sober (1983) claims that an equilibrium explanation can be given even when we have "considerable ignorance of the actual forces and initial conditions that in fact caused the system to be in its equilibrium state" (209). Although he thinks most equilibrium explanations can be captured as causal by his interventionist framework, Woodward (2003) largely agrees. Equilibrium explanations answer a broad range of what-if-things-had-been-different, or w-questions (Woodward 2003). However, rather than specifying which changes to factors in the explanans would have resulted in changes in the explanandum phenomenon, equilibrium explanations focus "on the fact that if various factors had been different in various ways, the explanandum would *not* have been different" (Woodward 2018, 131). Ensemble explanations of allostery do not have this feature. To give an ensemble explanation, one must have information about the preexisting equilibrium distribution of microstates. The initial equilibrium of microstates in a population of protein molecules determines the allosteric behavior of the system: "The observed [allosteric] response is dependent on where the equilibrium is poised before activation" (Motlagh et al. 2014, 333). In other words, if a certain subset of microstates is thermodynamically stable and therefore dominant in an ensemble, the addition of an allosteric ligand might lead to one outcome (e.g., positive allostery), whereas if a different subset of microstates is stable in the preexisting equilibrium, then the same allosteric ligand might lead to another outcome (e.g., negative allostery or no allosteric coupling at all). Because the preexisting equilibrium, or how the ensemble is "poised," can make such a big difference, we cannot be ignorant of initial conditions if we wish to give an ensemble explanation of allostery.

Fifth and finally, unlike equilibrium explanations, ensemble explanations are not chiefly focused on explaining patterns of behavior. One of the unique features of equilibrium explanations is that they show how many of the differences between systems are irrelevant to their behavior.

They afford "deep" explanations that rely only on structural features of the system. For this reason, some philosophers have argued that what is special about equilibrium explanations is not their ability to explain token events but rather their ability to explain why seemingly disparate systems exhibit the same behavior (Kuorikoski 2007, Rice 2015). For example, Kuorikoski (2007) wonders "if the intended target of Fisher's argument is really a 1:1 sex ratio of a single population at a given time?" He concludes it is not. Instead, he claims, "the perplexing explanandum is undoubtedly the pervasiveness of the 1:1 sex ratio across different populations, species, and environments" (Kuorikoski 2007, 152). This "perplexing explanandum" is an example of what Batterman (2002) calls a type ii question: it asks why a pattern of behavior manifests over a range of different systems in different circumstances. Since they highlight structural features that show why certain differences between systems do not matter, equilibrium explanations are well-suited to answer type ii questions.

In contrast, ensemble explanations of allostery provide equally good answers to both type i and type ii questions. On the one hand, they can identify general features of allosteric systems and the energetic connections between different binding sites on a protein. Consider the case of agonism-antagonism switching, in which the addition of an allosteric ligand sometimes leads to positive cooperativity and other times leads to negative cooperativity. Using the mathematical description of the ensemble for a three domain protein, researchers have identified the necessary features an allosteric protein must have in order to exhibit this type of conditional cooperativity or switching behavior. They have shown how the binding domains on a protein must be energetically linked in order to yield this observed behavior:

> The importance of the result is that it demonstrates how a single thermodynamic architecture, within the framework of the most simple three-domain model, can poise the energy landscape to respond in a "functionally pluripotent" manner. (Motlagh and Hilser 2012, 4136)

From their analysis, the scientists were able to identify the "thermodynamic 'ground rules' for conferring the ability to switch responses," thereby explaining why all proteins capable of such functional behavior share a single thermodynamic architecture (Ibid., 4137). Ensemble explanations can thus reveal general features of allosteric systems, showing how the energetic connections between binding sites will yield similar allosteric behavior in different protein-ligand systems.

On the other hand, ensemble explanations are often used to answer type i questions about why a system-level behavior manifests in a particular system under certain conditions at a given time (Batterman 2002). They show why allostery occurs in certain proteins with certain allosteric ligands, and the details of the case matters. To explain the allosteric behavior in a given case, such as between hemoglobin and oxygen, we must include details about the preexisting equilibrium of protein microstates, which microstates are dominant, and the nature of the allosteric ligand. Ensemble explanations of protein function aim to explain particular cases, many of which are biomedically relevant. For example, an entire classes of HIV antiretroviral drugs are allosteric inhibitors of HIV enzymes (Mehellou and De Clerq 2010). Researchers interested in explaining the behavior of these drugs and designing new allosteric inhibitors of viral proteins are not primarily interested in type ii questions. Instead, they want to understand the behavior of these ligands in specific contexts and the way in which they inhibit the action of viral enzymes. Although scientists develop ensemble explanations to answer type ii questions, these explanations are also used in biomedical contexts in which scientists want to explain what happens in a particular protein-ligand system.[142]

[142] Since some philosophers have claimed that equilibrium explanations answer type i questions while others have claimed they can answer type ii questions, it is worth further elaborating on the distinction I see between equilibrium

From this analysis of the five dimensions that distinguish equilibrium explanation from paradigmatic causal explanation, we see that ensemble explanations are significantly different from equilibrium explanations. Ensemble explanations are deeper than mechanistic explanations of allostery. But along the other dimensions, ensemble explanations have more in common with these paradigmatic causal explanations than they do with equilibrium explanations. We can therefore conclude that ensemble explanations are not a species of equilibrium explanation, as characterized in the philosophical literature. In the next section, I turn to consider recent philosophical work that attempts to show how the interventionist framework can apply to explanations in thermodynamics. This analysis, I argue, can help shed light on how ensemble explanations work.

5.4 Interventionist Causation and Explanation in Equilibrium Thermodynamics

One branch of thermodynamics is exclusively concerned with thermodynamic systems at equilibrium. Equilibrium thermodynamics characterizes such systems and analyzes the processes that lead them to transition between equilibrium states. Because ensemble explanations of allostery involve thermodynamic descriptions of protein ensembles and focus on the transition between

---

explanations, like Fisher's explanation of the one-to-one sex ratio, and ensemble explanations of allostery. Although Sober (1983) presented Fisher's equilibrium explanation as explaining the sex ratio of a given population at a given time, philosophers have more recently suggested that this is the wrong explanandum. For instance, Kuorikoski (2007) argues that equilibrium explanations account for patterns that occur across systems. He suggests that the existence of the same patterns of behavior across wildly different systems is what is curious and in need of explanation. Furthermore, it is the explanation of patterns—by showing why certain details do not matter—that justifies philosophers' focus on equilibrium explanations. In contrast, I suggest that it is not the case that the type ii questions that ensemble explanations can answer are necessarily more interesting or more in need of explanation. Using ensemble explanations to answer type i questions in biomedical contexts is an important feature and the primary justification for the development of this type of explanation in the first place (Hilser et al. 2012, Motlagh et al. 2014).

equilibrium states, the philosophical literature on thermodynamics offers a useful framework for the analysis and interpretation of ensemble explanations. Although philosophers of science interested in equilibrium explanations often discuss the ideal gas law as an example, Zwier's (2017) recent account of causation in thermodynamics provides the most systematic treatment of equilibrium thermodynamics. She argues that the interventionist account of causation can be used to describe thermodynamic theorizing. In what follows, I first present Zwier's account of causation in thermodynamics (Sec. 5.4.1). I then show how it can be applied to cover ensemble explanations of allostery (Sec. 5.4.2). From this analysis, I argue that ensemble explanations are causal (Sec. 5.4.3). Finally, I argue that ensemble explanations are novel explanations, distinct from both the common explanations in thermodynamics as well as the mechanistic explanations that dominate structural and molecular biology (Sec. 5.4.4).

5.4.1 Intervention on Equilibrium States

Thermodynamic analysis focuses on the equilibrium states of a system and the processes that move the system from one equilibrium state to another. Zwier (2017) argues that thermodynamic theorizing of this sort can be understood within the interventionist framework of causation. To make sense of her argument, we must first clarify two thermodynamic concepts: equilibrium and system. A system is in thermodynamic equilibrium when "all natural processes of change have come to an end and the observable state of the system is constant in time" (Frigg 2008, 99). It describes the state of the system in which there is no change in macroscopic variables. As Zwier (2017) points out, however, equilibrium states do not occur spontaneously, since "natural thermodynamic systems are in constant flux" (1304). Any candidate system is likely to be undergoing internal processes of change as well as exchanging energy with its environment. In

fact, thermodynamic systems themselves are largely theoretical constructs that we impose upon the world in order to do thermodynamic analysis and bookkeeping. For a system to be in thermodynamic equilibrium, it must be given sufficient time without any external perturbations from the environment or other systems. The only way for this condition to be met, according to Zwier, is by carefully setting and maintaining the boundary conditions that isolate the system from its surroundings.

Consider a plant cell as a candidate natural thermodynamic system, using the cell wall as the boundary of the system. For a healthy plant in a greenhouse, for example, the pressure and volume of the cell will be relatively fixed because the cell wall is rigid, and a well-watered plant will be able to maintain a relatively constant turgor pressure. In addition, even though most plants are exothermic, if the greenhouse is temperature-controlled, the temperature of the cell will also remain relatively constant. Finally, if we only consider the plant during the day, when photosynthesis is occurring, we might assume that the chemical reactions of photosynthesis and metabolism are occurring at a relatively steady state. Given these assumptions, we might reasonably think of the plant cell as thermodynamic system at a quasi-equilibrium state. It is clear from this description that the quasi-equilibrium state requires active maintenance of the boundary conditions. Stricter equilibrium states will require even more isolation of the system and control of the boundaries, and true equilibrium states "exist only in theory," since they "require idealized boundaries," such as perfect insulators and perfectly rigid containers (Zwier 2017, 1305). To even conceive of a system at thermodynamic equilibrium, we must first consider the constraints on the system imposed by the boundary conditions.

Zwier draws upon Woodward's (2003) interventionist account of causation to analyze the relationship between a system's boundary conditions and its equilibrium states. According to

Woodward, causal claims must correctly capture the counterfactual dependence relations between cause variables and effect variables. These patterns of counterfactual dependence between the variables are understood in terms of interventions. An intervention sets a cause variable to a certain value. According to Woodward, C is a cause of E if there exists a possible intervention on C that would lead to a change in the value of E. The intervention must be ideal, in that the change in E must be brought about only through the change in C and not by changes to other causes elsewhere. Woodward uses this account of causation as the cornerstone of his account of causal explanation. On his view, causal explanations identify patterns of counterfactual dependence in order to answer a range of what-if-things-had-been-different questions. That is, they explain by showing how certain changes in the explanans (i.e., changes in the value of C) would have yielded changes in the explanandum phenomenon (i.e., changes in the value of E).

According to Zwier (2017), this interventionist framework captures thermodynamic reasoning in two ways. She first argues that thermodynamic equilibrium is an inherently manipulated state, with the boundary conditions acting on a system as interventionist causes. These manipulated equilibrium states are central to thermodynamic theorizing, since most thermodynamic explanations describe the movement of a system from one equilibrium state to another. Second, according to Zwier, the "driving force" that pushes a thermodynamic system from one equilibrium to another is an external intervention on the system's boundary conditions. Because interventionist causes play these two important roles in thermodynamics, Zwier concludes that "interventionist reasoning is inseparable from the structural foundation of thermodynamic theory" (Ibid., 1307).

Let us look more closely at these two ways in which boundary conditions can function as interventions on thermodynamic systems. First, they intervene on a thermodynamic system in

order to enable it to reach equilibrium. To even conceive of a system at thermodynamic equilibrium, we must consider the constraints imposed by the boundaries on the system: it is impossible to have a system at equilibrium "unless the boundaries of the system are well defined" (Ibid., 1305). Boundary conditions constrain the system by setting the value of certain thermodynamic variables. For instance, the temperature-controlled greenhouse in the example above serves to set the temperature of the thermodynamic system—i.e, the plant cell—to the ambient temperature, which is set to a particular value by the thermostat. Even a natural thermodynamic system, such as a plant cell, is at quasi-equilibrium only because it is constantly being intervened upon to maintain its boundary conditions. It is for this reason that Zwier considers thermodynamic equilibrium to be an "inherently manipulated state" (1306). The equilibrium states that figure in most thermodynamic theorizing are similarly manipulated. Since the boundary conditions always intervene upon the system in order to fix certain thermodynamic variables, Zwier argues they "constitute external interventions on the system" that are "entirely consistent with the concept of an intervention that has been developed by Woodward" (Zwier 2017, 1305-6). On her view, because the boundary conditions distinguish the system from its surroundings, they are not part of the system itself and therefore can be construed as external interventions.

As external interventions, boundary conditions themselves can be the cause and can therefore explain why certain thermodynamic variables take on particular values. On this view, even for a thermodynamic system at equilibrium, such as a gas in a box, we might say that the temperature of the gas is the cause of its pressure, if the box has rigid walls and the temperature is maintained by a temperature reservoir. Given this setup, the fact that the experimenter set the water bath to a certain temperature rather than another explains why the system is at a particular pressure rather than another. Even though no change in temperature actually occurs, because we can

interpret the temperature as changeable or variable, it reveals a counterfactual dependence relation between temperature and pressure and therefore tells us something about the conditions under which the pressure would have been different.[143] In sum, boundary conditions can be construed as intervening on a system, setting thermodynamic (cause) variables to particular values which can then be cited to explain the values of other (effect) variables at equilibrium.

The second way boundary conditions function as interventionist causes fits more naturally within the interventionist framework. In this case, those boundary conditions that define the system at equilibrium and distinguish it from its surroundings are relegated to the status of background conditions. The conditions that are directly manipulated are construed as interventionist causes; they are the ones that lead to a change in the equilibrium state of the system. Consider this example from Frigg (2008). An ideal gas, initially confined by a barrier to one half of a box, with perfectly rigid and insulated walls, is at equilibrium. The barrier is then removed, and the gas molecules disperse throughout the new volume until the system reaches a final equilibrium state. This final state differs from the initial state in that the volume V will have doubled upon removal of the barrier and the pressure P will have halved. We know this from the application of the ideal gas law, PV = nRT, to the system. Because of the thermally insulated walls, the temperature T is constant.[144] Therefore, this manipulation of the system, which doubles the volume, also causes the pressure to halve.

---

[143] See Woodward (2003), p. 234 for a similar example. The details of the setup matter, according to Woodward. If a particular thermodynamic parameter is construed as a constant rather than a variable, then even though it can be derived from the ideal gas law and the values of the other thermodynamic parameters, it is not explained by them. For instance, given the above experimental setup, the pressure does not explain the volume of the system, even though the volume can be derived using the ideal gas law and the pressure and temperature. In this case, the pressure does not explain the volume because the box in this setup is rigid and will not vary given changes in pressure.

[144] Because of the insulating walls, the energy remains the same and hence also the temperature, since for an ideal gas, internal energy is proportional to temperature. In this case, the moles of gas (n) also remain constant.

Unlike the previous case, this second case does not focus on the role of boundary conditions in enabling and maintaining the equilibrium state, but instead focuses on a particular manipulation of the boundary conditions that drives the system from one equilibrium to a new equilibrium state. It is thus quite natural to think of this manipulation of the boundary conditions—i.e., the removal of the barrier—as an ideal intervention on the system. It changes the volume of the system by setting this variable to a new value, which in turn drives the system to a new equilibrium with a new value for the pressure variable. The boundary condition that was manipulated in this second case counts as an interventionist cause, and the new volume of the gas, along with the ideal gas law and the fixed temperature, can be cited to explain the new pressure of the system. This second case considers boundary conditions to be causes only insofar as they are manipulated to drive the system from one equilibrium state to another.

Zwier (2017) considers in further detail this second way changes in the boundary conditions can act as interventions in thermodynamic analysis. The First and Second Laws of thermodynamics which require the conservation of energy and the maximization of entropy, respectively, coupled with the boundary conditions, determine the equilibrium state of any given thermodynamic system. Thermodynamic reasoning begins with equations of state, such as the ideal gas law, and thermodynamic potential functions that relate small changes in state variables to changes in internal energy (U) and entropy (S). If we consider what happens to two isolated thermodynamic systems A and B at different temperatures ($T_A \neq T_B$), which are thermally isolated from their surroundings, we can derive an equation that relates the total entropy of the combined systems to their internal energies, volumes, and particle numbers. Since we assume that the particle numbers and volumes are fixed by the boundary conditions, we then obtain the following equation:

$$dS_{total} = \left(\frac{1}{T_A} - \frac{1}{T_B}\right) dU_A \ .$$

By definition, at equilibrium, there will be no change in entropy in the system ($dS_{total} = 0$), and this will occur when the temperatures of the two bodies are equal ($T_A = T_B$).

Zwier argues that this thermodynamic analysis does more than derive the well-known result that two bodies at different temperatures will reach equilibrium when they are both at the same temperature. It also shows that the difference in temperatures between system A and B—i.e., the nonzero value of this term—"act[s] as a driving force," causing a change in the internal energy of system A (Zwier 2017, 1310). According to Zwier, this "'driving force' language," as well as the thermodynamic changes it describes, "matches the way in which thermodynamic variables would be modeled in the interventionist account of causation" (Ibid.). In the temperature equilibration case, the act of bringing system B into thermal contact with system A is the primary intervention. The fixed boundary conditions that set and maintain the other thermodynamic variables such as the volume and particle number are what Zwier calls "auxiliary interventions" (Ibid., 1311). These auxiliary interventions are set up in just such a way to render the primary intervention an ideal intervention. Since all other variables are held fixed by these interventions, we can see how the primary intervention affected the internal energy of the system. According to Zwier, the interventionist framework demonstrates that "T_A is an interventionist cause of U_A" (Ibid.).

Most thermodynamic theorizing fits this model: some boundary condition is modified or some constraint on a system is lifted, and the system then tends toward a new equilibrium that maximizes the entropy of the system. It is the intervention on the system at equilibrium—in the above case, the removal of the thermal barrier between systems—that leads the system to move toward a new equilibrium state. That intervention thus reveals one thermodynamic parameter (T) to be the cause of another (U). Within the interventionist framework, we can therefore cite the

manipulation of the boundary conditions to explain why a system will have a new value for other parameters when it reaches its new thermodynamic equilibrium.

5.4.2 Interventionist Interpretation of Ensemble Explanations

In order to apply the insights from thermodynamics to the case of ensemble explanation, we first need to clarify the meaning of certain concepts—viz., system and equilibrium—in this context. The relevant thermodynamic system, in this case, is the set of protein molecules. The representation of the system used in ensemble explanations does not explicitly include other features of the local environment. It does not represent the solvent or other chemical or biological factors. However, these contextual environmental factors are implicitly included in the representation of the system, since these factors will affect the thermodynamic stability of the microstates and will thus be reflected in the ΔG_j for each microstate. Like the physical thermodynamic systems Zwier (2017) discusses, these biological systems also require the maintenance of boundary conditions. For example, *in vitro*, the temperature of the system will remain approximately constant because the ambient temperature of the lab is kept constant, and *in vivo*, metabolic activity will maintain the cellular temperature of endothermic organisms. The maintenance of the boundary conditions is a necessary precondition for a system to reach equilibrium. However, in ensemble explanations of protein behavior, these boundary conditions are not typically manipulated and thus function as implicit background conditions.

The relevant equilibrium of the system is the equilibrium distribution of protein molecules across microstate structures. Because of theoretical and empirical evidence, we know that individual protein molecules in solution are in constant motion, wandering around conformation space. At any moment in time, a population of the same type of protein molecules in solution will

be structurally heterogeneous. Although individual molecules do not have stable 3D conformations, the ergodic hypothesis tells us that the temporal evolution of all the individual molecules is replicated in the instantaneous ensemble distribution of microstates (Hilser et al. 2006). When the system is at equilibrium, a certain fraction of individual molecules will occupy each microstate. More stable, lower energy microstates will be more populated than less stable, higher energy microstates. (Recall that the thermodynamic stability of each microstate (ΔG_j), along with the stabilities of the other microstates in the ensemble, determines the probability for the microstate (P_j).) At thermodynamic equilibrium, individual molecules will continually move through different conformations, but there will be no net change in the fraction of molecules in each microstate at any time. Since the distribution of molecules across the ensemble of microstates is unchanging, this is called the equilibrium distribution.

Since the ensemble representation is imported from statistical mechanics, it is unsurprising that this representation of proteins is similar to that used to describe simpler thermodynamic systems, such as an ideal gas in a box. In both cases, at equilibrium the molecular arrangement—i.e., which molecules are in which microstates—is constantly changing, but the overall distribution of molecules across microstates remains constant (Frigg 2008). The equilibrium distribution of microstates is fixed by the set of thermodynamic parameters, or boundary conditions, of the system. In other words, the probability associated with each microstate at equilibrium will be fixed for any particular set of thermodynamic parameters (e.g., temperature, volume, pressure, internal energy, etc.). Changes in these parameters or other external interventions will cause the system to re-equilibrate at a new equilibrium state. For example, in the protein case, changes to the temperature of the system or other parameters that affect the relative stability of microstates (e.g., polarity or pH of the solvent) will lead the ensemble of protein molecules to adopt a different

equilibrium distribution, causing certain microstates to become more stable (and therefore more frequent) and others to become less stable (and less frequent).

Having clarified the key concepts, we can now turn to consider the role equilibrium states play in ensemble explanations. Compared to their role in canonical equilibrium explanations, the equilibrium states in ensemble explanations are less informative and less explanatory. What is cited to explain protein behavior is not the equilibrium distribution itself, but rather the change in equilibrium distribution upon perturbation of the system. As I will argue in Section 4.3, it is this focus on the change in equilibrium distribution and the cause of the perturbation that makes ensemble explanations of allostery causal.

To highlight this difference in the explanatory role of equilibrium states, let us have another look at Sober's (1983) account of equilibrium explanation. Reasoning from Fisher's explanation for the one-to-one sex ratio, he claims that equilibrium explanations explain how and why a population exhibits the equilibrium ratio. The explanation works by revealing structural features of the system as well as the internal dynamics that describe its evolution. In Fisher's case, the system is conceived as having a stable equilibrium state, but the internal dynamics of the system are such that a population could temporarily move to a non-equilibrium state. The population could be perturbed from equilibrium by any number of evolutionary forces (e.g., drift, mutation, migration, etc.). But when the population moves away from equilibrium, it will be brought back by the force of natural selection acting against parental pairs that overproduce the majority sex. Forces that perturb the system away from equilibrium, as well as the countervailing force of natural selection, are thus built into the internal dynamics of the system. According to Sober, explanations of this sort "are made possible by theories that describe the dynamics of systems in certain ways" (207). Or, more specifically, on this account of equilibrium explanation "the occurrence of an

equilibrium state can be explained by exhibiting it as an equilibrium" (Kuorikoski 2007, 150).[145] Fisher's theoretical description of the system thus counts as an explanation, according to Sober, because it demonstrates that within a system governed by these particular dynamics, the one-to-one sex ratio is a stable equilibrium state.

The thermodynamic equilibrium in the protein or ideal gas case plays a very different role than the equilibrium cited in Fisher's explanation. Both the representation of the system at equilibrium differs, as well as the role of the equilibrium in explaining the behavior of the system. The thermodynamic equilibrium is construed as an ideal equilibrium, occurring in a perfectly isolated system, such that once it has reached an equilibrium distribution nothing internal to the system will shift it from that distribution. Once the system has reached its equilibrium state there are no internal dynamics that can move it away from that state. Although changes occur all the time at the micro-level, by definition, no change occurs at the macro-level. (Otherwise, the system would not be at thermodynamic equilibrium.) Therefore, any perturbations to the system caused by changes in thermodynamic parameters or boundary conditions that lead the system away from equilibrium are represented as external to the system. They are not included in the internal dynamics of the system.

Defined in this way, the equilibrium in an ensemble explanation, by itself, is not that informative.[146] By using the ensemble representation, scientists aim to represent the dynamic properties of protein molecules so that these dynamics can be taken into account in explanations of protein behavior. But the mere fact that individual protein molecules will adopt an equilibrium

[145] In this quotation, Kuorikoski summarizes Sober's position. Kuorikoski (2007) ultimately argues that this view is mistaken, since he thinks equilibrium explanations primarily explain patterns rather than token events.
[146] For a canonical equilibrium explanation, showing that a particular state is an equilibrium state can be cited to explain why a given system is at that state. (I will discuss this further in the following section.)

distribution does not explain allostery or any protein behavior. Moreover, the goal of an ensemble explanation is not to show that or explain why a protein has a particular equilibrium distribution. Instead, the fact that protein molecules (or any molecules, for that matter) will adopt an equilibrium distribution is an assumption, supported by the laws of thermodynamics, used to justify the ensemble representation of protein molecules.

The thermodynamic equilibrium is therefore not a property of the structure and dynamics of the system that can be used, on its own, to explain the evolution of the system. It is thus unlike canonical equilibrium explanations that reveal the structure and dynamics of a system to show why a particular equilibrium state is to be expected. Consider Potochnik's (2015) example of an everyday equilibrium explanation to explain the temperature of her coffee. To explain the fact that her once hot coffee is now at room temperature, she cites the fact that it is in an open ceramic mug and has been sitting her desk for the past four hours. These two conditions show that whatever the initial temperature of the coffee, in four hours, it will have thermally equilibrated with the ambient temperature of her office. On her view, the explanation "must include the assumptions needed to generate the domain of attraction, that is, the range of conditions that would lead to the equilibrium value" (1179). This counts as an explanation because it shows that the system has an attractor and that the initial conditions in the given case fall within the range of that attractor. This everyday example differs from ensemble explanations regarding its use of the equilibrium. In Potochnik's case, the explanandum is the temperature of her stale coffee, and she can therefore explain its temperature by showing that it is at the equilibrium state, as expected. The goal of an ensemble explanation, in contrast, is never to explain why a given system is at equilibrium, since by definition the system is at equilibrium unless perturbed. A single equilibrium distribution, by itself, cannot be used to explain the behavior of proteins. Although they cite equilibria, ensemble

explanations do not explain the behavior of a system by identifying the equilibrium state or explaining why it emerges from the dynamics of the system. Nor does an ensemble explanation explain why a particular system is in a particular state by showing that state to be an equilibrium distribution.

Instead, the relevant feature of an ensemble explanation is the *change in equilibrium* upon perturbation of the system. That protein molecules adopt an equilibrium distribution of microstates is taken as given. This is not to say that there is no explanation for why a population of protein molecules in solution will obtain an equilibrium distribution across a set of structurally distinct microstates. To the contrary, we have good theoretical and empirical reasons to believe that protein molecules in solution are dynamic, undergoing significant structural fluctuations (Cooper 1976, 1980; Cui and Karplus 2008). However, an ensemble explanation of allostery, or other protein function, does not seek to explain these facts about protein equilibria. Merely showing that a particular distribution is an equilibrium distribution cannot explain protein function; a single equilibrium distribution, on its own, is not explanatory sufficient. In this case, the equilibrium is a red herring, since the equilibrium does not explain anything. Rather, it is the *change in equilibrium* that does the explanatory work in ensemble explanations.

5.4.3 Ensemble Explanations are Causal

This emphasis on the change in equilibrium and the cause of that change, supports my contention that ensemble explanations are more similar to paradigmatic cases of causal explanation than equilibrium explanation. Recall that an explanation for allostery involves two different why-questions: (1) why does a protein-ligand system exhibit changes in affinity upon addition of an allosteric ligand, and (2) why (and how) does the allosteric ligand lead to a change in the

equilibrium distribution of microstates. The ensemble explanation answers both of these by citing external causes that move the system from one equilibrium state to another. The explanatory account given by an ensemble explanation in response to the first question focuses on the change in equilibrium distribution of microstates caused by the perturbation to the system, in this case, the allosteric ligand. The observed increase in binding affinity for the second ligand is attributed to the shift in the equilibrium distribution caused by the perturbation. Compared to the previous equilibrium, the new one favors microstates that can bind the second ligand. The change in equilibrium distribution upon perturbation therefore explains the observed changes in ligand binding.

The second why-question could be construed as a follow-up question, a request for additional information. To answer the first question, one need only cite the change in equilibrium distribution of microstates caused by the perturbation and demonstrate that the new equilibrium favors certain microstates over others. But one might then ask, how does the addition of the allosteric ligand perturb the system? Here the answer follows from the ensemble representation of the system, along with the laws of chemistry and physics. The allosteric ligand causes the change in equilibrium distribution by binding to the subset of microstates that have the appropriate ligand binding site (Hilser et al. 2012, Motlagh et al. 2014). When protein molecules in a given microstate bind the allosteric ligand, it increases the stability of those microstates, lowering their free energy (ΔG_j). The addition of the allosteric ligand thus intervenes on the system by setting the free energy values for certain microstates lower than they were before. This increases their stability relative to those microstates the ligand was unable to bind, which in turn leads to the change in probabilities (P_j) of all the microstates in the ensemble. This change in probabilities constitutes a re-equilibration of the system, leading individual protein molecules to redistribute from higher energy

microstates into lower energy ones. The redistribution of molecules across the microstates caused by the addition of the allosteric ligand to the system is at the heart of the ensemble explanation of allostery: allostery occurs, on this account, if the new equilibrium favors microstates that can bind the second ligand. The allosteric ligand, which caused the change in equilibrium, explains this second why-question.

The answer to both why-questions cites the addition of the allosteric ligand as the cause of the change that is used to explain the allosteric effect. In the first case, the allosteric ligand is represented as intervening on the ensemble distribution, perturbing the existing equilibrium state and leading it to adopt a new equilibrium state. This intervention on the equilibrium distribution of microstates is then cited as the explanation for the change in protein behavior. In Figure 13, the causal structure of this explanation is represented by a directed graph.

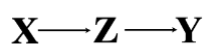

Figure 13 Causal Representation of Allostery.

When it is added to the system, the allosteric ligand (X) alters the equilibrium distribution (Z) in such a way as to result in allosteric behavior (Y). The precise details of how ligand binding leads to this change in equilibrium are blackboxed from this abstract causal explanation. Pitched at this level, the ensemble explanation is similar to abstract mechanistic explanations for allostery. Mechanistic explanations of allostery exhibit the same causal structure, but the intermediate causal variable differs. In mechanistic explanations, the allosteric ligand (X) acts upon the structure of the protein molecule (Z) in order to bring about a particular allosteric effect (Y). Abstract mechanistic and ensemble explanations share the same general causal structure, but cite changes to different intermediate causal variables as the proximate cause of the allosteric effect. Therefore,

if we think that mechanistic explanations of allostery are straightforwardly causal explanations, then ensemble explanations, which share the same structure, must also be causal explanations.

To answer the second why-question, we must flesh out this abstract causal explanation to reveal the intermediate causal variables that connect the allosteric ligand to the equilibrium distribution. In other words, we must show how the allosteric ligand ultimately leads to change in equilibrium distribution. For an ensemble explanation of allostery, this causal chain is something like the following: when added to the system, the allosteric ligands bind a subset of microstates, which in turn alters the thermodynamic stabilities of those microstates (ΔGj). The change in their stabilities leads to a change in their probabilities (Pj), which in turn leads to a change in the probabilities for all the microstates, and this change in the probabilities of all the microstates constitutes a change in the equilibrium distribution of the system. We can represent this chain of intermediate causes using another directed graph (Fig. 14).

$$\mathbf{X \rightarrow Z_1 \rightarrow Z_2 \rightarrow Z \rightarrow Y}$$

Figure 14 Causal Representation of Allostery with Intermediate Causal Variables.

Again, the allosteric ligand (X) causes the observed allosteric effect (Y), but additional intermediate causes are also specified (Z_1, Z_2, etc.) that ultimately lead to a particular equilibrium distribution (Z) that gives rise to the observed behavior of the system (Y). A similar elaboration of intermediate causal variables could be requested of the abstract mechanistic explanation given in response to the first why-question, which cited a change in structure caused by the allosteric ligand. However, unlike in the ensemble explanation, the intermediate causes would likely be unique to a particular protein-ligand system. They would have to show how the binding of the allosteric ligand at one site leads to local structural changes that ripple through the protein to ultimately affect the structure of the second binding site. This is precisely the level of causal detail that Perutz (1970)

aimed to provide to explain the allosteric behavior of hemoglobin oxygen-binding. His account showed how oxygen binding led to a structural change in hemoglobin, just as Figure 14 shows how the addition of the allosteric ligand to the system leads to a change in the equilibrium distribution of microstates. If we take Perutz's mechanistic explanation for allostery to have the similar causal structure to the more concrete ensemble explanation, with its additional intermediate variables, we once again must conclude that ensemble explanations of allostery are causal explanations.

This analysis of ensemble explanations shows that they can be interpreted as causal using the interventionist account. In that respect, ensemble explanations of allostery are similar to mechanistic explanations of allostery. Both recognize the allosteric ligand as the cause of the allosteric effect, but they differ in the intermediate causes they posit. While mechanistic explanations cite changes in protein structure, ensemble explanations cite thermodynamic changes in the equilibrium distribution of microstates. The latter are therefore similar to the causal explanations in thermodynamics that Zwier (2017) analyzes. She describes cases in which the boundary conditions of a thermodynamic system are manipulated and argues that changes in the boundary conditions are "driving forces" that move the system from one equilibrium state to another. In the case of allostery, the intervention on the system is more easily interpreted as external to the system. First, the addition of the allosteric ligand satisfies the requirements for an intervention, since it manipulates the ΔG_j values for the microstates to which it binds, setting their values lower than before. Second, the addition of the allosteric ligand is an intervention on the system that is unambiguously external. The allosteric ligand is not part of the boundary conditions that enable the maintenance of equilibrium nor can the addition of allosteric ligand be conceived as removing a constraint on the system. In the laboratory setting, an experimenter adds an aliquot

of the allosteric ligand to a cuvette or test tube containing a known concentration of the protein and measures the response. This is precisely the sort of surgical intervention required by the interventionist account. Thus, even those skeptical of the application of the interventionist approach to thermodynamics should therefore agree that the allosteric ligand is the external cause of the change in the equilibrium distribution of protein microstates.

5.4.4 Novelty of Ensemble Explanations in Biological Sciences

I have just argued that ensemble explanations are a type of causal explanation, similar to explanations in equilibrium thermodynamics. In this section, I briefly highlight the novelty of ensemble explanations. I first distinguish these explanations in structural and molecular biology from the cases of causal explanation in physics that are the focus of Zwier's analysis. I then show how they differ markedly from the mechanistic explanations that tend to dominate these biological fields. Finally, I suggest that one major advantage of ensemble explanations of allostery is their ability to unify phenomena and provide a unified explanatory strategy for understanding the behavior of proteins.

Even though they import theoretical and mathematical structure from statistical mechanics, ensemble explanations differ from similar causal explanations in physics. For one, ensemble explanations are quite different from thermodynamic explanations that invoke the ideal gas law. Causal explanations that cite the ideal gas law to explain changes in the thermodynamic parameters of a system are silent about the properties of the microscopic constituents of the system. They do not even cite the distribution of gas molecules across microstates but instead focus on their average thermodynamic properties. As we have seen, ensemble explanations do require consideration of the distribution of molecules across microstates. However, this difference from ideal gas law

explanations does not distinguish ensemble explanations from the majority of explanations in statistical thermodynamics. The ideal gas law, as it turns out, is a rather unique higher-level generalization. Most other explanations in statistical thermodynamics require partitioning the system into microstates and then tracking the change in distribution of molecules across microstates (Frigg 2008).

Although they also track changes in the distribution of molecules across microstates, ensemble explanations differ from these thermodynamic explanations as well. In the physics case, the change in the system is explained by showing how the intervention caused a change in the distribution of molecules across microstates. After a perturbation that increases the thermal energy of the system, for example, more molecules will be occupying higher energy microstates that were rarely populated before. An explanation of the system's thermodynamic parameters after the perturbation will cite this change in the distribution of molecules to account for the new equilibrium values. But the microstates themselves do not change. That is, although the intervention affects the distribution of molecules across microstates, it in no way alters the microstates or their properties.[147]

In the biological case, in contrast, the microstates are themselves altered by the perturbation. Consider the ensemble explanation for allostery. The system is partitioned into a set of structural microstates, which each have three explanatorily relevant properties—thermodynamic stability (ΔG_j), binding affinity, and probability (P_j). The addition of the allosteric ligand perturbs the system both directly and indirectly. Directly, it binds those microstates that

[147] To elaborate, an ideal gas molecule traveling with a certain momentum has exactly the same energy in a hot and a cold gas. In a hot gas there are just more molecules that have high momentum. In the ensemble representation of allostery, ligand binding changes the energy associated with the individual microstates. The microstates thus have different properties, and the changes in distributions are the result of these changes in the microstates. I thank Gal Ben Porath for assistance with this example.

have the appropriate binding site and stabilizes them, lowering their ΔG values. This change in the thermodynamic stability of certain microstates then indirectly affects the Pj values for all the microstates, since Pj is a relative rather than an individual property of a microstate. The perturbation of the ensemble by the allosteric ligand therefore remodels the energy landscape of the entire system: it alters the properties of all the microstate structures in the system, resetting their ΔGj or Pj values (Motlagh et al. 2014). The perturbation to the system leads to the redistribution of individual protein molecules across microstates, similar to the physics case, but in this case, the redistribution occurs because the energetic properties of the microstates themselves were differentially altered by the allosteric ligand. In this way, ensemble explanations of protein behavior are different than the change of distribution associated with change of temperature of a gas in a box.

The major innovation of ensemble explanations, however, is the application of insights from thermodynamics to a biological context. In particular, these explanations have been developed as an alternative to mechanistic explanations of protein behavior. Since the heyday of classical molecular biology, most research on protein function has been driven by a commitment to the structure-function rule, which holds that protein structure determines protein function (Sarkar 2008). Commitment to this rule has led to the dominance of mechanistic explanations in the molecular life sciences (Machamer et al. 2000, Craver 2007).

Under the mechanistic framework, scientists cite structural changes in proteins to explain their behavior. For example, mechanistic explanations of allostery look for a causal pathway through the protein that connects the allosteric binding site to the second binding site. Structural changes in the protein that occur at the allosteric site upon binding the allosteric ligand are presumed to ripple through the protein, causing a series of local structural changes that ultimately

affect the second binding site. Biophysicist Alan Cooper describes this type of explanation for allostery as following a "clockwork model," which posits "a very discrete…sequence of events like a clockwork mechanism" that connects the two binding sites. This mechanism aims to explain allostery through analogies to "rods, pullies, [and] a mechanical mechanism."[148] Since the earliest abstract mechanistic explanations of allostery were developed in the 1960s, many scientists working in this field have tried to develop mechanistic explanations of this sort, analyzing structural changes in allosteric proteins (Perutz 1970).

Although this research has produced some explanatory successes, the mechanistic strategy has a major shortcoming in that it leads to an almost exclusive focus on protein structure rather than dynamics (Hilser et al. 2012). There is good theoretical and empirical evidence that proteins are highly dynamic molecules, and we now know that many proteins exhibit functional behavior, such as allostery, even though they lack a rigid 3D structure (Wright and Dyson 1999, Hilser and Thompson 2007, Tompa 2011). But mechanistic explanations of protein behavior cannot easily take into account protein dynamics, since they rely on average 3D protein structures.

Ensemble explanations were developed to correct this shortcoming. In ensemble explanations, the representation of the protein as an ensemble of microstate structures replaces the average structure used in mechanistic explanations. The ensemble representation encodes the dynamic properties of the protein in solution: more thermodynamically stable structures are more probable within the equilibrium distribution. Reasoning about changes in the equilibrium distribution of microstates rather than changes in the average structure therefore ensures that ensemble explanations consider both structural and dynamic changes as potential causes of protein behavior. They can provide better explanations of protein behavior, since they consider all and

[148] Video interview by the author with Alan Cooper, 27 May 2019.

only explanatorily relevant features of proteins. In cases in which the allosteric effect is caused solely by changes in dynamics, ensemble explanations provide the only explanation. In fact, the earliest account of dynamic allostery focused on just such cases—i.e., cases in which allostery coupling occurred because of changes in protein dynamics without any change in the average structure of the protein (Cooper and Dryden 1984). The thermodynamic description of the ensemble also facilitates the development of quantitative accounts of allosteric coupling, identifying both the sign and magnitude of the energetic coupling between two sites on a protein (Hilser et al. 2012). This allows for the comparison of allosteric effects within different protein-ligand pairs. In sum, one major innovation of ensemble explanations arises from the application of thermodynamic reasoning to proteins and ability of the ensemble to represent the dynamic features of proteins typically excluded from mechanistic explanations.

A final advantage of ensemble explanations—one emphasized by their scientific promoters (Motlagh et al. 2014)—is their ability to provide a unified framework for explaining allostery and other protein behaviors. These explanations cite a change in the equilibrium distribution of microstates to explain allostery. The allosteric ligand perturbs the thermodynamic system driving it from one equilibrium distribution to another, and this change in the distribution explains allosteric coupling. For example, if the new distribution favors microstates that can bind the second ligand, then the allosteric ligand is a positive effector and the protein-ligand system exhibits positive allostery. This explanation of allostery focuses attention on the equilibrium distribution of microstates rather than the average protein structure. The equilibrium distribution is an intermediate causal variable that is shared by all cases of allostery. Thus, one important result of the development of ensemble explanations of allostery is that these explanations reveal the equilibrium distribution to be a causal variable in a shared causal pathway (Ross 2016, 2018). Any

perturbation to the system that alters the equilibrium distribution of microstates has the potential to produce a positive or negative allosteric effect (Fig. 15).

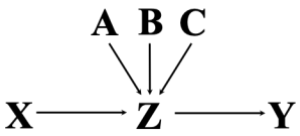

Figure 15 Unified Causal Structure of Ensemble Explanation.

The directed graph shows how interventions on the equilibrium distribution (Z) can mediate the causal relationship between the allosteric ligand (X) and the allosteric effect (Y). Anything that can alter the equilibrium distribution of microstates, by affecting their relative stabilities, has the potential to alter a protein's affinity for a ligand. In the case of allostery, the intervention is an allosteric ligand (X), but other perturbations to the system will also drive it from one equilibrium distribution to another. The additional causal variables in the directed graph in Figure 15—viz., A, B, and C—could be temperature, pH, or another ligand. Any perturbation to the system is funneled through the common pathway variable: Z, the equilibrium distribution.

Ensemble explanations thus have additional benefits over mechanistic explanations of allostery. For one, they can capture all cases of allostery within this unifying framework, even those cases in which there are no changes in average protein structure. Moreover, the causal structure of ensemble explanations reveals the equilibrium distribution as a common pathway variable that accounts for this protein behavior.[149] By shifting scientific attention from away from average protein structure and toward the equilibrium distribution of structural microstates, the ensemble framework helps guide attempts to explain and intervene upon protein behavior.

[149] Although I have focused on the causal structure here, this account of unification has affinities with Kitcher's (1989) account of explanatory unification. The thermodynamic framework of the ensemble explanation is a shared pattern of explanation that can be used to explain any case of allostery.

5.5 Conclusion

In this chapter, I have presented and analyzed a novel type of explanation in molecular and structural biology. These ensemble explanations merit philosophical analysis since they are becoming increasingly common and are strikingly different from mechanistic explanations. My analysis has shown that ensemble explanations of protein behavior do not match philosophical accounts of equilibrium explanation. Although they involve equilibrium states, any single equilibrium state does little explanatory work, since the system is assumed to be at equilibrium unless perturbed. I have argued that it is the *change in equilibrium*, caused by the allosteric ligand, that ensemble explanations cite to explain allostery. The addition of the allosteric ligand perturbs the system, driving it from one equilibrium distribution to another, thereby affecting its binding behavior with respect to a second ligand. This description suggested not an equilibrium interpretation but rather an interventionist interpretation of ensemble explanations. I therefore turned to Zwier's (2017) account of interventionist causation in thermodynamics. Most thermodynamic reasoning aims to explain the movement of a system from one equilibrium state to another, caused by some manipulation of the system. According to Zwier, thermodynamic explanations of this sort fit the interventionist account and are therefore causal explanations. Following similar reasoning, I have argued that ensemble explanations are also best interpreted as causal explanations. Although the interventionist account of thermodynamic reasoning sheds light on how ensemble explanations work, I have shown that there are also certain key differences between causal explanations in the thermodynamic context and in the biological context. Ultimately, I have argued that the major import of ensemble explanation of protein behavior lies in their differences from the mechanistic explanations that dominate structural and molecular biology.

6.0 Conclusion

Proteins, we can conclude, are an object as fruitful for historical and philosophical study as they are for scientific study. This dissertation has demonstrated that proteins, specifically, as well as the history of the molecular life sciences, more generally, offer fertile ground for philosophers of science and biology interested in scientific explanation, representation, and conceptual change.

I have focused on the recent history of structural biology, during a time in which the molecular sciences underwent a dramatic shift in worldview from static to dynamic. In the first half of the dissertation, I have developed a historical account of protein science over the last century, explaining the origins of the static view in the early twentieth century and the eventual emergence of the dynamic view of proteins in the latter half of the century. I have argued that explanations of protein function, such as allostery and catalysis, did not represent a serious challenge to the static view in the 1950s and 1960s. In fact, even though they admitted a limited role for protein dynamics, I have argued that these early accounts of protein function were compatible with the structure-function rule as well as the static view of proteins. They were representative of most explanations of protein behavior at that time, providing mechanistic accounts that treated proteins like molecular machines. Looking at the transition from the static to the dynamic view, I have argued that theory and theoretical commitments played an important and overlooked role in the emergence and spread of the dynamic view of proteins. Before the great technological advances in structural biology at the turn of the century, certain theoretically trained scientists were already committed to treating proteins as small thermodynamic systems, and these

scientists led the search for empirical anomalies in the 1980s and 1990s that would support the dynamic rather than static view of proteins.

In the second half of the dissertation, I have analyzed certain representational and explanatory practices in protein science. I have characterized the epistemic relationships that obtain between the structural and dynamic protein concepts and their associated representations. I have argued that conceptual replacement is happening in protein science. On my account, the dynamic concept of the protein, which maintains the importance of both structure and dynamics, is replacing the structural concept, with its exclusive focus on structure. Representations compatible with the structural and dynamic concept, I have argued, are typically related via abstraction. Turning to explanation in structural biology, I have presented an account of ensemble explanations, which were developed to explain the role of dynamics in protein function. I have argued that these are distinct from mechanistic explanations, which rely on static structural representations of proteins. As I have shown, these explanations capture protein dynamics by representing proteins not as single, rigid structures but as ensembles of structurally distinct microstates. Although they involve equilibrium distributions, I have argued that these explanations are different from the standard accounts of equilibrium explanation in the philosophical literature. Instead, I have argued that ensemble explanations are a type of causal explanation. My account has thus demonstrated how scientists have had to develop new explanatory strategies to account for the complexity introduced by dynamics.

Looking ahead, there is still much untapped potential for philosophical work in the molecular life sciences. I have touched on some areas for additional work during the course of the dissertation, but two future directions are worth elaborating here. First, my account of ensemble explanation has left the question of reduction largely unexplored. While I have argued that this

type of explanation is distinct from mechanistic explanation, I have not argued that ensemble explanations of protein behavior are either reductive or emergent. However, I believe that further analysis of these explanations and similar explanations that consider dynamic will prove illuminating for this philosophical topic. I have a hunch that these explanations resist the standard classification of reductive and emergent explanation, although they have affinities with both. If my hunch proves correct, then the failure of ensemble explanations to fit neatly within this debate will suggest that the current philosophical debate between reductionists and anti-reductionists offers us a false choice, obscuring a class of explanations that falls between the two extremes. Considering this novel type of explanation in protein science, therefore, might help us to rethink this seemingly intractable debate.

Second, my dissertation has only scratched the surface when it comes to understanding how scientists working in the molecular life sciences incorporate knowledge from different perspectives. I have shown how borrowing from the representational and explanatory practices in physics has been important for the emergence of the dynamic view of proteins and the development of ensemble explanations in structural biology. However, more work needs to be done to fully understand the successes and failures of different methodological and epistemic strategies, such as reduction, unification, and integration, in solving the problem of connecting knowledge from different perspectives. Because the molecular life sciences are at the intersection of biology, chemistry, and physics, successful explanations and interventions likely require different strategies to combine these perspectives. With careful philosophical analysis of these strategies and the conditions that determine their success, we can hopefully address questions of how science ought to proceed if we want to intervene in the world in various ways.

Bibliography

Akke, M., R Bruschweiler, and A.G. Palmer. 1993. "NMR Order Parameters and Free-Energy: An Analytical Approach and Its Application to Cooperative Ca2+ Binding by Calbindin-D(9K)." *Journal of the American Chemical Society* 115 (21): 9832–33. https://doi.org/10.1021/ja00074a073.

Akke, M., P.A. Carr, and A.G. Palmer. 1994. "Heteronuclear-Correlation NMR-Spectroscopy with Simultaneous Isotope Filtration, Quadrature Detection, and Sensitivity Enhancement Using Z-Rotations." *Journal of Magnetic Resonance Series B* 104 (3): 298–302. https://doi.org/10.1006/jmrb.1994.1090.

Akke, M., S. Forsen, and W.J. Chazin. 1991. "Molecular-Basis for Cooperativity in Ca2+ Binding to Calbindin-D9k - H-1 Nuclear-Magnetic-Resonance Studies of (Cd2+)1-Bovine Calbindin-D9k." *Journal of Molecular Biology* 220 (1): 173–89. https://doi.org/10.1016/0022-2836(91)90389-N.

Anfinsen, Christian B., and Robert R. Redfield. 1956. "Protein Structure in Relation to Function and Biosynthesis." In *Advances in Protein Chemistry*, edited by M. L. Anson, Kenneth Bailey, and John T. Edsall, 11:1–100. Academic Press. https://doi.org/10.1016/S0065-3233(08)60420-9.

Barnett, James A., and Frieder W. Lichtenthaler. 2001. "A History of Research on Yeasts 3: Emil Fischer, Eduard Buchner and Their Contemporaries, 1880–1900." *Yeast* 18 (4): 363–88. https://doi.org/10.1002/1097-0061(20010315)18:4<363::AID-YEA677>3.0.CO;2-R.

Batterman, Robert. 2002. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford: Oxford University Press.

———. 2013. "The Tyranny of Scales." In *The Oxford Handbook of Philosophy of Physics*, edited by Robert W Batterman, 255–86. New York: Oxford University Press.

Beatty, John. 1997. "Why Do Biologists Argue like They Do?" *Philosophy of Science* 64: S432–43. https://doi.org/10.1086/392620.

Bechtel, William, and Adele Abrahamsen. 2005. "Explanation: A Mechanist Alternative." *Studies in History and Philosophy of Biological and Biomedical Sciences* 36 (2): 421–41. https://doi.org/10.1016/j.shpsc.2005.03.010.

Bechtel, William, and Robert C Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, Mass.: MIT Press.

Berman, H. M. 2008. “The Protein Data Bank: A Historical Perspective.” *Acta Crystallographica Section A: Foundations of Crystallography* 64 (1): 88–95. https://doi.org/10.1107/S0108767307035623.

Bernal, J. D. 1939. “Structure of Proteins.” *Nature* 143 (3625): 663. https://doi.org/10.1038/143663a0.

Bernal, J. D., and D. Crowfoot. 1934. “X-Ray Photographs of Crystalline Pepsin.” *Nature* 133 (3369): 794–95. https://doi.org/10.1038/133794b0.

Bonner, Philip L. R. 2019. *Protein Purification*. London: Taylor and Francis.

Brandon, Robert N. 1990. *Adaptation and Environment*. Princeton, N.J.: Princeton University Press.

Braxton, B. L., L. S. Mullins, F. M. Raushel, and G. D. Reinhart. 1996. “Allosteric Effects of Carbamoyl Phosphate Synthetase from Escherichia Coli Are Entropy-Driven.” *Biochemistry* 35 (36): 11918–24. https://doi.org/10.1021/bi961305m.

Brigandt, Ingo. 2010. “The Epistemic Goal of a Concept: Accounting for the Rationality of Semantic Change and Variation.” *Synthese* 177 (1): 19–40. https://doi.org/10.1007/s11229-009-9623-8.

Brigandt, Ingo, and Alan Love. 2017. “Reductionism in Biology.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2017. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2017/entries/reduction-biology/.

Brooks, Charles L, Martin Karplus, and Bernard Montgomery Pettitt. 1988. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*. New York: Wiley.

Brunori, M. 1999. “Hemoglobin Is an Honorary Enzyme.” *Trends in Biochemical Sciences* 24 (4): 158–61. https://doi.org/10.1016/s0968-0004(99)01380-8.

Bueno, Otávio, Steven French, and James Ladyman. 2012. “Models and Structures: Phenomenological and Partial.” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 43 (1): 43–46. https://doi.org/10.1016/j.shpsb.2011.12.001.

Cartwright, Nancy, Towfic Shomar, and Mauricio Suárez. 1995. “The Tool Box of Science: Tools for the Building of Models with a Superconductivity Example.” *Poznan Studies in the Philosophy of the Sciences and the Humanities* 44: 137–49.

Chadarevian, Soraya de. 2018. “John Kendrew and Myoglobin: Protein Structure Determination in the 1950s.” *Protein Science* 27 (6): 1136–43. https://doi.org/10.1002/pro.3417.

Chang, Raymond. 2003. *General Chemistry: The Essential Concepts*. Dubuque, Iowa: McGraw-Hill.

Changeux, J.P. 1961. “The Feedback Control Mechanism of Biosynthetic L-Threonine Deaminase by L-Isoleucine.” *Cold Spring Harbor Symposia on Quantitative Biology* 26 (January): 313–18. https://doi.org/10.1101/SQB.1961.026.01.037.

———. 1963. “Allosteric Interactions on Biosynthetic L-Threonine Deaminase from E. Coli K12.” *Cold Spring Harbor Symposia on Quantitative Biology* 28 (0): 497–504. https://doi.org/10.1101/SQB.1963.028.01.066.

———. 1993. “Allosteric Proteins: From Regulatory Enzymes to Receptors--Personal Recollections.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 15 (9): 625–34. https://doi.org/10.1002/bies.950150909.

———. 2011. “50th Anniversary of the Word ‘Allosteric.’” *Protein Science : A Publication of the Protein Society* 20 (7): 1119–24. https://doi.org/10.1002/pro.658.

———. 2012. “Allostery and the Monod-Wyman-Changeux Model After 50 Years.” *Annual Review of Biophysics* 41 (1): 103–33. https://doi.org/10.1146/annurev-biophys-050511-102222.

———. 2013. “The Origins of Allostery: From Personal Memories to Material for the Future.” *Journal of Molecular Biology*, Allosteric Interactions and Biological Regulation (Part I), 425 (9): 1396–1406. https://doi.org/10.1016/j.jmb.2013.02.033.

Changeux, J.P., and S.J. Edelstein. 2006. “Allosteric Receptors after 30 Years.” *Rendiconti Lincei* 17 (1): 59–96. https://doi.org/10.1007/BF02904502.

Clarkson, M. W., and A. L. Lee. 2004. “Long-Range Dynamic Effects of Point Mutations Propagate through Side Chains in the Serine Protease Inhibitor Eglin c.” *Biochemistry* 43 (39): 12448–58. https://doi.org/10.1021/bi0494424.

Cooper, A. 1976. “Thermodynamic Fluctuations in Protein Molecules.” *Proceedings of the National Academy of Sciences* 73 (8): 2740–41. https://doi.org/10.1073/pnas.73.8.2740.

———. 1980. “Conformational Fluctuation and Change in Biological Macromolecules.” *Science Progress* 66 (264): 473–97. https://www.jstor.org/stable/43420497.

———. 1983. “Protein Structure and Dynamics: A Daresbury Study Weekend (Nov 19-20, 1983).” In *Information Quarterly for Protein Crystallography*, edited by P. Machin, J.W. Campbell, and M. Elder, 12:36.

———. 1984. “Protein Fluctuations and the Thermodynamic Uncertainty Principle.” *Progress in Biophysics and Molecular Biology* 44 (3): 181–214. https://doi.org/10.1016/0079-6107(84)90008-7.

Cooper, A., and D. T. F. Dryden. 1984. “Allostery without Conformational Change.” *European Biophysics Journal* 11 (2): 103–9. https://doi.org/10.1007/BF00276625.

Cooper, A., B. Sanders, and D.T.F. Dryden. 1989. "Dynamics of Genetically-Engineered Enzymes - Fluorescence and Depolarized Rayleigh-Scattering Studies of Yeast Phosphoglycerate Kinase." *Journal of Molecular Liquids* 42 (October): 99–111. https://doi.org/10.1016/0167-7322(89)80028-5.

Craver, Carl F. 2007a. *Explaining the Brain*. Oxford: Oxford University Press.

———. 2007b. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon Press.

Craver, Carl F., and David M. Kaplan. 2020. "Are More Details Better? On the Norms of Completeness for Mechanistic Explanations." *The British Journal for the Philosophy of Science* 71 (1): 287–319. https://doi.org/10.1093/bjps/axy015.

Creager, Angela N. H., and Jean-Paul Gaudillière. 1996. "Meanings in Search of Experiments and Vice-Versa: The Invention of Allosteric Regulation in Paris and Berkeley, 1959-1968." *Historical Studies in the Physical and Biological Sciences* 27 (1): 1–89. https://doi.org/10.2307/27757769.

Creighton, T.E. 1978. "Experimental Studies of Protein Folding and Unfolding." *Progress in Biophysics & Molecular Biology* 33 (3): 231–97. https://doi.org/10.1016/0079-6107(79)90030-0.

Cui, Qiang, and Martin Karplus. 2008. "Allostery and Cooperativity Revisited." *Protein Science* 17 (8): 1295–1307. https://doi.org/10.1110/ps.03259908.

Darden, Lindley. 1996. "Generalizations in Biology." *Studies in History and Philosophy of Science Part A* 27 (3): 409–19. https://doi.org/10.1016/0039-3681(95)00050-X.

Debrunner, P.G., and H. Frauenfelder. 1982. "Dynamics of Proteins." *Annual Review of Physical Chemistry* 33 (1): 283–99. https://doi.org/10.1146/annurev.pc.33.100182.001435.

Demchenko, A.P. 1986. "Fluorescence Analysis of Protein Dynamics." *Essays in Biochemistry* 22: 120–57.

Dixon, Malcolm, and Edwin C Webb. 1964. *Enzymes*. New York, NY: Academic Press.

Drenth, Jan. 1994. *Principles of Protein X-Ray Crystallography*. New York, NY: Springer.

Dryden, David T. F. 1988. "Functional Consequences of Protein Dynamics." PhD, Ann Arbor, MI: ProQuest Dissertations & Theses. http://theses.gla.ac.uk/77669/.

Dupré, John. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.

Edelstein, Stuart J. 2013. "Allosteric Interactions after 50 Years." *Journal of Molecular Biology*, Allosteric Interactions and Biological Regulation (Part I), 425 (9): 1391–95. https://doi.org/10.1016/j.jmb.2013.03.026.

Englander, S. Walter, and Neville R. Kallenbach. 1984. “Hydrogen Exchange and Structural Dynamics of Proteins and Nucleic Acids.” *Quarterly Reviews of Biophysics* 16 (4): 521–655. https://doi.org/10.1017/S0033583500005217.

Epstein, Brian, and Patrick Forber. 2013. “The Perils of Tweaking: How to Use Macrodata to Set Parameters in Complex Simulation Models.” *Synthese* 190 (2): 203–18. https://doi.org/10.1007/s11229-012-0142-7.

Feynman, Richard P. 1963. *The Feynman Lectures on Physics*. Vol. Vol I, Ch 3. Reading, MA: Addison-Wesley.

Fischer, Emil. 1894. “Einfluss Der Configuration Auf Die Wirkung Der Enzyme.” *Berichte Der Deutschen Chemischen Gesellschaft* 27 (3): 2985–93. https://doi.org/10.1002/cber.18940270364.

———. 1906. *Untersuchungen Über Aminosäuren, Polypeptide Und Proteine*. Vol. 2. J. Springer.

Fisher, R.A. 1931. *The Genetical Theory of Natural Selection*. New York: Dover.

Forman-Kay, Julie D., and Tanja Mittag. 2013. “From Sequence and Forces to Structure, Function, and Evolution of Intrinsically Disordered Proteins.” *Structure* 21 (9): 1492–99. https://doi.org/10.1016/j.str.2013.08.001.

Frauenfelder, H. 1983. “Summary and Outlook.” In *Mobility and Function in Proteins and Nucleic Acids*, edited by R. Porter, M. O’Connor, and J. Whelan, 93:329–39. London: Pitman.

———. 2012. “Giorgio Careri: A Physicist in the Life Sciences.” *Journal of Biological Physics* 38 (1): 3. https://doi.org/10.1007/s10867-011-9258-0.

Frauenfelder, H., G.A. Petsko, and D. Tsernoglou. 1979. “Temperature-Dependent x-Ray-Diffraction as a Probe of Protein Structural Dynamics.” *Nature* 280 (5723): 558–63. https://doi.org/10.1038/280558a0.

Frederick, Kendra King, Michael S. Marlow, Kathleen G. Valentine, and A. Joshua Wand. 2007. “Conformational Entropy in Molecular Recognition by Proteins.” *Nature* 448 (7151): 325-U3. https://doi.org/10.1038/nature05959.

Frigg, Roman. 2008. “A Field Guide to Recent Work on the Foundations of Statistical Mechanics.” In *The Ashgate Companion to Contemporary Philosophy of Physics*, edited by Dean Rickles. Aldershot, UK: Ashgate Publishers.

Fruton, Joseph S. 1999. *Proteins, Enzymes, Genes: The Interplay of Chemistry and Biology*. New Haven, CT: Yale University Press.

Garrett, Reginald H, and Charles M Grisham. 2010. *Biochemistry*. Boston, MA: Brooks/Cole.

Gerhart, J. C., and A. B. Pardee. 1963. “The Effect of the Feedback Inhibitor, CTP, on Subunit Interactions in Aspartate Transcarbamylase.” *Cold Spring Harbor Symposia on Quantitative Biology* 28 (0): 491–96. https://doi.org/10.1101/SQB.1963.028.01.065.

———. 1964. “Aspartate Transcarbamylase, an Enzyme Designed for Feedback Inhibition.” *Federation Proceedings* 23 (June): 727–35.

Gerhart, J.C. 2014. “From Feedback Inhibition to Allostery: The Enduring Example of Aspartate Transcarbamoylase.” *FEBS Journal* 281 (2): 612–20. https://doi.org/10.1111/febs.12483.

Gerhart, J.C., and A.B. Pardee. 1962. “The Enzymology of Control by Feedback Inhibition.” *Journal of Biological Chemistry* 237 (3): 891–96. https://doi.org/10.1016/S0021-9258(18)60389-8.

Gingras, Yves. 2010. “Revisiting the ‘Quiet Debut’ of the Double Helix: A Bibliometric and Methodological Note on the ‘Impact’ of Scientific Publications.” *Journal of the History of Biology* 43 (1): 159–81. https://doi.org/10.1007/s10739-009-9183-2.

Goodsell, David S., and Arthur J. Olson. 2000. “Structural Symmetry and Protein Function.” *Annual Review of Biophysics and Biomolecular Structure* 29 (1): 105–53. https://doi.org/10.1146/annurev.biophys.29.1.105.

Gould, S. J., and R. C. Lewontin. 1979. “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme.” *Proceedings of the Royal Society of London B: Biological Sciences* 205 (1161): 581–98. https://doi.org/10.1098/rspb.1979.0086.

Green, D. W., Vernon Martin Ingram, and Max Ferdinand Perutz. 1954. “The Structure of Haemoglobin-IV. Sign Determination by the Isomorphous Replacement Method.” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 225 (1162): 287–307. https://doi.org/10.1098/rspa.1954.0203.

Green, Eric D., James D. Watson, and Francis S. Collins. 2015. “Human Genome Project: Twenty-Five Years of Big Biology.” *Nature* 526 (7571): 29–31. https://doi.org/10.1038/526029a.

Griffiths, Paul, and Karola Stotz. 2013. *Genetics and Philosophy: An Introduction*. Cambridge University Press.

Gunasekaran, K., Buyong Ma, and Ruth Nussinov. 2004. “Is Allostery an Intrinsic Property of All Dynamic Proteins?” *Proteins: Structure, Function, and Bioinformatics* 57 (3): 433–43. https://doi.org/10.1002/prot.20232.

Hilser, Vincent J. 2010. “An Ensemble View of Allostery.” *Science* 327 (5966): 653–54. https://doi.org/10.1126/science.1186121.

Hilser, Vincent J., Bertrand García-Moreno E., Terrence G. Oas, Greg Kapp, and Steven T. Whitten. 2006. “A Statistical Thermodynamic Model of the Protein Ensemble.” *Chemical Reviews* 106 (5): 1545–58. https://doi.org/10.1021/cr040423+.

Hilser, Vincent J., and E. Brad Thompson. 2007. “Intrinsic Disorder as a Mechanism to Optimize Allosteric Coupling in Proteins.” *Proceedings of the National Academy of Sciences* 104 (20): 8311–15. https://doi.org/10.1073/pnas.0700329104.

Hilser, Vincent J., James O. Wrabl, and Hesam N. Motlagh. 2012. “Structural and Energetic Basis of Allostery.” *Annual Review of Biophysics* 41 (1): 585–609. https://doi.org/10.1146/annurev-biophys-050511-102319.

Hitchcock, Christopher, and James Woodward. 2003. “Explanatory Generalizations, Part II: Plumbing Explanatory Depth.” *Noûs* 37 (2): 181–99. https://www.jstor.org/stable/3506081.

Hofmeister, F. 1902. “Über Bau und Gruppierung der Eiweisskörper.” *Ergebnisse der Physiologie* 1 (1): 759–802. https://doi.org/10.1007/BF02188398.

Hull, David L. 1972. “Reduction in Genetics–Biology or Philosophy?” *Philosophy of Science* 39 (4): 491–99. https://doi.org/10.1086/288470.

Huttemann, A., and A. C. Love. 2011. “Aspects of Reductive Explanation in Biological Science: Intrinsicality, Fundamentality, and Temporality.” *The British Journal for the Philosophy of Science* 62 (3): 519–49. https://doi.org/10.1093/bjps/axr006.

Igumenova, Tatyana I., Kendra King Frederick, and A. Joshua Wand. 2006. “Characterization of the Fast Dynamics of Protein Amino Acid Side Chains Using NMR Relaxation in Solution.” *Chemical Reviews* 106 (5): 1672–99. https://doi.org/10.1021/cr040422h.

Jacob, Francois, and Jacques Monod. 1963. “Genetic Repression, Allosteric Inhibition, and Cellular Differentiation.” In *Cytodifferentiation and Macromolecular Synthesis*, edited by Michael Locke, 30–64. New York, NY: Academic Press.

Jameson, D.M. 1998. “Gregorio Weber, 1916-1997: A Fluorescent Lifetime.” *Biophysical Journal* 75 (1): 419–21. https://doi.org/10.1016/S0006-3495(98)77528-9.

Johnson, J.L., and G.D. Reinhart. 1994a. “Influence of Mgadp on Phosphofructokinase from Escherichia-Coli - Elucidation of Coupling Interactions with Both Substrates.” *Biochemistry* 33 (9): 2635–43. https://doi.org/10.1021/bi00175a036.

———. 1994b. “Influence of Substrates and Mgadp on the Time-Resolved Intrinsic Fluorescence of Phosphofructokinase from Escherichia-Coli - Correlation.” *Biochemistry* 33 (9): 2644–50. https://doi.org/10.1021/bi00175a037.

———. 1997. “Failure of a Two-State Model to Describe the Influence of Phospho(Enol)Pyruvate on Phosphofructokinase from Escherichia Coli.” *Biochemistry* 36 (42): 12814–22. https://doi.org/10.1021/bi970942p.

Kaiser, David. 2012. “In Retrospect: The Structure of Scientific Revolutions.” *Nature* 484 (7393): 164–65. https://doi.org/10.1038/484164a.

Kaplan, David Michael, and Carl F. Craver. 2011. “The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective.” *Philosophy of Science* 78 (4): 601–27. https://doi.org/10.1086/661755.

Karplus, Martin. 1988. “Preface.” In *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, edited by Charles L Brooks, Martin Karplus, and Bernard Montgomery Pettitt. New York: Wiley.

———. 2006. “Spinach on the Ceiling: A Theoretical Chemist’s Return to Biology.” *Annual Review of Biophysics and Biomolecular Structure* 35 (1): 1–47. https://doi.org/10.1146/annurev.biophys.33.110502.133350.

Karplus, Martin, and J. Andrew McCammon. 1986. “The Dynamics of Proteins.” *Scientific American* 254 (4): 42–51. https://doi.org/10.1038/scientificamerican0486-42.

Ke, Qing, Emilio Ferrara, Filippo Radicchi, and Alessandro Flammini. 2015. “Defining and Identifying Sleeping Beauties in Science.” *Proceedings of the National Academy of Sciences* 112 (24): 7426–31. https://doi.org/10.1073/pnas.1424329112.

Kendrew, J. C., G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips. 1958. “A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis.” *Nature* 181 (4610): 662–66. https://doi.org/10.1038/181662a0.

Kendrew, John C. 1961. “The Three-Dimensional Structure of a Protein Molecule.” *Scientific American* 205 (6): 96–111. https://doi.org/10.1038/scientificamerican1261-96.

Kitcher, Philip. 1978. “Theories, Theorists and Theoretical Change.” *The Philosophical Review* 87 (4): 519–47. https://doi.org/10.2307/2184458.

———. 1982. “Genes.” *The British Journal for the Philosophy of Science* 33 (4): 337–59. https://doi.org/10.1093/bjps/33.4.337.

———. 1984. “1953 and All That. A Tale of Two Sciences.” *The Philosophical Review* 93 (3): 335–73. https://doi.org/10.2307/2184541.

———. 1989. “Explanatory Unification and the Causal Structure of the World.” In *Scientific Explanation*, edited by Philip Kitcher and Wesley C Salmon. Minneapolis, MN: University of Minnesota Press.

Koerdel, Johan, Nicholas J. Skelton, Mikael Akke, Arthur G. Palmer, and Walter J. Chazin. 1992. “Backbone Dynamics of Calcium-Loaded Calbindin D9k Studied by Two-Dimensional Proton-Detected Nitrogen-15 NMR Spectroscopy.” *Biochemistry* 31 (20): 4856–66. https://doi.org/10.1021/bi00135a017.

Koonin, Eugene V. 2016. “Splendor and Misery of Adaptation, or the Importance of Neutral Null for Understanding Evolution.” *BMC Biology* 14 (1): 114. https://doi.org/10.1186/s12915-016-0338-2.

Koshland, D. E. 1958. “Application of a Theory of Enzyme Specificity to Protein Synthesis.” *Proceedings of the National Academy of Sciences* 44 (2): 98–104. https://doi.org/10.1073/pnas.44.2.98.

———. 1959. “Enzyme Flexibility and Enzyme Action.” *Journal of Cellular and Comparative Physiology* 54 (S1): 245–58. https://doi.org/10.1002/jcp.1030540420.

———. 1963. “Correlation of Structure and Function in Enzyme Action.” *Science* 142 (3599): 1533–41. https://doi.org/10.1126/science.142.3599.1533.

Koshland, D. E., G. Némethy, and D. Filmer. 1966. “Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits.” *Biochemistry* 5 (1): 365–85. https://doi.org/10.1021/bi00865a047.

Koshland, D.E. 1994. “The Key–Lock Theory and the Induced Fit Theory.” *Angewandte Chemie International Edition in English* 33 (23–24): 2375–78. https://doi.org/10.1002/anie.199423751.

Kuhn, Thomas S. 1977. *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: The Univ. of Chicago Press.

———. 2012. *The Structure of Scientific Revolutions*. Chicago, Ill.: The University of Chicago Press.

Kuorikoski, Jaakko. 2007. “Explaining with Equilibria.” In *Rethinking Explanation*, edited by Johannes Persson and Petri Ylikoski, 149–62. Boston Studies in the Philosophy of Science. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-5581-2_11.

Lean, Oliver M. 2020. “Binding Specificity and Causal Selection in Drug Design.” *Philosophy of Science* 87 (1): 70–90. https://doi.org/10.1086/706093.

Levy, A. 2014. “What Was Hodgkin and Huxley’s Achievement?” *The British Journal for the Philosophy of Science* 65 (3): 469–92. https://doi.org/10.1093/bjps/axs043.

Liu, Jin, and Ruth Nussinov. 2016. “Allostery: An Overview of Its History, Concepts, Methods, and Applications.” *PLOS Computational Biology* 12 (6): e1004966. https://doi.org/10.1371/journal.pcbi.1004966.

Liu, T., S. T. Whitten, and V. J. Hilser. 2006. “Ensemble-Based Signatures of Energy Propagation in Proteins: A New View of an Old Phenomenon.” *Proteins-Structure Function and Bioinformatics* 62 (3): 728–38. https://doi.org/10.1002/prot.20749.

London, Fritz. 1950. *Superfluids, Vol I.* New York, NY: Wiley.

Low, Barbara W. 1953. “The Structure and Configuation of Amino Acids, Peptides, and Proteins.” In *The Proteins: Chemistry, Biological Activity, and Methods*, edited by Hans Neurath and Kenneth Bailey. Vol. Vol. 1, Part A. New York, NY: The Academic Press, Inc.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. “Thinking about Mechanisms.” *Philosophy of Science* 67 (1): 1–25. https://doi.org/10.2307/188611.

McCammon, J. Andrew, Bruce R. Gelin, and Martin Karplus. 1977. “Dynamics of Folded Proteins.” *Nature* 267 (5612): 585. https://doi.org/10.1038/267585a0.

McMullin, Ernan. 1985. “Galilean Idealization.” *Studies in History and Philosophy of Science Part A* 16 (3): 247–73. https://doi.org/10.1016/0039-3681(85)90003-2.

Mehellou, Youcef, and Erik De Clercq. 2010. “Twenty-Six Years of Anti-HIV Drug Discovery: Where Do We Stand and Where Do We Go?” *Journal of Medicinal Chemistry* 53 (2): 521–38. https://doi.org/10.1021/jm900492g.

Michaelis, Leonor, and Maud L. Menten. 1913. “Die Kinetik Der Invertinwirkung.” *Biochem. z* 49 (333–369): 352.

Mitchell, Sandra D. 2002. “Integrative Pluralism.” *Biology and Philosophy* 17 (1): 55–70. https://doi.org/10.1023/A:1012990030867.

———. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge, U.K.; New York: Cambridge University Press.

———. 2012. “Emergence: Logical, Functional and Dynamical.” *Synthese* 185 (2): 171–86. https://doi.org/10.1007/s11229-010-9719-1.

———. 2019. “Perspectives, Representation, and Integration.” In *Understanding Perspectivism*, edited by Michela Massimi and Casey D McCoy. Routledge.

Mitchell, Sandra D., and Angela M. Gronenborn. 2017. “After Fifty Years, Why Are Protein X-Ray Crystallographers Still in Business?” *The British Journal for the Philosophy of Science* 68 (3): 703–23. https://doi.org/10.1093/bjps/axv051.

Monod, J. 1968. “On Symmetry and Function in Biological Systems.” In *Symmetry and Function of Biological Systems at the Macromolecular Level: Proceedings of the Eleventh Nobel Symposium*, edited by A. Engström and B. Strandberg. Vol. 11. Stockholm: Almqvist & Wilksell Forlag.

Monod, J., J.P. Changeux, and F. Jacob. 1963. “Allosteric Proteins and Cellular Control Systems.” *Journal of Molecular Biology* 6 (4): 306–29. https://doi.org/10.1016/S0022-2836(63)80091-1.

Monod, J., and F. Jacob. 1961. “General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation.” *Cold Spring Harbor Symposia on Quantitative Biology* 26 (January): 389–401. https://doi.org/10.1101/SQB.1961.026.01.048.

Monod, J., A. Lwoff, and A. Ullmann. 1979. *Origins of Molecular Biology: A Tribute to Jacques Monod*. New York; London: Academic Press.

Monod, J., J. Wyman, and J. P. Changeux. 1965. “On the Nature of Allosteric Transitions: A Plausible Model.” *Journal of Molecular Biology* 12: 88–118.

Monod, Jacques. 1966. “From Enzymatic Adaptation to Allosteric Transitions.” *Science* 154 (3748): 475–83. https://doi.org/10.1126/science.154.3748.475.

Morange, Michel. 1998. *A History of Molecular Biology*. Cambridge, Mass.: Harvard University Press.

———. 2006. “The Protein Side of the Central Dogma: Permanence and Change.” *History and Philosophy of the Life Sciences* 28 (4): 513–24. https://www.jstor.org/stable/23334183.

———. 2012. “What History Tells Us XXVII.: A New Life for Allostery.” *Journal of Biosciences* 37 (1): 13–17. https://doi.org/10.1007/s12038-012-9186-6.

———. 2017. “What History Tells Us XLII. A ‘New’ View of Proteins.” *Journal of Biosciences* 42 (1): 11–14. https://doi.org/10.1007/s12038-017-9673-x.

———. 2020. *The Black Box of Biology: A History of the Molecular Revolution*. Translated by Matthew Cobb. Cambridge, MA: Harvard University Press.

Motlagh, Hesam N., and Vincent J. Hilser. 2012. “Agonism/Antagonism Switching in Allosteric Ensembles.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (11): 4134–39. https://doi.org/10.1073/pnas.1120519109.

Motlagh, Hesam N., James O. Wrabl, Jing Li, and Vincent J. Hilser. 2014a. “The Ensemble Nature of Allostery.” *Nature* 508 (7496): 331–39. https://doi.org/10.1038/nature13001.

———. 2014b. “The Ensemble Nature of Allostery.” *Nature* 508 (7496): 331–39. https://doi.org/10.1038/nature13001.

Muirhead, Hilary, and M. F. Perutz. 1963. “Structure Of Hæmoglobin: A Three-Dimensional Fourier Synthesis of Reduced Human Haemoglobin at 5.5 Å Resolution.” *Nature* 199 (4894): 633. https://doi.org/10.1038/199633a0.

Nagel, Ernest. 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation.* London: Routlege & K. Paul.

Neal, Jacob P. 2019. “When Causal Specificity Does Not Matter (Much): Insights from HIV Treatment.” *Philosophy of Science* 86 (5): 836–46. https://doi.org/10.1086/705510.

Novick, Aaron, and Leo Szilard. 1954. “Experiments with the Chemostat on the Rates of Amino Acid Synthesis in Bacteria.” In *Dynamics of Growth Processes*, 21. Princeton, N.J.: Princeton University Press.

Nussinov, Ruth, and Chung-Jung Tsai. 2015. “Allostery without a Conformational Change? Revisiting the Paradigm.” *Current Opinion in Structural Biology* 30 (February): 17–24. https://doi.org/10.1016/j.sbi.2014.11.005.

Oldfield, Christopher J., and A. Keith Dunker. 2014. “Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions.” *Annual Review of Biochemistry* 83 (1): 553–84. https://doi.org/10.1146/annurev-biochem-072711-164947.

Pauling, Linus, Robert B. Corey, and H. R. Branson. 1951. “The Structure of Proteins: Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain.” *Proceedings of the National Academy of Sciences* 37 (4): 205–11. https://doi.org/10.1073/pnas.37.4.205.

Perutz, M. F. 1970. “Stereochemistry of Cooperative Effects in Haemoglobin: Haem–Haem Interaction and the Problem of Allostery.” *Nature* 228 (5273): 726–34. https://doi.org/10.1038/228726a0.

Perutz, M. F., M. G. Rossmann, Ann F. Cullis, Hilary Muirhead, Georg Will, and A. C. T. North. 1960. “Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis.” *Nature* 185 (4711): 416. https://doi.org/10.1038/185416a0.

Petter, R.C., J.S. Salek, C.T. Sikorski, G. Kumaravel, and F.T. Lin. 1990. “Cooperative Binding by Aggregated Mono-6-(Alkylamino)-Beta-Cyclodextrins.” *Journal of the American Chemical Society* 112 (10): 3860–68. https://doi.org/10.1021/ja00166a021.

Phillips, DC. 1981. *Biomolecular Stereodynamics*. Edited by Ramaswamy H Sarma. II. New York, NY: Adenine Press.

Porter, Ruth, Maeve O’Connor, and Julie Whelan, eds. 1983. *Mobility and Function in Proteins and Nucleic Acids: Symposium on Mobility and Function in Protein and Nucleic Acids, Held at the Ciba Foundation, London, 2-4 March 1982*. London: Pitman.

Rebek, Julius. 1984. “Binding Forces, Equilibria and Rates: New Models for Enzymic Catalysis.” *Accounts of Chemical Research* 17 (7): 258–64. https://doi.org/10.1021/ar00103a006.

Rebek, Julius, Timothy Costello, Luann Marshall, Ruth Wattley, Robert C. Gadwood, and Kay Onan. 1985. “Allosteric Effects in Organic Chemistry: Binding Cooperativity in a Model for Subunit Interactions.” *Journal of the American Chemical Society* 107 (25): 7481–87. https://doi.org/10.1021/ja00311a043.

Reinhart, G.D. 1983. “The Determination of Thermodynamic Allosteric Parameters of an Enzyme Undergoing Steady-State Turnover.” *Archives of Biochemistry and Biophysics* 224 (1): 389–401. https://doi.org/10.1016/0003-9861(83)90225-4.

———. 1985. “Influence of PH on the Regulatory Kinetics of Rat-Liver Phosphofructokinase -- A Thermodynamic Linked-Function Analysis.” *Biochemistry* 24 (25): 7166–72. https://doi.org/10.1021/bi00346a022.

———. 1988. “Linked-Function Origins of Cooperativity in a Symmetrical Dimer.” *Biophysical Chemistry* 30 (2): 159–72. https://doi.org/10.1016/0301-4622(88)85013-0.

Reinhart, G.D., and S.B. Hartleip. 1986. “Relationship Between Fructose 2,6-Bisphosphate Activation and MgATP Inhibition of Rat-Liver Phosphofructokinase at High PH -- Kinetic Evidence for Individual Binding-Sites Linked by Finite Couplings.” *Biochemistry* 25 (23): 7308–13. https://doi.org/10.1021/bi00371a010.

———. 1987. “Perturbation of the Quaternary Structure and Allosteric Behavior of Rat-Liver Phosphofructokinase by Polyethylene-Glycol.” *Archives of Biochemistry and Biophysics* 258 (1): 65–76. https://doi.org/10.1016/0003-9861(87)90323-7.

Reinhart, G.D., S.B. Hartleip, and M.M. Symcox. 1989. “Role of Coupling Entropy in Establishing the Nature and Magnitude of Allosteric Response.” *Proceedings of the National Academy of Sciences of the United States of America* 86 (11): 4032–36. https://doi.org/10.1073/pnas.86.11.4032.

Richards, FM. 1983. “Introduction.” In *Mobility and Function in Proteins and Nucleic Acids: Symposium on Mobility and Function in Protein and Nucleic Acids, Held at the Ciba Foundation, London, 2-4 March 1982*, edited by Ruth Porter, Maeve O’Connor, and Julie Whelan, 1–3. London: Pitman.

Robson, Barry. 1977. “Biological Macromolecules: Outmoding the Rigid View.” *Nature* 267 (5612): 577–78. https://doi.org/10.1038/267577a0.

Rosen, Gideon. 2020. “Abstract Objects.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2020. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2020/entries/abstract-objects/.

Ross, Lauren N. 2016. “Explanation in Contexts of Causal Complexity.” University of Pittsburgh. http://d-scholarship.pitt.edu/id/eprint/27571.

Ross, Lauren N. 2018. “Causal Selection and the Pathway Concept.” *Philosophy of Science* 85 (4): 551–72. https://doi.org/10.1086/699022.

Sarkar, Sahotra. 1998. *Genetics and Reductionism*. Cambridge: Cambridge University Press.

———. 2008. “Genomics, Proteomics, and Beyond.” In *A Companion to the Philosophy of Biology*, edited by Sahotra Sarkar and Anya Plutynski. Malden, Mass.: Blackwell.

———. 2015. “Nagel on Reduction.” *Studies in History and Philosophy of Science Part A*, Special Section: Formal Epistemology and the Legacy of Logical Empiricism, 53 (October): 43–56. https://doi.org/10.1016/j.shpsa.2015.05.006.

Schulz, Georg E, and R.Heiner Schirmer. 1979. *Principles of Protein Structure*. New York, NY: Springer.

Skelton, N.J., A.G. Palmer, M. Akke, J. Kordel, M. Rance, and W.J. Chazin. 1993. “Practial Aspects of 2-Dimensional Proton-Detected N-15 Spin Relaxation Measurements.” *Journal of Magnetic Resonance Series B* 102 (3): 253–64. https://doi.org/10.1006/jmrb.1993.1095.

Sober, Elliott. 1983. “Equilibrium Explanation.” *Philosophical Studies* 43 (2): 201–10. https://doi.org/10.1007/BF00372383.

———. 1984. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge, Mass.: MIT Press.

Sperry-Taylor, Ashton T. 2019. “Reassessing Equilibrium Explanations: When Are They Causal Explanations?” *Synthese*, October. https://doi.org/10.1007/s11229-019-02423-2.

Sterelny, Kim, and Paul E Griffiths. 1999. *Sex and Death: An Introduction to Philosophy of Biology*. Chicago, Ill.: University of Chicago Press.

Strevens, Michael. 2008. *Depth: An Account of Scientific Explanation*. Cambridge, Mass.: Harvard University Press.

Sturtevant, J.M. 1977. “Heat Capacity and Entropy Changes in Processes Involving Proteins.” *Proceedings of the National Academy of Sciences of the United States of America* 74 (6): 2236–40. https://doi.org/10.1073/pnas.74.6.2236.

Suárez, Mauricio, and Nancy Cartwright. 2008. “Theories: Tools versus Models.” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 39 (1): 62–81. https://doi.org/10.1016/j.shpsb.2007.05.004.

Tanford, Charles, and Jacqueline A Reynolds. 2003. *Nature’s Robots: A History of Proteins*. New York; Oxford: Oxford University Press.

Tompa, Peter. 2011. “Unstructural Biology Coming of Age.” *Current Opinion in Structural Biology* 21 (3): 419–25. https://doi.org/10.1016/j.sbi.2011.03.012.

Tsai, Chung-Jung, Antonio Del Sol, and Ruth Nussinov. 2009. “Protein Allostery, Signal Transmission and Dynamics: A Classification Scheme of Allosteric Mechanisms.” *Molecular Biosystems* 5 (3): 207–16. https://doi.org/10.1039/b819720b.

Tsai, Chung-Jung, and Ruth Nussinov. 2014. “A Unified View of ‘How Allostery Works.’” *PLOS Computational Biology* 10 (2): e1003394. https://doi.org/10.1371/journal.pcbi.1003394.

Umbarger, H. E., and B. Brown. 1958. “Isoleucine and Valine Metabolism in Escherichia Coli. VII. A Negative Feedback Mechanism Controlling Isoleucine Biosynthesis.” *The Journal of Biological Chemistry* 233 (2): 415–20.

Umbarger, H.E. 1956. “Evidence for a Negative-Feedback Mechanism in the Biosynthesis of Isoleucine.” *Science* 123 (3202): 848. https://doi.org/10.1126/science.123.3202.848.

———. 1961. “Feedback Control by Endproduct Inhibition.” *Cold Spring Harbor Symposia on Quantitative Biology* 26 (January): 301–12. https://doi.org/10.1101/SQB.1961.026.01.036.

Waters, C. Kenneth. 1990. “Why the Anti-Reductionist Consensus Won’t Survive: The Case of Classical Mendelian Genetics.” *PSA: Proceedings of the Biennial Meeting of the Phil of Sci. Assoc.* 1990: 125–39. https://doi.org/10.1086/psaprocbienmeetp.1990.1.192698.

Weber, Gregorio. 1975. “Energetics of Ligand Binding to Proteins.” In *Advances in Protein Chemistry*, edited by C. B. Anfinsen, John T. Edsall, and Frederic M. Richards, 29:1–83. Academic Press. https://doi.org/10.1016/S0065-3233(08)60410-6.

Williamson, M. P., T. F. Havel, and K. Wüthrich. 1985. “Solution Conformation of Proteinase Inhibitor IIA from Bull Seminal Plasma by 1H Nuclear Magnetic Resonance and Distance Geometry.” *Journal of Molecular Biology* 182 (2): 295–315. https://doi.org/10.1016/0022-2836(85)90347-x.

Woodward, James F. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

———. 2018. “Some Varieties of Non-Causal Explanation.” In *Explanation beyond Causation: Philosophical Perspectives on Non-Causal Explanations*, edited by Alexander Reutlinger and Juha Saatsi. Oxford: Oxford University Press.

———. 2020. “Levels: What Are They and What Are They Good For?” In *Levels of Analysis in Psychopathology: Cross-Disciplinary Perspectives*, edited by Kenneth S Kendler, Josef Parnas, and Peter Zachar. Cambridge: Cambridge University Press.

Woody, Andrea I. 2004. “More Telltale Signs: What Attention to Representation Reveals about Scientific Explanation.” *Philosophy of Science* 71 (5): 780–93. https://doi.org/10.1086/421416.

Worrall, John. 1989. “Fresnel, Poisson, and the White Spot: The Role of Successful Predictions in the Acceptance of Scientific Theories.” In *The Uses of Experiment: Studies in the Natural Sciences*, edited by David Gooding, Trevor Pinch, and Simon Schaffer, 135–57. Cambridge University Press.

Wright, Peter E, and H. Jane Dyson. 1999. “Intrinsically Unstructured Proteins: Re-Assessing the Protein Structure-Function Paradigm.” *Journal of Molecular Biology* 293 (2): 321–31. https://doi.org/10.1006/jmbi.1999.3110.

Wyman, J. 1948. “Heme Proteins.” *Advances in Protein Chemistry* 4: 407–531. https://doi.org/10.1016/s0065-3233(08)60011-x.

———. 1963. “Allosteric Effects in Hemoglobin.” *Cold Spring Harbor Symposia on Quantitative Biology* 28 (0): 483–89. https://doi.org/10.1101/SQB.1963.028.01.064.

Wyman, J., and D.W. Allen. 1951. “The Problem of the Heme Interactions in Hemoglobin and the Basis of the Bohr Effect.” *Journal of Polymer Science* 7 (5): 499–518. https://doi.org/10.1002/pol.1951.120070506.

Yates, Richard A., and Arthur B. Pardee. 1956a. “Pyrimidine Biosynthesis in Escherichia Coli.” *The Journal of Biological Chemistry* 221 (2): 743–56.

———. 1956b. “Control of Pyrimidine Biosynthesis in Escherichia Coli by a Feed-Back Mechanism.” *Journal of Biological Chemistry* 221 (2): 757–70.

Zwier, Karen R. 2017. “Interventionist Causation in Thermodynamics.” *Philosophy of Science* 84 (5): 1303–15. https://doi.org/10.1086/694168.