Computer Simulations and Explanations in the Nanosciences

A Philosophical Study

Zur Erlangung des akademischen Grades eines
DOKTORS DER PHILOSOPHIE (Dr. phil.)

von der KIT-Fakultät für Geistes- und Sozialwissenschaften
des Karlsruher Instituts für Technologie (KIT)
angenommene

DISSERTATION

von

Julie Schweer

KIT-Dekan: Prof. Dr. Michael Mäs

1. Gutachterin: Prof. Dr. Dr. Rafaela Hillerbrand
2. Gutachter: Prof. Dr. Marcus Elstner

Tag der mündlichen Prüfung: 17.10.2024

Contents

Zusammenfassung

Die Nanowissenschaften sind ein multi- und interdisziplinäres Forschungsfeld, das sich in den vergangenen Jahrzehnten stark entwickelt hat und erheblich durch den Einsatz von Computersimulationen geprägt ist. Die Welt der Nanowissenschaften wirft einige Herausforderungen auf und hält zahlreiche Überraschungen bereit. So zeigt Materie auf der Nanoskala oft ungewöhnliche und interessante Eigenschaften – und in der Regel lassen sich Ansätze zur Beschreibung von Systemen auf Skalen der Alltagswelt nicht so einfach auf den nanoskaligen Bereich übertragen.

Sowohl in Anbetracht ihrer Phänomene als auch in Bezug auf die zu ihrer Untersuchung verwendeten Werkzeuge geben die Nanowissenschaften daher zu Fragen Anlass, die nicht nur von wissenschaftlichem, sondern auch von philosophischem Interesse sind. Insbesondere stellt sich die Frage, was und auf welche Art und Weise eigentlich in den Nanowissenschaften erklärt wird – und welche Rolle Computersimulationen dabei spielen.

Die vorliegende Arbeit nähert sich der Praxis der Nanowissenschaften aus wissenschaftsphilosophischer Perspektive und verfolgt dabei zwei übergeordnete Anliegen. Das primäre Ziel der Arbeit ist es, am Beispiel der Nanowissenschaften zu einer philosophischen Klärung der Rolle von Computersimulationen in Erklärungszusammenhängen beizutragen und dabei auch auf mögliche Vorbehalte hinsichtlich ihrer Eignung für das Ziel wissenschaftlicher Erklärungen zu reagieren. Neben einer Diskussion der Frage, ob und inwiefern die vielfach attestierte epistemische Opazität von Computersimulationen ihrem Einsatz zu Erklärungszwecken entgegensteht, widmet sich die Arbeit den konkreten Beiträgen, die Simulationen zu Erklärungen leisten können.

Das sekundäre Anliegen der Arbeit ist es, eine philosophische Perspektive auf Computersimulationen in den Nanowissenschaften vorzustellen, die ihren Ausgangspunkt in einer nahen Beschäftigung mit tatsächlicher Simulations- und Erklärpraxis findet. Auf diese Weise soll zu einem nuancierten Verständnis der Rolle von Computersimulationen und den in ihnen zum Einsatz kommenden Modellen beigetragen werden und Anknüpfungsmöglichkeiten für aktuelle Diskussionen gegeben werden, welche etwa Herausforderungen beim Erklären komplexer Phänomene oder das Verhältnis zwischen verschiedenen Beschreibungsebenen in Multiskalenkontexten betreffen.

Die philosophischen Studien dieser Arbeit sind dabei in großen Teilen an zwei Fallstudien orientiert, die jeweils unterschiedlichen für nanowissenschaftliche Forschung relevanten Wissenschaftszweigen entstammen; der Molekularbiologie einerseits und der physikalischen Chemie andererseits.

Auch wenn sich in der gegenwärtigen Wissenschaftsphilosophie ein Trend abzeichnet, philosophische Studien nah an tatsächlicher Wissenschaftspraxis zu orientieren, so wirft der philosophische Rekurs auf Fallstudien doch einige Herausforderungen auf, die etwa das Risiko von Übergeneralisierungen und Verzerrungen betreffen. Nach einer Einleitung wird daher im zweiten Kapitel der Arbeit erörtert, auf welche Art und Weise Fallstudien eine berechtigte Rolle in philosophischen Untersuchungen spielen können. Es wird vorgeschlagen, Fallstudien weniger als partikulare Ressourcen für induktive Verallgemeinerungen zu betrachten, sondern sie vielmehr als mögliche Ausgangspunkte zur Abstraktion zu verstehen. Mit diesem Alternativvorschlag verschieben sich allerdings auch die Fragen, zu deren Beantwortung der Rekurs auf Fallstudien dienlich erscheint. Statt der Suche nach einer übergeordneten Theorie wissenschaftlicher Erklärungen rückt etwa die Frage nach den Charakteristiken und Mustern unterschiedlicher Erklärungspraktiken in den Fokus.

Anschließend setzt sich das dritte Kapitel der Arbeit mit der vielfach attestierten Opazität von Computersimulationen auseinander. In philosophischen Diskussionen wurden Bedenken dahingehend geäußert, ob und wie Computersimulationen angesichts ihrer Opazität dazu geeignet sind, wissenschaftliche Erklärungen zu ermöglichen oder Verstehen zu befördern. Anhand eines Fallbeispiels, welches die biologische Funktion bestimmter Kanalproteine an den Membranen biologischer Zellen betrifft, diskutiert das Kapitel, ob die attestierte Opazität von Computersimulationen ihrem Einsatz zu Erklärungszwecken entgegensteht. Es wird aufgezeigt, dass Computersimulationen eine Schlüsselrolle für Erklärungen spielen können; und am Beispiel der Kanalproteine an Zellmembranen wird illustriert, wie sie ein effektives Instrument darstellen können, um mit der Komplexität molekularer Systeme umzugehen.

Im vierten Kapitel der Arbeit werden die verschiedenen Hinsichten, in denen Computersimulationen zu Erklärungen beitragen können, genauer in den Blick genommen. Ausgangspunkt bietet eine Fallstudie aus der physikalischen Chemie, welche das überraschende Vorkommen bestimmter Ionen an der Oberfläche von Wassertropfen in atmosphärischen Aerosolen betrifft. Es werden drei zentrale Hinsichten herausgestellt, in denen Computersimulationen zu wissenschaftlichen Erklärungen beitragen können. Erstens können Computersimulationen dazu dienen, kontrollierte und anderenfalls oft nicht durchführbare Manipulationen an komplexen Systemen vorzunehmen. Dies erlaubt es, gezielt zu überprüfen, welche Faktoren in Bezug auf das jeweils zu erklärende Phänomen einen Unterschied machen. Zweitens können Computersimulationen dabei helfen, systematisch von

(atomistischen) Details zu abstrahieren und epistemischen Zugriff auf diejenigen Variablen zu erlangen, welche für Erklärungen besonders relevant sein können, da sie der Stabilität von Strukturen oder der Erwartbarkeit bestimmter molekularer Prozesse Rechnung tragen. Drittens können Computersimulationen eine Rolle für Erklärungen spielen, indem sie dabei helfen, für das Verhalten komplexer Systeme potentiell relevante Faktoren quantitativ besser zu beschreiben und damit ein detaillierteres Bild von der Gewichtung verschiedener Einflussfaktoren zu erhalten.

Zum Abschluss richtet das fünfte Kapitel der Arbeit ausblickend den Fokus auf gegenwärtige Diskussionen zu Multiskalensimulationen. Vielfach werden in den Nanowissenschaften Simulationen eingesetzt, bei denen Beschreibungen auf verschiedenen Zeit- oder Längenskalen miteinander verbunden werden. In der jüngeren Vergangenheit wurde die Entwicklung und der erfolgreiche Einsatz von Multiskalensimulationen in den Wissenschaften – besonders in den Nanowissenschaften – zum Anlass in der Wissenschaftsphilosophie genommen, um Zweifel an der Angemessenheit traditioneller Ansätze zur Charakterisierung des Verhältnisses verschiedener Beschreibungsebenen zueinander anzumelden. Diese Diskussionen aufgreifend untersucht das Kapitel die Zusammenarbeit verschiedener Theorien und Modelle in Multiskalensimulationen. Auf einen vieldiskutierten Vorschlag bezugnehmend, demzufolge Modelle als teilweise autonome Vermittlungsinstanzen zwischen Theorien und der Welt zu verstehen sind, unterstreicht das Kapitel, inwiefern die Verbindung verschiedener Teilmodelle in Multiskalensimulationen das Ergebnis einer Reihe von Modellierungsschritte und -entscheidungen ist, die mehr als nur Theoriewissen erfordern.

Summary

The nanosciences are a multi- and interdisciplinary field of research that has grown rapidly over the last decades and that is significantly shaped by the use of computer simulations. The nanoscale presents many interesting scientific challenges and holds various surprises. Matter at the nanoscale often shows unusual and unexpected properties, and approaches to describing phenomena at macroscopic scales are often not straightforwardly applicable to phenomena at the nanoscale.

Against this backdrop, the nanosciences raise a number of questions that are not only of scientific but also of philosophical interest. Specifically, there is the question of how explanations are given in the nanosciences – and what role computer simulations play in these explanations.

This dissertation addresses research practices in the nanosciences from a philosophical perspective. There are two overarching research aims. The first aim is to contribute to a philosophical clarification of the role of computer simulations in explanatory contexts and to respond to possible concerns regarding their explanatory power. In particular, I zoom in on the concrete contributions that computer simulations can make to explanations.

The second aim is to offer a philosophical perspective on computer simulations in the nanosciences that draws inspiration from an inspection of actual explanatory practice and simulation practice. In this way, I seek to contribute to a nuanced understanding of the role of computer simulations and models in scientific practice, and offer connecting points for further discussions concerning challenges in explaining complex phenomena or the relationship between different levels of description in multiscale simulation contexts.

The philosophical studies in this dissertation are largely centered around two case studies from distinct branches of science relevant for nanoscientific research: molecular biology and physical chemistry. Even though the philosophy of science has recently experienced a trend of 'practice-orientation', philosophers' use of case studies introduces a number of challenges. These challenges include the risk of overgeneralization and bias. After an introduction, the second chapter of the dissertation therefore discusses how case studies can play a valid role in philosophical inquiries of scientific practice. Rather than regarding case studies primarily as resources for inductive gen-

eralizations, it is suggested to consider them as potential starting points for abstractions. This alternative approach, however, has an effect on the kinds of questions that case studies can help address. Instead of searching for an 'overarching' theory of scientific explanation, the focus shifts towards the characteristics of explanatory practices and the patterns present in them.

The third chapter starts with the often-attested epistemic opacity of computer simulations. In philosophical discussions, concerns have been raised regarding whether and how computer simulations, considering their opacity, are suitable means for obtaining scientific explanations or achieving understanding. Drawing on a case study that concerns the biological function of certain channel proteins in the membranes of biological cells, it is discussed whether the supposed opacity of computer simulations is at odds with their explanatory power. It is proposed that computer simulations can play a key role in explanations, and the example of channel proteins in cell membranes illustrates how they can effectively help scientists handle the complexity of molecular systems.

Building on that, the fourth chapter of the dissertation is concerned with the concrete contributions that computer simulations can make to scientific explanations. The starting point is a case study from physical chemistry, which is about the surprising occurrence of certain ions near the surface of water droplets in atmospheric aerosols. Three key contributions that computer simulations can make to scientific explanations are outlined. First, computer simulations allow scientists to perform controlled manipulations that might otherwise be difficult or even impossible to conduct. In this sense, they enable targeted manipulations of factors which potentially make a difference for the respective phenomenon under consideration. Second, computer simulations can help with a systematic abstraction from (atomistic) details, providing epistemic access to variables that code for the stability of structures or the expectedness of certain molecular processes. Third, computer simulations can play a crucial role in explanations by capturing the effects of particular factors in quantitative terms, thereby contributing to a better assessment of their relative importance for the explanandum.

In the fifth chapter, I zoom in on current philosophical discussions about multiscale simulations. In the nanosciences, simulations are frequently used to examine systems that span various time or length scales. Recently, the development and successful use of multiscale simulations in the sciences – particularly in the nanosciences – have prompted discussions in the philosophy of science, casting doubt on traditional approaches to describing the relationships between different levels of description. Addressing these discussions, I draw on an widely-discussed idea according to which scientific models are partially autonomous mediators between theories and the world. I suggest that the successful combination of different sub-models in multiscale simulation contexts is the result of various modeling steps and decisions and requires more than knowledge of the respectively

involved scientific theories.

Acknowledgements

While being a doctoral student at Karlsruhe Institute of Technology (KIT), I was member of both a philosophical research group, PhilETAS[1], and an interdisciplinary research training group, the RTG "Tailored Scale-Bridging Approaches to Computational Nanoscience"[2]. I am most grateful that I had the chance to pursue my doctoral studies in such an academically diverse environment. Working on my philosophical dissertation while also being part of a scientific research training group was certainly an experiment, and a rewarding and a highly stimulating time.

There are numerous people I would like to thank. First of all, I would like to express my profound gratitude to my supervisor Rafaela Hillerbrand. You struck the right balance between being a guiding academic mentor and giving me the freedom to develop and pursue my own research interests during my time as a doctoral student. I deeply appreciate your feedback, your encouragement and support, and I am thankful that you provided me with all the resources I needed for my studies.

To Marcus Elstner, thank you for your dedicated support, for asking hard questions, and for highly productive exchange. I was always looking forward to discussing my ideas with you and my interests in the philosophy of science owe much to you. Thank you for encouraging me to look closely into scientific practice and to discuss my ideas with people from various academic backgrounds.

To the members of my research training group, thank you for your openness, for lively discussions and insightful exchange. I am particularly grateful for exchange with Peter Gumbsch and Pascal Friederich, whose perspective and input were highly valuable for me.

I met many other great people at KIT whom I would like to thank. First of all, I was most lucky to have Giovanni Frigo as my mentor in the PhilETAS group. Giovanni, thank you for being a great listener, for offering me thoughtful advise and for navigating me through challenging times of my PhD. Also, my thanks to the members of PhilETAS and to various members of the Institute

[1]Which is the research group "Philosophy of Engineering, Technology Assessment, and Science", led by Rafaela Hillerbrand, at the Institute for Technology Assessment and Systems Analysis, Karlsruhe Institute of Technology.

[2]Which is the research training group (RTG) 2450 at Karlsruhe Institute of Technology, funded by the Deutsche Forschungsgemeinschaft (DFG).

of Technology Futures at KIT with whom I had the chance to discuss philosophy. I am especially thankful for productive and fruitful exchange with Michael Poznic, Paul Grünke, Marcel Krüger, Heinrich Blatt, Marianne van Panhuys, and Oliver Bott. Also, my special thanks to Mathias Gutmann for invaluable feedback and inspiring discussions.

During my time as a doctoral student, I had the chance to visit the Department of Philosophy at the University of Pittsburgh as a visiting scholar. I owe my profound gratitude to Robert Batterman for hosting me, for stimulating exchange and for invaluable feedback. In Pittsburgh I also met Sander Beckers, whom I thank for feedback and inspiration, and with whom I spent many hours exploring Pittsburgh by foot, talking about philosophy and everything else.

In 2023, I joined the Section for Translational Medical Ethics at the National Center for Tumor Diseases in Heidelberg as a research associate. I thank the members of the AG Winkler for their warm welcome, and for their understanding during the completion of my dissertation.

Last but not least, my biggest thanks to my family and friends. You know who you are – thank you so much for making these past years so much more enjoyable and for keeping me going. My special thanks to Philipp, Anke and Herbert, to Henning, to Mila and Jonas, and to Christian. I am most grateful for quality time with you, for welcome distraction, and for emotional support. Thanks to my parents for their love and understanding. I am indebted to my mother for her unconditional support and for her encouragement during my journey. Thank you for having my back!

List of included articles

Parts of this dissertation have been published in or submitted to philosophical journals.

Schweer, J.: Scientific Explanations in the Wild: On Case Studies in the Philosophy of Science in Practice, *under peer review*. (**Chapter 2**)

Schweer, J., Elstner, M.: Dealing with Molecular Complexity. Atomistic Computer Simulations and Scientific Explanation, published in: *Perspectives on Science*, 31 (5): 594–626, ©2023 by the Massachusetts Institute of Technology. `https://doi.org/10.1162/posc_a_00594`.
(**Chapter 3**)

Schweer, J., Elstner, M.: Computer Simulations, Scientific Explanation and Difference-Making: Lessons from Simulations of Ions in Aerosols, *under peer review*. (**Chapter 4**)

Introduction

Contents

Summary

Over the last decades, the increased use of computer simulations has greatly shaped scientific research practices in various fields and sparked lively discussions in the philosophy of science. One scientific field that heavily relies on computer simulations is the nanosciences. Not only is the study of nanoscale systems relevant for many areas of life and science, but systems at

the nanoscale often exhibit surprising or unexpected behavior. In this dissertation, I examine research practices in the nanoscience from a philosophical perspective and suggest that specifically the role of computer simulations in scientific explanations deserves philosophical attention. After providing a brief introduction to the nanosciences and their use of computer simulations, I offer an overview of various recent topics in the philosophy of computer simulation. Building on this, I present the two main research aims of this dissertation. Firstly, the aim is to contribute to a better understanding of how computer simulations can figure scientific explanations. Secondly, the aim is to offer a philosophical perspective on computer simulations in the nanosciences that draws inspiration from actual explanatory practice. In pursuing these aims, the dissertation seeks to contribute to a more nuanced understanding of the explanatory role of computer simulations in scientific research, while also offering a fresh outlook on philosophical discussions related to explanations of complex systems, models, and multiscale simulations.

1 Entering the Nanoscale

Many processes that have a profound impact on our life occur at very small time or length scales. Take the case of proteins and their manifold roles in maintaining our bodily functioning: they synthesize hormones, carry nutrients or regulate our immune system. Compared to the organisms whose functioning they affect, proteins are extremely small. A hemoglobin molecule, for example, has the size of only about five nanometers – which is way too tiny for it to be visible through a standard optical microscope, let alone the naked eye. One nanometer amounts to a billionth of a meter. A single small atom has the size of approximately one tenth of a nanometer. As a comparison, an average human hair has a diameter of about 60 000 nanometers (Bayda et al. 2019).

Scientists and scholars have been fascinated with the world of the extremely small ever since. As early as four hundred years BC, early Greek natural philosophers including Leucippus and Democritus suggested that the world is made up of tiny building blocks that bring about the macroscopic world (Berryman 2023).

As concerns modern science, Dalton's atomic theory in the early 19th century set important groundwork for studies about the interactions between atoms and molecules. With the birth of quantum mechanics in the 20th century, a precise mathematical framework enabling a detailed understanding of atomic structure and a quantitative description of molecules came into reach (Kleppner and Jackiw 2000).

In his visionary 1959 lecture *There's Plenty of Room at The Bottom*, physicist Richard Feynman famously speculated how systems at the scale of atoms and molecules could be inspected and

manipulated. The Nobel prize winning invention of the scanning tunneling microscope (STM) in 1981 was an important factor for this vision to become true (Lindsay 2010, 2). The STM generates atomic resolution images using a very thin metal probe tip that 'scans' the surface of the respective specimen.[1] Being affordable and relatively easy to build, the availability of the STM led to a rapidly growing scientific interest in the study of the atomic-scale properties of materials (Lindsay 2010, 2). Only a few years after its invention, the first images of biological molecules were reported (Dahn et al. 1988; Lindsay 2010, 3).[2]

However, despite advances in the study or imaging of individual atoms and molecules, many particularly interesting structures and processes can be found not at the atomic scale, but at a scale between 1 – 100 nm, which is the nanoscale. Systems at the nanoscale typically consist of several hundreds to tens of thousands of atoms (Lindsay 2010, 1). At the nanoscale, matter often shows remarkable physical, chemical, or biological properties, e.g. as concerns chemical reactivity or electrical conductivity. The scientific field dedicated to studying the behavior and structure of matter at the nanoscale is nowadays known as the nanosciences.

Having experienced fast advances over the last decades, research on nanoscale phenomena and processes today contributes to a wide array of scientific domains, including physics, materials science, chemistry, the life sciences and engineering. The targets of nanoscale research can be many different things, be it a better understanding of biological processes such as the working of enzymes in our bodies, the study of material failure, or the development of technologies such as carbon nanotubes which, with their unique properties, hold promises for applications in many areas, ranging from energy storage to biomedicine, just to name a few (see e.g. Yang et al. 2007; Kumar et al. 2018).

1.1 Scale Matters

> *"Atoms on a small scale behave like nothing on a large scale, for they satisfy the laws of quantum mechanics. So, as we go down and fiddle around with the atoms down there, we are working with different laws, and we can expect to do different things."*
>
> – Richard Feynman (1960, 36)

Entering the nanoscale means more than just 'getting smaller'. Rather, size has an enormous ef-

[1] The height of the probe is constantly aligned with the specimen such that it maintains constant electron tunneling between the tip and the surface (Stroscio and Eigler 1991). This, in turn, allows it to obtain an image of the 'surface topology' of the inspected material.

[2] Besides imaging of atoms, the STM even made it possible to perform *manipulations* at an atomic scale by displacing single atoms and molecules into novel arrangements, thereby constructing new artificial material compositions (Marcovich and Shinn 2014, 1.1.1.). An example for how atomic-scale properties could be controlled by means of STM is the spelling of the company name IBM with 35 individual xenon atoms performed by IBM employée Donald Eigler in 1989 (Marcovich and Shinn 2014, 1.1.1.).

fect on the properties and behavior of matter. Much of what our theories and everyday experiences teach us about matter on larger scales figures very differently or fails to be applicable for systems at the nanoscale.

As an example, consider the case of sodium chloride – which is common salt. At macroscopic scales, sodium chloride is known as a rigid material. At the nanoscale, however, sodium chloride has been found to show superplasticity under certain conditions, meaning that the material can be stretched to more than twice its original size (Moore et al. 2009, 2295). The case of salt gives a good example of how material properties at the nanoscale can differ strikingly from those at higher scales:

> "[...] [T]he idea that common salt, which crushes brittlely between one's fingers, can be superplastic is a striking and unexpected example of how material properties can change when viewed at the nanoscale." (Moore et al. 2009, 2298)

Another example is gold. Studying the behavior of materials at the nanoscale, Uzi Landman states that "small is different in an essential way, with the physical and chemical properties becoming emergent in nature, i.e., they can no longer be deduced from those known for larger sizes" (Landman 2005, 6672). Having performed computer simulations of gold at the nanoscale, Landman et al. noted that gold atoms would assemble in a certain way so as to form small wires when being brought in the proximity of a nickel tip (Landman et al. 1996). Such a behavior is different from what would be expected based on knowledge about gold on larger scales. Hence, the finding that gold can form small wires at the nanoscale was certainly a surprise (Lenhard 2006, 606).

There are various reasons for why things are different on very small scales. First of all, the surface-to-volume ratio of nano-sized materials is significantly larger than in bulk materials (Cortie 2004, 14). Quantum effects can become crucial for materials at the nanoscale, impacting their optical or magnetic properties, among other things. Also, fluctuations often begin to play a dominant role (Lindsay 2010, 9), and factors such as static electricity or gravity figure quite differently in materials of nanoscopic size.

As Lindsay (2010, 1) puts it, the nanoscale is where quantum mechanics and the physics and chemistry of complex systems meet:

> "In nanostructures, we have, layered on top of quantum mechanics, the statistical behavior of a large collection of interacting atoms. From this mixture of quantum behavior and statistical complexity, many phenomena emerge." (Lindsay 2010, 1)

To illustrate how scale matters, consider two samples of the same material, one of nanoscopic and one of macroscopic size. Given the relatively larger number of atoms at the surface compared

to the inside, surface reactions will be significantly amplified in the nano-sized material (Cortie 2004, 14). Moreover, given that the number of atomic bonds that 'hold together' the nanoscopic material sample is much smaller than in a macroscopic sample of the same material, it will be more sensitive to factors such as its thermal environment – and whereas the macroscopic sample may be quite stable, the effects of random thermal agitation may even cause the nanoscopic sample to be prone to falling apart (Lindsay 2010, 9).

1.2 Computer Simulations: Peeking Through the Computational Microscope – And Zooming Out Again

Examining nanoscale phenomena requires sophisticated scientific methods and instruments. Not only are the systems addressed in the nanosciences extremely small, they also usually comprise a high number of components that interact in complicated ways. Analytical methods cannot be applied if systems get too complex, and techniques and structural tools such as X-ray crystallography can reach their limits when the aim is to understand the dynamic aspects relevant for many molecular systems (see e.g. Srivastava et al. 2018).

Having profited from great advances in method development and rapid gains in computing power, computer simulations have emerged as particularly important means for addressing the properties and behavior of matter at the nanoscale.

The Birth Hour of Monte Carlo Simulations

The scientific use of computer simulations traces back to the late 1940s and the military context of the Second World War. Being concerned with the development of nuclear weapons, physicists affiliated with the Manhattan Project studied the diffusion of neutrons but found themselves unable to solve the problem using traditional mathematical methods (Galison 1996, 120). This was the birth hour of Monte Carlo simulations, a family of stochastic methods based on the execution of random operations. Monte Carlo simulations are commonly used for tackling molecular problems and they still play a highly relevant role in the zoo of methods not only in the nanosciences, but in a plethora of branches ranging from statistical physics over biology to medicine, just to name a few.

With the development of the microprocessor and the availability of personal computers in the 1980s, computer simulations became increasingly popular, finding manifold applications throughout the sciences, and leading to the formation of whole new computational research departments (Marcovich and Shinn 2014, 1.1.2.).

Computer simulations have today become ubiquitous in almost all areas in the sciences, ranging from the climate sciences or biology to the social sciences and economics. In recent years, they are increasingly used together with machine learning methods which can help prepare, analyze or enhance the simulations (see e.g. Wang et al. 2020).

Molecular Dynamics Simulations

Besides the aforementioned Monte Carlo simulations, an especially relevant class of simulations in the nanosciences are molecular dynamics simulations. Rather than being based on stochastic methods, conventional molecular dynamics simulations proceed by numerically solving Newton's equations of motion for the atoms of the respectively investigated system, thereby tracing the evolution of the respective system in time.

What makes molecular dynamics simulations particularly useful for research in the nanosciences is that they are, for many problems, considered to meet a good balance between computational efficiency and accuracy. Even though the theory to give a detailed account of atomic interactions is quantum mechanics, the complexity of molecular problems often prohibits a quantum mechanical treatment. Therefore, conventional molecular dynamics simulations are based on inter-atomic potentials, or force fields, which are approximative models of the interactions between the atoms. More specifically, interatomic potentials approximate the potential energy of the system based on a description of interactions between the atoms as a function of their positions.

The basic idea of conventional molecular dynamics simulations is the following. Take an atomic starting configuration (obtained, e.g., on the basis of experimental or computational techniques) and divide time into suitable discrete steps, usually a few femtoseconds. Then, using suitable interatomic potentials, calculate the forces acting on the atoms as a derivative of the energy with regard to changes in the respective atoms' positions. In this way, obtain trajectories which show how the atoms of the investigated system move over time.

From Micro to Macro

Despite starting from descriptions of atoms and their interactions, the overarching research questions tackled by molecular dynamics simulations typically do not concern particular trajectories, but the ways in which the microscopic details of a system give rise to behaviors and properties on a higher scale. In other words, when using molecular dynamics simulations, we are typically not interested in the detailed atomic trajectories generated in single simulation runs, but we would like to process such information in order to learn about higher-scale, e.g. thermodynamic, properties

of a system.

The relevant theoretical ingredient for bridging from the atomic to the macroscopic world is statistical mechanics. At an atomic resolution, the state of a system is described in terms of the positions and momenta of all atoms. Statistical mechanics uses such microstate information about a system in order to infer higher-scale properties – such as temperature, pressure, or free energy. Very briefly put, this is facilitated by drawing on *ensembles* of microstates and treating them statistically.[3] With computer simulation methods such as molecular dynamics or Monte Carlo methods, microstate ensembles can be generated, and then, by means of statistical mechanical methods, processed to enable insight, for example, into the free energy of the investigated system.

The free energy is a particularly relevant thermodynamic quantity because it reflects (or codes for) the stability or likelihood of certain molecular configurations.[4] If we are for, example, studying protein folding dynamics, we may be interested in tracing how the free energy of the respectively investigated system changes along the course of the folding reaction. Such information can be obtained based on microstate ensembles provided by molecular dynamics simulations or Monte Carlo simulations.

In a nutshell, to understand the relevance of molecular dynamics simulations in the nanosciences, it is important to acknowledge what kinds of insights such simulations can provide. By tracing the dynamic evolution of molecular systems and by acting as instruments for 'bringing statistical mechanics to work', molecular dynamics simulations can be used to study a plethora of interesting processes and phenomena including protein folding, the dynamics and function of biomolecules, or the behavior of materials on a molecular scale.

1.3 Multiscale Simulations

Even though molecular dynamics simulations alone are already instrumental for bridging between a description of atomic movement and the world of molecular structures or molecular behavior, there are contexts in which the aim is to address systems which involve processes or properties occurring at an even broader range of spatial or temporal scales. Again, this is perhaps especially the case in the nanosciences, because for many practical applications – think of materials research or drug design – the peculiar behavior of matter at the nanoscale is assumed to have relevant consequences at higher scales.

As an example of a multiscale phenomenon from materials science, consider the emergence of cracks in solid materials such as steel or silicon. Solid materials can be described on various length

[3] In chapter 3 and chapter 4, a more detailed picture of how this 'bridging' works will be provided.

[4] Dealing with the examples of aquaporin selectivity and ions in aerosols, this will be further elaborated on in chapters 3 and 4.

scales: The atomic scale of less than 1 nm is where chemical bonds figure (Nosonovsky and Bhushan 2007, 163). At a scale from 1 nm to 1 μm, structures such as dislocations (i.e., irregularities in the molecular lattice of the material) or small material defects occur. At macroscopic continuum scale, solid materials are treated as continua. Here, factors such as the relationship between stress and strain are commonly examined.

The Example of Cracks in Silicon

The failure of materials is typically considered a 'multiscale phenomenon', meaning that factors figuring at various scales are assumed to matter for how materials fail.[5] Think of the dynamics of tiny cracks in silicon – a semiconducting material with countless of uses in everyday life and industry, ranging from solar cells to bricks. On a macroscopic scale, the material properties of silicon can be studied by means of continuum mechanics, i.e., by treating the investigated silicon sample as a continuum rather than a compositum of individual atoms (see table 1.1). Yet, neglecting the atomic structure of the material comes with certain limitations, e.g. regarding the extent to which losses in the local cohesion of the material can be taken account of (Abraham et al. 1998, 783). At an atomic resolution, crack dynamics in silicon can be modeled in terms of empirical forces acting between classical atoms. However, in order to study the breaking of molecular bonds in more detail in the region of the emerging crack tip, an even finer-grained view using quantum mechanical principles can be useful.

Length scale	**Theory**	**Objects investigated**
0.1-1 nm	Quantum mechanics	Atoms and molecules
1 nm - 1 μm	Dislocation dynamics	Dislocations, material defects
$> 1\mu$m	Continuum mechanics	Continuum bodies

Table 1.1: Investigation of the properties of solid materials at multiple length scales. Table based on Nosonovsky and Bhushan 2007.

In the case of silicon cracks – just as in numerous other scientific contexts – it is interesting to examine how processes on different scales affect each other: How do small molecular defects eventually lead to larger-scale material failure? How is the development of small defects in a certain material sensitive to changes in its temperature? In order to study the relationship between processes on different scales, bridges between these scales need to be established.

[5] Another example for a multiscale phenomenon would be friction between two solid bodies, i.e., the force resisting the relative motion of pushing the two bodies on each other. At a macroscopic scale, the friction coefficient is one of the central parameters for describing the interaction between two bodies (Bucher et al. 2006). At the same time, processes occurring at nanoscopic length sales, such as plastic deformation and fracture have been suggested to play a relevant role for friction (ibid.).

Simulating Crack Propagation Across Scales

Computer simulations have become key tools for studying processes occurring and interacting at multiple scales. As concerns crack propagation in silicon, a very illustrative example of a multi-scale simulation linking various spatial domains of the material at different length scales is that by Broughton, Abraham et al. (1998; 1999). Basically, the silicon crack simulation connects three sub-models each describing a certain spatial region of the material (see table 1.2).

In the spatial region where the tip of the crack emerges, an approximative quantum-mechanical (QM) approach – called tight-binding (TB) method – was used to describe the electronic structure of the material. Such a model captures the kinetics of the system in a very detailed fashion by explicitly describing the breaking of bonds, but it is computationally expensive and can hence only be employed for a small spatial region of the material (Broughton et al. 1999, 2392). Therefore, the scientists combined it with a molecular dynamics (MD) model based on classical forces between atoms that captures the neighboring region of the crack (Winsberg 2006, 587). This modeling choice is motivated as follows:

> "This MD region is required because it correctly captures the necessary thermal fluctuations and the pressure waves emitted by the bond rupturing and/or defective region. The MD region may also contain those areas of the system that are defective but for which the primary dynamics are no longer important." (Broughton et al. 1999, 2392)

In addition to the TB model of the crack tip and the MD model of the surrounding region, the material region farer from the tip was modeled with a finite element (FE) approach, i.e., by treating the material as a continuum that is approximated and discretized in the form of a mesh consisting of various subparts called finite elements (Abraham et al. 1998, 784).

Spatial domain	**Model (and theory)**	**Investigated structures**
Crack tip	Tight-binding (Quantum mechanics)	Breaking of molecular bonds
Neighboring regions of the crack	Molecular dynamics (Classical mechanics)	Effects of the emerging crack on the domain surrounding it
Regions spatially distant from the crack	Finite element (Continuum mechanics)	Displacement and strain

Table 1.2: The different scales and the according models used to study various structures relevant for crack propagation in silicon. Table based on Abraham et al. (1998) and Bursten (2018).

Finally, to study the propagation of the crack through the different spatial domains of the silicon block, Broughton, Abraham et al. employed sophisticated algorithms specifying how information is passed from one sub-model to the other. Considering that the sub-models at the different spatial

regions are based on different theoretical frameworks and make use of various approximating ideas, crafting such algorithms is a challenging and non-trivial task.

2 A Philosophical Perspective

> *"If the philosophy of computational science is a footnote to Plato, it's an exceedingly long one."*
>
> – Paul Humphreys (2004: 56, footnote)

Besides challenges such as the aforementioned development of suitable algorithms, the increasing use of computer simulations in the sciences, especially multiscale simulations, has sparked a host of questions that are philosophical in nature. In the following sections, I start with a conceptual clarification of the term 'computer simulation' and zoom in on the role of models in the context of computer simulations. Building on that, some of the topics pertinent in the philosophy of computer simulation will be touched upon, and I discuss how they resonate with research practice in the nanosciences.

2.1 What are Computer Simulations?

So far, this introduction has drawn on a rather implicit and vague understanding of the term 'computer simulation', using mainly the history and exemplary cases of computer simulations in the nanosciences as a starting point. However, when being concerned with epistemic aspects of computer simulations, it is important to clarify in more detail as to what this work is concerned with when dealing with *computer simulations*.

First of all, it is useful to distinguish between different types of computer simulations. One commonly-drawn distinction is that between 'equation-based simulations' and 'agent-based simulations'. Whereas the former rely on equations which describe the respectively investigated system's overall dynamics, the latter use rules that capture the dynamics and interactions of a system's constituent parts *locally* (e.g. in the form of behavioral rules) and, in this way, *generate* the overall dynamics of the system (Grüne-Yanoff and Weirich 2010, 32).[6] As a third type of computer simulations, there are stochastic approaches, such as the earlier mentioned Monte Carlo methods, which use random operations to explore the characteristics of a target (see Winsberg 2019).

Besides distinguishing between various types of computer simulations, one can ask more broadly as to what it means to use a computer simulation in science: What are we doing when we are using

[6] Agent-based simulations are used, e.g, in the social sciences, philosophy of science or social epistemology to study topics such as polarization dynamics or scientific collaborations, see Šešelja (2023).

a computer simulation? One idea is to characterize computer simulations as means for *imitation*. According to Stephan Hartmann, for example,

> "[a] simulation imitates one process by another process. [...] If the simulation is run on a computer, it is called a computer simulation." (Hartmann 1996, 82)

A drawback of Hartmann's characterization is that it does not cover all types of computer simulations equally well. In the case of the earlier mentioned molecular dynamics simulations, it may seem acceptable to say that the simulation *imitates* a molecular process by numerically integrating Newton's equation of motion for all atoms, thereby tracing the motion of the atoms in the system.[7] The situation is different in the case of Monte Carlo simulations. Whereas Monte Carlo simulations can similarly be used to study e.g. energetic and structural properties of molecular systems, they rely on stochastic operations rather than an integration of equations of motion. Against this backdrop, it is not entirely clear as to what the *imitating process* would be in the case of Monte Carlo simulations (see Winsberg 2019).[8]

A different working definition has been offered by Paul Humphreys. According to Humphreys, computer simulations can be regarded as computationally implemented methods for "exploring the properties of mathematical models where analytic methods are unavailable" (Humphreys 1991, 501). Humphreys' definition points to the fact that many scientific problems are addressed with the help of (partial) differential equations. Since such problems are often hard to come by with analytical methods, computer simulations have become popular means for an approximative treatment using numerical methods. However, as it has been pointed out in the literature (Hartmann 1996; Hillerbrand 2013; Winsberg 2019; cf. also Schweer et al. forth), the proposed definition raises a few problems: on the one hand, computer simulations can also be useful in situations in which analytical solutions may be available; on the other hand, not every method used to explore the properties of mathematical models is a computer simulation.[9]

According to Eric Winsberg, "no single definition of computer simulation is appropriate" (Winsberg 2019). Rather, Winsberg distinguishes between a narrow and a broad sense of the term 'computer simulation'. According to him, 'computer simulation' in the narrow sense refers to the execution of a program that takes as input a set of numerically-described rules and a description of a system's state at a certain point in time (Winsberg 2019). It then approximates the behavior of said

[7] However, the metaphor of imitation has limitations even for molecular dynamics simulations because it is frequently taken for granted from the beginning on that individual trajectories do not have a straightforward physical interpretation in terms of a natural process that corresponds to them (see chapter 3, section 4.2.).

[8] Grüne-Yanoff and Weirich (2010) explicitly consider Monte Carlo simulations to be distinct from what they call *imitating* simulations (Grüne-Yanoff and Weirich 2010, 30).

[9] Humphreys himself later stated that this early definition was too broad and too narrow at the same time (Humphreys 2004, 108). In his 2004 work, he considers Hartmann's definition mentioned afore as a suitable starting point (see main body of the text).

system by algorithmically calculating its state at subsequent discrete time steps (Winsberg 2018, 39f., Winsberg 2019). Differently put, 'computer simulation' in the narrow sense amounts to the run of a simulation program leading from a certain input to an output.

Complementing this narrow definition, Winsberg suggests regarding 'computer simulation' in a *broad sense* as a 'comprehensive method' for studying the properties or behavior of systems (Winsberg 2019). As he puts it, 'computer simulation' broadly understood amounts to an entire process, ranging from model choice over implementation to running the program and visualizing the output (Winsberg 2019).[10]

I agree with Winsberg that – given the variety of computer simulation types and their manifold uses in the sciences – it is difficult to give a unifying definition of the term 'computer simulation'. However, rather than characterizing computer simulation in a broad sense as a 'comprehensive method',[11] I suggest adopting slightly different approach that sets more emphasis on the epistemic *functions* of computer simulations in scientific practice. Following Alvarado (2022), I propose that it is instructive to understand and approach computer simulations as *scientific instruments*.

Characterizing computer simulations as scientific instruments means, according to Alvarado, recognizing them as technical artifacts, and as functional instantiations of certain designs and processes (Alvarado 2022, 1186).

On the one hand, instruments require careful design and construction. Addressing computer simulations as instruments means being attentive to the processes and choices involved in their crafting, use, and assessment. On the other hand, instruments can be used for different purposes in various scientific settings. Approaching computer simulations as instruments means recognizing that they can serve different epistemic purposes in scientific practice, be it measuring (Morrison 2009), controlling (Lenhard 2006) or – as will be the main concern of this work – explaining.

Instruments are often considered as means to amplify or extend our human abilities – and this certainly resonates with the case of computer simulations: they are frequently used to examine processes occurring at scales too small for the naked eye or too large to be tackled in the laboratory, and they can greatly expand our inferential abilities. Discussing how computational methods have shaped and changed science, Humphreys introduces the concept of 'epistemic enhancers' (Humphreys 2004). As Alvarado argues, computer simulations may be regarded as epistemic enhancers in a 'hybrid' sense because using them can involve both *producing* and *intervening on* the object of investigation (Alvarado 2022, 1202f.). Not only may computer simulations expand our

[10] Another common expression capturing 'computer simulations in a broad sense' would be the term 'computer simulation *studies*' (Winsberg 2019).

[11] Looking at the use of the term 'simulation method' in scientific practice, it seems useful to acknowledge that there exist a number of *distinct* computer simulation methods, Monte Carlo methods and molecular dynamics methods being two groups of such methods.

inferential abilities, they also shape *how* something becomes available as an object of investigation in the first place.[12] This has essentially to do with computer simulations' reliance on models, as I will outline in the following.

2.2 Simulations and Models

The idea that computer simulations can be instruments for measurement, prediction, control, or other purposes requires elaboration. When using a computer simulation, we do not intervene directly on the target we are interested in.[13] Rather, inferences from computer simulations rely on *conceptual* or *theoretical*[14] models that capture various relevant characteristics of the target in a mathematical form, typically in the shape of sets of (partial) differential equations.[15]

In the view suggested above, according to which computer simulations can fruitfully be approached as instruments, computer simulations are based on theoretical models, but not identical to them. As Alvarado puts it:

> "While computer simulations may be the product of, or contain within them the specifications of a conceptual model, computer simulations are something other than the model itself: they are the implementation (through hardware architecture and software specifications) of said models." (Alvarado 2022, 1188-89)

I suggest that using a computer simulation is, in a sense, a way of putting theoretical models to work via further modeling steps and implementation: They are instruments that – in a way to be specified – draw on theoretical models and can be utilized for various epistemic purposes.[16]

The construction and scientific employment of theoretical models is certainly a non-trivial task: Although typically being 'theory-inspired', it is acknowledged in contemporary philosophical debates[17] that the theoretical models upon which computer simulations are based should not be seen

[12] While this twofold character of instruments is not specific for computer simulations (cf. e.g. Rheinberger 2006), it underscores that we should look carefully at their relationship to the respective phenomenon of interest.

[13] It is important to note that this is, however, not specific for computer simulations but applies to laboratory experiments, too. Think, e.g., of the San Francisco Bay Model used to study the potential effects of dam building activities, as discussed by Weisberg (2012).

[14] In the philosophical literature, some use the term 'conceptual model', while others speak of 'theoretical models'. I will treat these terms interchangeably, but mostly stick to the term 'theoretical model'.

[15] There are, however, also the earlier mentioned agent-based simulations whose respective theoretical models express rules describing the local interactions between the 'agents' in the system.

[16] Even though I am setting a slightly different emphasis, I consider this approach to resonate with Morgan and Morrison who more broadly consider *models as instruments* for various purposes, including the application of theories, experimental manipulations or measurement (Morrison and Morgan 1999, 20ff.). I here mainly want to suggest a more fine-grained grammatical distinction between computer simulations and their models: using a computer simulation means working with models. Computer simulations are vehicles for using models for various possible purposes, i.e., they act as as instruments in scientific practice.

[17] Current discussions on the role of models and their relationship to theories owe much to the influential work of Margaret Morrison and Mary Morgan. In the traditional syntactic view of theories, the latter were regarded as uninterpreted axiomatized systems and models as sets of statements used to interpret them (Morrison 2007, 198). In contrast, the semantic tradition essentially saw theories as families of models (Morrison 2007, 198). Going beyond both traditions, Morrison and

as straightforward 'derivations' from our theories (Winsberg 1999, Morrison 2009). Besides the need to carefully decide as to which features of the world shall be reflected in a theoretical model (and how), further steps are required which allow for mathematical tractability. Furthermore, before being 'put to work' in the context of a simulation, models need to be adequately discretized and algorithmically implemented.

Recognizing the various steps and activities involved in the construction and evaluation of theoretical models in the context of computer simulations, it is useful to draw an even finer-grained conceptual distinction between 'theoretical models' or 'conceptual models' on the one hand, and 'computational models' or 'numerical models'[18] on the other hand (e.g. Morrison 2015; Boge 2019). Just as the depiction and specification of theoretical models are not a matter of straightforward derivation from theory, the choice of computational models is not dictated by the respective theoretical model. Rather, finding or developing appropriate numerical algorithms is an important research topic in the respective scientific communities, and many different factors – be it empirical considerations or computational constraints – can play a role in how such models are specified.[19]

In a sense, the models upon which computer simulations rely have a 'life on their own' (Morrison and Morgan 1999, 18): They are informed and constrained by scientific theories, but not determined by them. Rather, various considerations playing a role at several stages in modeling processes can affect the form and scope of these models. Following Morrison and Morgan, we can consider models as autonomous 'mediators': On the one hand, they can be means for bringing theoretical resources in a mathematically tractable and computationally implemented form. On the other hand, they are not straightforwardly derived from these theories, but partly independent from them.

2.3 On the Epistemology of Computer Simulation

In the philosophical literature, special attention has been devoted to the epistemic status of computer simulations as compared to experiments and theory. As has been pointed out by Rohrlich (1990), computer simulations seem to be located somewhere *in between* theory and experimenta-

Morgan prominently suggested to consider models as autonomous agents that are partly independent from both theories and the world (Morgan and Morrison 1999), see main body of the text.

[18] Sometimes, the terms 'numerical model' and 'computational model' are not used interchangeably. Rather, an additional conceptual distinction is drawn to emphasize that numerical models still need to be brought in an algorithmically implemented form – and that this process is, again, not a matter of straightforward derivation from theoretical models but relies on making various contingent modeling decisions (Boge 2019). 'Computational models' or 'simulation models' are, according to this distinction, numerical models implemented on a computer in a certain programming language. Adhering to how the terms 'computational model' or 'simulation model' are used in the context of the case studies investigated later, I will in this work stick to the less fine-grained picture and distinguish conceptually between theoretical models and computational models, with the latter being another name for the numerical models in a simulation context. When applicable, the implemented form of numerical models will be referred to as a numerically implemented model.

[19] See e.g. chapter 4, subsection 2.3 of this work.

tion, methodologically speaking. Whereas some have located computer simulations in the methodological proximity of experiments (e.g. Parker 2009; Morrison 2009; Massimi and Bhimji 2015; Boge 2019), others have stressed their inferential character and argued that they can be regarded as (very sophisticated) arguments or inferences (Beisbart 2012; Beisbart and Norton 2012).

Furthermore, discussions have emerged on whether the use of computer simulations in the sciences introduces novel epistemic challenges (Frigg and Reiss 2009; Grüne-Yanoff and Weirich 2010). Besides the question of whether and it what sense computer simulations are epistemically 'opaque' (e.g. Humphreys 2004; Lenhard 2006), it has been asked if they require a shift towards a new 'non-anthropocentric epistemology' (Humphreys 2009).[20]

Simulations and Arguments

Being interested in how computer simulations can be used to generate new knowledge, Beisbart argues that they can be reconstructed as (deductive) arguments with the simulation input and algorithmic rules forming the premises and the simulation output giving the conclusion[21] (Beisbart 2012, 407f.). According to Beisbart, computer simulations contribute to the generation of new knowledge in the same way as arguments do: besides being *reconstructible* as arguments, they have – as Beisbart puts it – the same *epistemic power* as arguments (Beisbart 2012, 403).

Whereas Beisbart convincingly underscores the inferential character of computer simulations[22], it is questionable if the argument view adequately captures how computer simulations 'work', epistemically speaking. As Boge (2019) argues, computer simulations may respect the *logic* of arguments, but they differ significantly in their pragmatics and epistemology. Furthermore, it is often not the output of one particular simulation run that is of scientific interest. Rather, learning from computer simulations often means engaging in inductive reasoning practices and, for example, interpreting and comparing the outputs of *various* simulation runs (cf. Grünke and Schweer 2023).

An Epistemic Privilege?

Of those who have examined the relationship between computer simulations and experiments, some have asked if traditional experiments are in some way better suited for the task of learning

[20] Based on the idea that opacity limits human ability to understand new computational methods and their results, Humphreys has controversially stated that "[...] an exclusively anthropocentric epistemology is no longer appropriate because there now exist superior, non-human authorities" (Humphreys 2009, 617). Thus, we are – according to Humphreys – confronted with the so-called problem of "anthropocentric predicament", that is "[...] how we, as humans, can understand and evaluate computationally based scientific methods that transcend our own abilities" (ibid.).

[21] Beisbart emphasizes that while the computer takes numbers as inputs, these numbers have an interpretation in the context in which the computer simulation is used by scientists, hence they can reflect premises and conclusions: "What the computer program takes as input are of course mere numbers. But for the working scientist, these numbers have empirical meaning; they encode the values of physical characteristics of the system in certain units" (Beisbart 2012, 407).

[22] Cf. also Winsberg (1999) who states that computer simulations "involve a complex chain of inferences [...]".

something about the world. Concretely, there have been discussions on whether experiments are "epistemically privileged" over computer simulations (see Grüne-Yanoff and Weirich 2010, 27). According to Morgan (2005), for example, experiments allow stronger inferences back to the world because they "share the same stuff" as their target (Morgan 2005, 323). While experiments can *confound*, simulations can, in Morgan's view, at best *surprise* because the resources that go into the computational model are known beforehand by the modeler (Morgan 2005, 324).[23]

At first glance, the idea that laboratory experiments involving inventory from the 'real world' can support stronger epistemic claims than computer simulations based on abstract mathematical models may seem appealing. However, a problem with the idea of material sameness or "ontological equivalence" (Morgan 2005, 326) between the object and target of experimental investigation is that it does not fit for all contexts of experimentation equally well.[24] Firstly, there exist many experiments in which the object of experimental investigation is qualitatively very different from the actual target of the study, think of laboratory interventions on mice to study drug efficacy in humans. Secondly, experiments, too, can rely on highly sophisticated equipment and theoretical models that mediate between the object and target of investigation, e.g. in studies in particle physics (Morrison 2009, 43). In such cases, it is not entirely clear why inferences based on experiments should be considered more 'direct' or stronger than inferences based on computer simulations.

More generally, given that there exist various distinct types of both computer simulations and experiments, and considering that they are used for a wide array of scientific problems and for different aims, it seems that the question of whether computer simulations or traditional experiments are 'privileged' means to learn about the world depends on the respective research context and cannot be answered *in principle* (see Hillerbrand 2013; Parke 2014; Grünke 2023).

Epistemic Opacity

The chain of inferences connecting the input with the output of a computer simulation is usually highly complicated, to the extent that no human being could possibly inspect all the computational steps of the simulation process. In light of this, it has been suggested that computer simulations – or, specifically, the inferential processes they involve – are *epistemically opaque*. In Humphreys' famous wording: "In many computer simulations, the dynamic relationship between the initial and final states of the core simulation is epistemically opaque because most steps in the process are not open to direct inspection and verification" (Humphreys 2004, 147-148). According to

[23] Morgan is concerned with economics and discusses the case of behavior patterns that could be studied in an experimental setting or by means of mathematical (simulation) models. According to Morgan, the choice and specification of the mathematical model puts important constraints on the possible outcomes of the simulation study while the experimental participants' "potential for independent action" allows for greater epistemic power (Morgan 2005, 325).

[24] For a detailed discussion, see Parke (2014).

Humphreys, the opacity of computer simulations is an epistemic challenge because it can "lead to a loss of understanding" (ibid., 148). The idea here seems to be that if we, as human agents, cannot trace how the input of a computer simulation leads to a given output, we may end up lacking an understanding of why the computer simulation produced the result that it did.

Others, too, have expressed the concern that the enormous 'inferential capacity' of simulations and our human inability to survey the whole simulation process matter for how computer simulations can figure in scientific explanations or contribute to understanding. Drawing parallels between computer simulations and complicated computer-aided mathematical proofs (such as the four color theorem), Beisbart points out that due to the sheer number of derivational steps involved, the respective computational processes can usually not be assessed in a stepwise fashion by epistemic agents (Beisbart 2012, 427). As he puts it: "This is important for the question of whether computer simulations may be explanatory" (Beisbart 2012, 427).

A related point has been raised by Lenhard. According to Lenhard, the opacity of simulations is "a major obstacle to explanatory potential" (Lenhard 2019, 224). Considering the earlier-mentioned example of a molecular dynamics simulation used to study the behavior of gold at the nanoscale[25], Lenhard pointedly states that "[...] the simulation does not offer an explanation in the usual sense" and that "[d]espite obviously being theory based, the simulation does not offer something like a theory-based insight into behavior" (Lenhard 2006, 609). According to Lenhard, the simulated behavior of gold at the nanoscale was at odds with scientists' expectations, even though the simulation was "based on well-known physics" (Lenhard 2006, 606). As he puts it, the theoretical resources underlying the simulation certainly "produced" the surprising behavior, but the exact relationship between theory and simulation output nevertheless remained opaque (Lenhard 2006, 609). In other words, the point seems to be that whereas the simulation made it possible to *generate* a certain result, it remained at the same time unintelligible how exactly this result relates to the theoretical assumptions that were fed into the simulation.

However, it is an open question whether the idea of 'theory-based insight' constitutes a good benchmark for explanations sought with the help of computer simulations in the first place. Quite generally, to assess if and how epistemic opacity hinders computer simulations from contributing to scientific explanations or enhancing understanding, it seems important to take a closer look at how exactly explanatory aims are pursued with the help of computer simulations, and if they indeed turn out to be a matter of relating simulation outputs to underlying theory.

[25]The example and Lenhard's argumentation is investigated more closely in chapter 3.

2.4 Computer Simulations and Scientific Explanations

In scientific research articles, it is regularly – implicitly or explicitly – suggested that computer simulations *can* and actually *do* make relevant contributions to explanations or understanding. Consider, for example, studies about the biological function of biomolecules. Besides making predictions of molecular behavior, an interesting question is as to *why* molecules exert a certain biological function, or *why* they show the behavior that they do. Computer simulations are frequently used by scientists to tackle problems such as the biological function of biomolecules (see Hollingsworth and Dror 2018) – and in some cases, they are specifically regarded as means for *explaining* such function. For instance, being concerned with the selective permeability of certain tunnel proteins at cell membranes – known as aquaporins[26] – researchers explicitly state that computer simulations could "reveal molecular mechanisms underlying the efficiency and selectivity of aquaporins, and thus *explain* biological function" (Hub et al. 2009, 58, my emphasis). As another example, in a seminal paper by Karplus and Kuriyan, molecular dynamics simulations are presented as important means for establishing links between protein structure and dynamics, and the authors suggest that there exist various examples in which these simulations could further "our understanding of protein functions" or that they come with the prospect of enabling a "deeper understanding of particular biological systems" (Karplus and Kuriyan 2005, 6684).

The Explanatory Power of Simulations

Over the last years, more and more philosophers have addressed the question of how computer simulations can provide explanations. Krohs (2008), for example, raises the question of how computer simulations can explain real-world processes in the first place: How can digital computers help generate explanations about the world? Resonating with Morgan and Morrison's models-as-mediators approach (see subsection 2.2.), Krohs emphasizes that explanations with computer simulations crucially rely on the mediatory role of theoretical models: "In the triangle of real-world process, theoretical model, and simulation, explanation of the real-world process by simulation involves a detour via the theoretical model" (Krohs 2008, 284). The theoretical model is what 'connects' the computer simulation to the world and allows us to infer explanations about phenomena in the world, even though we are working with digital computers.

In a related fashion, Weirich (2011) proposes that being based on models, computer simulations can offer "partial explanations" of phenomena in the world (Weirich 2011, 159). To do so, certain features of the theoretical model in a computer simulation need, according to Weirich, to

[26] See subsection 3.2. The example of aquaporins will be one of the two main examples used in this work and will be elaborated on in chapter 3.

be isomorphic with features of the natural world (Weirich 2011, 163). This also resonates with Bokulich (2011) who more broadly examines as to how explanations with models work and highlights that models need to 'reproduce' – in a wide sense – the relevant properties of the respective target of investigation (Bokulich 2011, 39).[27]

Without settling on whether the relationship between theoretical models and the world can appropriately be captured in terms of 'isomorphisms' or 'reproduction', the aforementioned contributions underscore the central role of theoretical models regarding the explanatory power of computer simulations.

How Simulation-Based Explanations Work

Having emphasized the explanatory relevance of theoretical models, one may ask further how exactly explanations obtained with the help of computer simulations work. Durán (2017), for example, suggests that the epistemic contributions of computer simulations can best be addressed from within a unificationist framework of explanation and understanding. Durán argues that the unificationist idea of using the same patterns of derivation over and over again[28] resonates with the case of computer simulations which allow to obtain a multiplicity of results from several simulation runs using similar theoretical resources but varying input parameters (Durán 2017).

Furthermore, computer simulations can, according to Durán, enable scientists to incorporate particular findings into a greater body of scientific beliefs. As a rather simple example, Durán (2017; 2018) considers a computer simulation of a satellite orbiting around a planet under tidal stress which generates as output an orbital eccentricity profile that shows certain characteristic 'spikes' (see Durán 2018, 7). According to Durán, when explaining the spikes as they were generated by the simulation, researchers could broaden "their body of scientific knowledge by incorporating a case derived from Newtonian mechanics" (Durán 2018, 119).

Another author who discusses the characteristics of explanations obtained by means of computer simulations is Imbert (2017). Discussing how computer simulations fit into various traditional models of explanation, Imbert proposes that some simulation-based explanations look like instances of what is traditionally captured under the header of the deductive-nomological (DN) model of explanation. According to the DN model, explanations are deductive arguments that include scientific laws as essential premises. In DN explanations, the explanandum must be a logical consequence of the explanans (Hempel and Oppenheim 1948, 137). As Imbert puts it:

[27] For an application of Bokulich's approach to the context of computer simulations, cf. chapter 4, section 2.3.

[28] According to Kitcher who prominently sketched the unificationist idea, "[s]cience advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same pattern of derivation again and again, and in demonstrating this, it teaches us how to reduce the number of facts we have to accept as ultimate" (Kitcher 1989, 432).

> "According to the deductive-nomological [...] model, one explains a phenomenon when a sentence describing it is logically deduced from true premises essentially containing a scientific law [...]. For example, the explanation of the trajectory of a comet, by means of a computer simulation of its trajectory based on the laws of classical (or relativistic) mechanics together with the initial positions of all bodies significantly influencing its trajectory, seems to qualify as a perfect example of DN explanation – provided that computer simulations can be seen as deductions [...]." (Imbert 2017, 751)

At first glance, even if the theoretical equations upon which computer simulations rely oftentimes do not have the character of straightforward implementations of scientific laws, the idea that simulations generate explanations by showing as to what happens if we apply our theoretical equations to concrete initial conditions may have a certain appeal. Starting from a set of theoretical equations or rules that describe the behavior of systems, computer simulations can help explore the (approximate) consequences of applying these equations to particular conditions.

At closer look, however, the issue is far from clear. Even if explanations obtained by means of computer simulations may in some cases 'respect'[29] the structure of unificationist or deductive-nomological explanations, I think it is not settled that these frameworks can render the explanatory contributions of computer simulations in scientific practice intelligible. Computer simulations are often used to study complex systems – be it the aforementioned biomolecules, climate phenomena or material behavior. In such cases, even if our simulations draw on sets of well-established theoretical equations, it is not clear whether explanations achieved with the help of these simulations can properly be understood as a matter of nomic expectability or of obtaining descriptions of various phenomena using a few patterns of derivation. In the nanosciences (as well as in many other scientific domains), explanatory questions arguably look quite different than in the case of the satellite or the comet. Rather than trajectories, explanatory questions often concern – as I will outline – the structural, organizational or functional features of complex systems. In order to understand the explanatory power of computer simulations in such contexts, it is instructive to first take a closer look at the explanatory questions at stake and at how computer simulations and their results concretely play a role in answering them.

[29] In a similar way in which computer simulations may, as Boge (2019) has put it, respect the logic of arguments while not being arguments, epistemically speaking.

3 This Work

I have proposed that there are several reasons why the role of computer simulations in the context of scientific explanation requires further philosophical attention. First of all, quite generally, given that explanation is widely considered an important goal of science, and given that computer simulations have become ubiquitous almost everywhere in science, philosophical studies concerned with explanation should care about the role of computer simulations in explanatory contexts.

Second, with the development of sophisticated modeling and simulation methods, the objects of scientific study have, in many places, become more and more complex. As it is recognized in the current philosophical literature, the case of complex systems[30] presents challenges to traditional perspectives on scientific explanations (cf. e.g. Woodward 2003; Mitchell 2009, 13, 76ff.; Woodward and Ross 2021). Since computer simulations are often key instruments in studies of complex systems, we can reasonably assume that our understanding how explanations of complex systems work will profit from paying attention to how computer simulations figure in studies of such systems.

Third, as it has been outlined earlier, there is an interesting tension that deserves further attention. On the one hand, computer simulations have been deemed epistemically opaque by some authors – and this opacity has been regarded as a possible obstacle to their explanatory potential. On the other hand, when looking into scientific practice, it seems that the use of computer simulations is sometimes not only perceived as *compatible with*, but even as *conducive to* the aim of obtaining explanations.

In what follows, I propose that to enhance our understanding of the explanatory role of computer simulations in science, it is fruitful to adopt a practice-oriented approach to how they figure in contexts of scientific explanation. Drawing on that, I outline the program and research aims of this dissertation.

3.1 Explanation and Philosophy of Science in Practice

In recent years, philosophical discourse on scientific explanation has, to some extent, experienced a change of focus. Rather than discussing idealized textbook examples or the shortcomings of the DN model of explanation, the attention has increasingly shifted towards detailed examinations of explanatory practices in various branches of science (Love 2015, 88). Furthermore, philosophers are more and more zooming in on the role of models and idealizations in explanations, or on the

[30] Drawing on a suggestion by Simon, I here refer to 'complex systems' as systems consisting of a large number of parts which interact in non-trivial ways, broadly understood (according to Simon [1969, 86] as cited in Stöckler 2000, 363-364).

collaborative character of explanatory practices (see e.g. Fagan 2015; Woody 2015; Potochnik 2017; Huneman 2018).

This development in the philosophical literature on scientific explanations resonates with a more general recent trend according to which philosophical studies of science should be sensitive to actual scientific practice[31] (cf. e.g. Soler et al. 2014). Whereas traditional discussions in the philosophy of science have often put emphasis on physics as a 'foundational' discipline and on the relationship between theories and the world as a central topic of concern (Ankeny et al. 2011, 305; see also Mission statement of the SPSP 2023), this traditional focus has increasingly been called into question and many branches in the philosophy of science have experienced what may be referred to as a trend of 'practice-orientation'.

The idea is that when aiming at capturing science in its manifold facets and in its various disciplines, investigations in the philosophy of science should be attentive to how scientists actually *do* science. As, e.g., Kaiser puts it:

> "More and more philosophers of science agree that philosophical theories about science must account for how science actually is done and must be informed, for instance, by the explanations developed in scientific practice and by the investigative strategies that scientists in fact employ." (Kaiser 2019, 36)

According to Ankeny et al., "[p]ractice consists of organized or regulated activities aimed at the achievement of certain goals" (Ankeny et al. 2011, 304). In this sense, being concerned with scientific practice means being concerned not only or primarily with the *results* of science, but with the scientific actions and reasoning patterns by means of which these results are obtained (Woody 2015, 80). For example, rather than dealing with the structure of ideal explanatory texts, a practice-oriented approach to explanations in science might prefer to start from the idea that 'explaining' is a form of coordinated activity whose function within certain scientific communities deserves scrutiny (Woody 2015, 79).

Drawing inspiration from this "re-orientation" (Woody 2015) of traditional discussions around explanation, this dissertation aims to take actual scientific practice as a starting point for scrutinizing the role of computer simulations in science, particularly within the context of explanation. By paying attention to how computer simulations are concretely employed by scientists in explanatory settings, my aim is to further a nuanced understanding of their explanatory contributions and epistemic role in science.

[31] At the same time, even though there is a current 'trend' of practice-orientation, the idea that philosophers should zoom in on how scientists actually do science is certainly not a completely novel one (see chapter 2, section 2; cf. also Soler et al. 2014).

3.2 Two Case Studies

In order to approach the explanatory contributions of computer simulations in a way that is sensitive to how computer simulations are actually used by scientists in explanatory contexts, the philosophical examinations in this dissertation will largely be centered around two concrete case studies.[32] The case studies belong to different branches of nanoscientific research: molecular biology (or more precisely *structural* biology) and physical chemistry. In both case studies, computer simulations played a major role; and in both cases, molecular structures or behavior at the nanoscale were found to be surprising.

The first case study is concerned with the biological function of aquaporins. Aquaporins are channel proteins located at cell membranes, where they regulate the flow of molecules between the inside and outside of cells. Aquaporins are highly selective channels: while they enable a rapid passage of water molecules through the cell membrane, they strictly exclude protons (de Groot and Grubmüller 2005). This exclusion matters for maintaining electrochemical gradients across the membrane (Hub et al. 2009, 58; de Groot and Grubmüller 2005), but the selectivity of aquaporins is notwithstanding surprising. This is because protons are traditionally expected to easily 'hop' from water molecule to water molecule through hydrogen bond networks, and because other aqueous pores seem to efficiently conduct protons (de Groot and Grubmüller 2005, 176). Various scientific groups tackled the puzzle of aquaporin selecivity and conducted studies on the biological function of aquaporins (see Yarnell 2004). In these studies, molecular dynamics simulations played a key role (cf. Hub et al. 2009, 58ff.).

The second case study concerns the surprising presence of certain atomic ions near the surface of small water droplets in atmospheric aerosols. Traditionally, most scientists believed that ions are not present at the interface between small water droplets and air (cf. e.g. Petersen and Saykally 2006, 333-34). Recent studies indicated, however, that certain ions can be found near the surface of small droplets, and it was assumed that this surprising presence matters for reactions occurring in the atmosphere on a larger scale (Caleman et al. 2011, 108). Similar to the case of aquaporins, various scientific groups addressed the explanatory question of why certain ions seem to prefer the surface rather than the inside of water droplets – and again, computer simulations were extensively used to study the phenomenon in question.

A major part of this dissertation will be concerned with an investigation of how concretely computer simulations were used by scientists to tackle the explanatory puzzles of aquaporin selectivity and ion surface preference, respectively. In particular, the interest lies in illuminating the research

[32] I only very briefly introduce the two case studies here and offer a more detailed motivation and background in chapters 3 and 4.

contexts in which the use of computer simulations was embedded, and in tracing how concretely they could support explanatory reasoning about the respective problems at stake.

3.3 Difference-making

When examining the role of computer simulations in scientific explanations, I will largely take for granted that computer simulations *can* and in fact regularly *do* play a relevant role in contexts of scientific explanation. Rather than casting general doubt on the explanatory power of computer simulations, I start from the idea that scientists frequently use computer simulations in explanatory contexts, and the interesting question is not so much if simulations can *really* help with explanations, but how their explanatory role can best be understood.[33]

That is, being concerned with understanding how computer simulations can play a role in scientific explanations, I adopt what may – drawing inspiration from Kaiser[34] – be regarded as a 'bottom-up' approach: I presuppose that computer simulations frequently figure in explanatory contexts and can make valuable contributions to explanations. Based on this, my aim is to clarify in more detail how their role in such contexts can be made sense of.

When drawing on the two case studies – that of aquaporin selectivity and that of ion surface preference – I aim to carve out how the scientific puzzles addressed in the two case studies can be considered as concrete instances of more abstract sorts of explanatory problems; and I aim to clarify how computer simulations could help tackle such problems.[35]

Zooming in on the two case studies, I propose that explanatory interests have much to do with an interest in dependence relations that hold under a range of variations in the respectively investigated systems. In the philosophical literature, this idea is prominently endorsed by proponents of what is commonly referred to as the counterfactual account of explanation in the tradition of Woodward (2003). Woodward regards explanation as a matter of exhibiting systematic patterns of counterfactual dependence (Woodward 2003, 191). The bottom line of the counterfactual account of explanation is that successful explanations show how the explanandum would have been different if things had been different in various regards (Woodward 2003, 11). In other words,

[33] I consider this approach to resonate with how Woodward motivates his approach to causation and explanation. Woodwards states: "my project focuses on (what I take to be) the purposes or goals behind our practices involving causal and explanatory claims; it is concerned with the underlying point of our practices" (Woodward 2003, 7). Related to this, I am interested in what the point of scientists' use of computer simulations is in contexts where explanatory questions are at stake.

[34] Dealing with reduction and reductive explanations in biology, Kaiser aims to develop a 'bottom-up' account of reductive explanations that "emerges from a critical reconstruction of the reductive reasoning and investigative practices that are present in contemporary biology" (Kaiser 2015, 186). Kaiser starts from the idea that reductions are common in biology, and the question is as to *how* such reductions can be understood (Kaiser 2015, 9f.). In parallel to this, I suggest that it is fruitful to start with recognizing that scientists actually use computer simulations for explanatory purposes – and the interesting question is as to how their contributions to explanation and their explanatory merits can be understood.

[35] I suggest that the problems of aquaporin selectivity and ion surface preference are problems where the *stability* or *expectedness* of certain (molecular) structures or processes are at stake.

explanations cite various factors which make a difference for the explanandum.

According to Woodward, the dependence relations that are of interest in contexts of explanation are those exploitable for purposes of manipulation and control (Woodward 2003, 10).Woodward's account draws on the idea of *interventions* which can be understood heuristically as manipulations in an idealized experimental setting (Woodward 1997, 29). The dependence relations that figure in explanations are addressed in terms of relations between *variables*. As Woodward puts it:

> "My view is that the sorts of counterfactuals that matter for purposes of causation and explanation are [...] counterfactuals that describe how the value of one variable would change under interventions that change the value of another." (Woodward 2003, 15)

Drawing on interventions, Woodward's account is essentially an account of *causal* explanation.[36] The scope, merits and limits of the counterfactual account are widely discussed in the recent literature (e.g. Ylikoski and Kuorikoski 2010; Saatsi and Pexton 2013; Rice 2021; Ross and Woodward 2023) and the causal-interventionist underpinnings of Woodward's account are subject to critical discussion (see e.g. Bokulich 2008; Bokulich 2011; Rice 2015; Reutlinger 2018).[37]

Rather than delving into the respective discussions on the causal and interventionist backdrop of Woodward's account, I will, in this dissertation, broadly take up the idea that explanatory interests can fruitfully be approached in terms of an interest in difference-making factors, and I propose that the counterfactual account offers useful conceptual and analytical resources for illuminating the role of computer simulations in contexts of scientific explanation.

To give a brief first idea of this, consider how computer simulations are appreciated by scientists as instruments for performing highly controlled manipulations – particularly in situations in which experimental approaches reach their limits. Elaborating on the role of computer simulations for studies of nanoparticles, computational physicist Amanda Barnard states that

> "[u]nlike our experimental colleagues, we have the luxury of being able to systematically test the relationship between our results and a particular parameter (such as temperature, pressure, charge, etc) independently, and we can see precisely how each parameter affects a system, while keeping all others constant. In a conventional experiment, if one physical parameter is changed, there are consequences and other experimental parameters change in response to it. Using theory and simulations it is also possible to study nanoparticles transplanted into highly non-equilibrium environments,

[36] Woodward acknowledges, however, that the idea of counterfactual dependence can be applied to non-causal contexts, too: "the common element in many forms of explanation, both causal and non-causal, is that they must answer what-if-things-had-been-diferent questions" (Woodward 2003, 221).

[37] Bokulich, for example, suggests that the counterfactual dependence relations figuring in explanations can be of different origins, causal origins just being one option (Bokulich 2011, 40).

such as at extreme temperatures, pressures, or in strong electric or magnetic fields, or even in the human body." (Barnard 2010, 2)

As this quote illustrates, computer simulations seem to be valued not only for granting access to scientific problems that are difficult or impossible to tackle by means of experiment, but also for enabling manipulations in a particularly tractable and controlled way. I suggest that recognizing how computer simulations enable scientists to perform targeted manipulations is also central for understanding their role in contexts of scientific explanation. Very briefly put, computer simulations can be powerful instruments for perfoming controlled manipulations, and such manipulations can matter for explanatory purposes as they enable insight into the factors that make a difference for the phenomenon to be explained.

3.4 Overall Research Aim

The overarching research aim of this work is twofold. The first aim is to contribute to a more comprehensive understanding of the role of computer simulations in the nanosciences and in contexts of scientific explanation. The second aim is to sketch a philosophical approach to computer simulations in the nanosciences that draws on actual scientific practice, and, more specifically, on how scientists actually use computer simulations in explanatory contexts.

In particular, this dissertation aims to bring the attention to the various ways in which computer simulations can support explanatory reasoning. Rather than mainly focusing on a particular diagnosis, e.g. that computer simulations are epistemically opaque, and rather than discussing their 'epistemic status' in general terms, the aim is to look at how concretely they are employed by scientists in contexts of explanation, and at how their explanatory role can be understood.

Furthermore, being concerned with computer simulations and scientific explanations, I aim to build bridges to two related topics of philosophical discussion. Firstly, I seek to connect my studies of computer simulations in the nanosciences to recent discussions on scientific explanations of complex systems. As it has been recognized, traditional approaches to explanation may not fit very well with the case of complex systems, and in order to account for the behavior of complex systems, oftentimes access to variables that 'code' for the expectedness or stability of certain behaviors or patterns is crucial (e.g. Woodward 2010; Batterman and Green 2021; Batterman 2021). One contribution of this dissertation is to show how computer simulations can be valuable instruments for this task. In this sense, the aim is not only to contribute to a better understanding of the explanatory power of computer simulations, but also to offer a nuanced perspective on recent debates on scientific explanations of complex systems.

Secondly, I aim to offer connecting points for discussions on multiscale simulations and reduction. It has recently been argued that the success of multiscale simulations calls traditional reductionist perspectives on inter-level relationships and scientific explanation into question (e.g. Winsberg 2006; Green and Batterman 2017; Bursten 2018). Contributing to these discussions, I take a look at how bridges between different scales or levels of description in multiscale simulation contexts work, and I suggest that to investigate how descriptions and explanations of phenomena across multiple scales can be obtained, it is useful to pay close attention to the mediatory role of models in multiscale simulations.

3.5 Overview

While it is, in current discussions, increasingly stressed that "it is crucial to have productive interactions between philosophical analyses and the study of actual scientific practices" (Ankeny et al. 2011), philosophers' use of case studies has, at the same time, given rise to worries. In particular, the concern has been raised that the philosophical reliance on case studies could lead to overgeneralization or bias. Responding to such concerns, chapter 2 discusses how philosophical examinations of scientific explanations can be closely informed by actual scientific practice without simply confirming what particular scientists would consider a successful explanation, and without drawing "hasty conclusions" from a few "conveniently chosen" cases, as Chang puts it (Chang 2012, 109).

Both chapter 3 and chapter 4 zoom in on how computer simulations concretely figure in contexts of scientific explanation. In chapter 3, recent discussions on the (supposedly) opaque character of complex simulations are taken as a starting point for examining the role of computer simulations in scientific explanations. Concretely, it is investigated how, in the case of aquaporins, atomistic molecular dynamics simulations could help scientists with the search for an explanation of the surprising selectivity of aquaporins. By elucidating how precisely computer simulations allowed it to draw inferences about the molecular system under investigation, responses are offered to some concerns regarding their explanatory power. It is outlined that atomistic simulations can be instrumental for managing molecular complexity and for systematically assessing how the occurrence of the explanandum is sensitive to a range of factors.

In chapter 4, the second case study is discussed that concerns the surprising behavior of certain ions in aerosols. Following up on the earlier developed idea that computer simulations can be instruments for managing complexity, the chapter zooms in on the concrete contributions that computer simulations can make to scientific explanations. To that end, the chapter takes a look at how computer simulations have been employed by scientists to support explanatory reasoning

about the surprising behavior of certain ions in atmospheric aerosol droplets.

Drawing on conceptual resources from the counterfactual account of scientific explanation broadly understood, it is suggested that atomistic computer simulations can support explanatory reasoning in at least three relevant ways, namely by (1) providing an opportunity for effectively performing controlled manipulations that would otherwise largely remain unfeasible, by (2) granting epistemic access to levels of description at which the variables that potentially figure in explanations of complex systems are specified, and by (3) offering a framework for quantitative analysis.

In chapter 5, I address a further question relevant for discussions on the role of computer simulations in explanations, namely the question of how different models and theories can 'communicate' in multiscale simulation contexts. Multiscale simulations are simulations that capture the behavior or properties of systems across multiple scales or levels of description. Recently, various authors have suggested that the scientific success of multiscale simulations challenges traditional reductionist ideas concerning how different models and theories can be connected to each other, or how the behavior of multiscale systems can be described and explained. Contributing to these discussions, the final chapter of the dissertation suggests that the combination of different levels of description in multiscale simulations deserves further attention. By drawing on Morgan and Morrison's models-as-mediators approach, the chapter explores how models in multiscale simulation contexts act as partly autonomous mediators between theories and the world. It is suggested that recognizing the mediatory role of models in multiscale simulations offers a fruitful starting point for enhancing our understanding of how different levels of description can be integrated and 'work together' in multiscale simulation frameworks.

Bibliography

Abraham, F. F., Broughton, J. Q., Bernstein, N., and Kaxiras, E. (1998). Spanning the continuum to quantum length scales in a dynamic simulation of brittle fracture. *EPL (Europhysics Letters)*, 44(6):783.

Alvarado, R. (2022). Computer simulations as scientific instruments. *Foundations of Science*, (27):1183–1205.

Ankeny, R., Chang, H., Boumans, M., and Boon, M. (2011). Introduction: Philosophy of science in practice. *European Journal for Philosophy of Science*, 1:303–307.

Barnard, A. (2010). Modelling of nanoparticles: approaches to morphology and evolution. *Reports on Progress in Physics*, 73(8):086502.

Batterman, R. W. (2021). *A Middle Way: A Non-Fundamental Approach to Many-Body Physics*. Oxford University Press.

Batterman, R. W. and Green, S. (2021). Steel and bone: mesoscale modeling and middle-out strategies in physics and biology. *Synthese*, 199(1-2):1159–1184.

Bayda, S., Adeel, M., Tuccinardi, T., Cordani, M., and Rizzolio, F. (2019). The history of nanoscience and nanotechnology: from chemical–physical applications to nanomedicine. *Molecules*, 25(1):112.

Beisbart, C. (2012). How can computer simulations produce new knowledge? *European Journal for Philosophy of Science*, 2(3):395–434.

Beisbart, C. and Norton, J. D. (2012). Why monte carlo simulations are inferences and not experiments. *International Studies in the Philosophy of Science*, 26(4):403–422.

Berryman, S. (2023). Democritus. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Boge, F. J. (2019). Why computer simulations are not inferences, and in what sense they are experiments. *European Journal for Philosophy of Science*, 9(1):13.

Bokulich, A. (2008). Can classical structures explain quantum phenomena? *The British Journal for the Philosophy of Science*, 59(2):217–235.

Bokulich, A. (2011). How scientific models can explain. *Synthese*, 180:33–45.

Broughton, J. Q., Abraham, F. F., Bernstein, N., and Kaxiras, E. (1999). Concurrent coupling of length scales: methodology and application. *Physical Review B*, 60(4):2391.

Bucher, F., Dmitriev, A., Ertz, M., Knothe, K., Popov, V., Psakhie, S., and Shilko, E. (2006). Multiscale simulation of dry friction in wheel/rail contact. *Wear*, 261(7-8):874–884.

Bursten, J. R. (2018). Conceptual strategies and inter-theory relations: The case of nanoscale cracks. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 62:158–165.

Caleman, C., Hub, J. S., van Maaren, P. J., and van der Spoel, D. (2011). Atomistic simulation of ion solvation in water explains surface preference of halides. *Proceedings of the National Academy of Sciences*, 108(17):6838–6842.

Chang, H. (2012). Beyond case-studies: History as philosophy. In Mauskopf, S. and Schmaltz, T., editors, *Integrating history and philosophy of science: Problems and prospects*, pages 109–124. Springer.

Cortie, M. B. (2004). The weird world of nanoscale gold. *Gold Bulletin*, 37(1-2):12–19.

Dahn, D. C., Watanabe, M., Blackford, B., Jericho, M., and Beveridge, T. (1988). Scanning tunneling microscopy imaging of biological structures. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 6(2):548–552.

de Groot, B. L. and Grubmüller, H. (2005). The dynamics and energetics of water permeation and proton exclusion in aquaporins. *Current Opinion in Structural Biology*, 15(2):176–183.

Durán, J. M. (2017). Varying the explanatory span: scientific explanation for computer simulations. *International Studies in the Philosophy of Science*, 31(1):27–45.

Durán, J. M. (2018). *Computer Simulations in Science and Engineering: Concepts - Practices - Perspectives*. Springer.

Fagan, M. B. (2015). Collaborative explanation and biological mechanisms. *Studies in History and Philosophy of Science Part A*, 52:67–78.

Feynman, R. P. (1960). There's plenty of room at the bottom. *Int. J. Eng. Sci*, 23(8):22–36.

Frigg, R. and Reiss, J. (2009). The philosophy of simulation: hot new issues or same old stew? *Synthese*, 169:593–613.

Galison, P. (1996). Computer simulations and the trading zone. In Galison, P. and Stump, D. J., editors, *The Disunity of Science: Boundaries, Contexts, and Power*, pages 118–157. Stanford University Press.

Green, S. and Batterman, R. W. (2017). Biology meets physics: Reductionism and multi-scale modeling of morphogenesis. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 61:20–34.

Grüne-Yanoff, T. and Weirich, P. (2010). The philosophy and epistemology of simulation: A review. *Simulation & Gaming*, 41(1):20–50.

Grünke, P. (2023). *Computer-based methods of knowledge generation in science: What can the computer tell us about the world?* Karlsruhe Institute of Technology, Dissertation Thesis.

Grünke, P. and Schweer, J. (2023). How do we learn from computer simulations? a critical examination of the argument view of simulation. In *Computer-based methods of knowledge generation in science: What can the computer tell us about the world?*, Paul Grünke. Karlsruhe Institute of Technology, Dissertation Thesis.

Hartmann, S. (1996). The world as a process: Simulations in the natural and social sciences. In Hegselmann, R., Mueller, U., and Troitzsch, K. G., editors, *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*, pages 77–100. Springer.

Hempel, C. G. and Oppenheim (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2):135–175.

Hillerbrand, R. (2013). Order out of chaos? a case study in high energy physics. *Studia Philosophica Estonica*, pages 61–78.

Hollingsworth, S. A. and Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143.

Hub, J. S., Grubmüller, H., and De Groot, B. L. (2009). Dynamics and Energetics of Permeation through Aquaporins. What Do We Learn from Molecular Dynamics Simulations? In *Aquaporins*, pages 57–76. Springer.

Humphreys, P. (1991). Computer simulations. In Fine, A., Forbes, F., and Wessels, L., editors, *PSA 1990: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume 2, pages 497–506. University of Chicago Press.

Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press.

Humphreys, P. (2009). The philosophical novelty of computer simulation methods. *Synthese*, 169(3):615–626.

Huneman, P. (2018). Diversifying the picture of explanations in biological sciences: ways of combining topology with mechanisms. *Synthese*, 195(1):115–146.

Imbert, C. (2017). Computer simulations and computational models in science. In Magnani, L. and Bertolotti, T., editors, *Springer Handbook of Model-Based Science*, pages 735–781. Springer.

Kaiser, M. I. (2015). *Reductive Explanation in the Biological Sciences*. Springer.

Kaiser, M. I. (2019). Normativity in the philosophy of science. *Metaphilosophy*, 50(1-2):36–62.

Karplus, M. and Kuriyan, J. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences*, 102(19):6679–6685.

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In Kitcher, P. and Salmon, W., editors, *Scientific Explanation*, pages 410–505. Minneapolis: University of Minnesota Press.

Kleppner, D. and Jackiw, R. (2000). One hundred years of quantum physics. *Science*, 289(5481):893–898.

Krohs, U. (2008). How digital computer simulations explain real-world processes. *International Studies in the Philosophy of Science*, 22(3):277–292.

Kumar, S., Nehra, M., Kedia, D., Dilbaghi, N., Tankeshwar, K., and Kim, K.-H. (2018). Carbon nanotubes: A potential material for energy conversion and storage. *Progress in Energy and Combustion science*, 64:219–253.

Landman, U. (2005). Materials by numbers: Computations as tools of discovery. *Proceedings of the National Academy of Sciences*, 102(19):6671–6678.

Landman, U., Luedtke, W., and Gao, J. (1996). Atomic-scale issues in tribology: interfacial junctions and nano-elastohydrodynamics. *Langmuir*, 12(19):4514–4528.

Lenhard, J. (2006). Surprised by a nanowire: Simulation, control, and understanding. *Philosophy of Science*, 73(5):605–616.

Lenhard, J. (2019). *Calculated Surprises: A Philosophy of Computer Simulation*. Oxford University Press.

Lindsay, S. (2010). *Introduction to Nanoscience*. Oxford University Press.

Love, A. C. (2015). Collaborative explanation, explanatory roles, and scientific explaining in practice. *Studies in History and Philosophy of Science Part A*, 52:88–94.

Marcovich, A. and Shinn, T. (2014). *Toward a New Dimension: Exploring the Nanoscale*. Oxford University Press.

Massimi, M. and Bhimji, W. (2015). Computer simulations and experiments: The case of the higgs boson. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 51:71–81.

Mission statement of the SPSP (2023). Mission statement of the society for philosophy of science in practice. http://philosophy-science-practice.org/about/mission-statement/ (URL date 16.05.2023).

Mitchell, S. D. (2009). *Unsimple Truths: Science, Complexity, and Policy*. University of Chicago Press.

Moore, N. W., Luo, J., Huang, J., Mao, S. X., and Houston, J. (2009). Superplastic nanowires pulled from the surface of common salt. *Nano letters*, 9(6):2295–2299.

Morgan, M. S. (2005). Experiments versus models: New phenomena, inference and surprise. *Journal of Economic Methodology*, 12(2):317–329.

Morgan, M. S. and Morrison, M., editors (1999). *Models as Mediators*. Cambridge University Press.

Morrison, M. (2007). Where have all the theories gone? *Philosophy of Science*, 74(2):195–228.

Morrison, M. (2009). Models, measurement and computer simulation: the changing face of experimentation. *Philosophical Studies*, 143(1):33–57.

Morrison, M. (2015). *Reconstructing Reality: Models, Mathematics, and Simulations*. Oxford Studies in Philosophy of Science.

Morrison, M. and Morgan, M. S. (1999). Models as mediating instruments. In Morrison, M. and Morgan, M. S., editors, *Models as Mediators*, pages 10–37. Cambridge University Press.

Nosonovsky, M. and Bhushan, B. (2007). Multiscale friction mechanisms and hierarchical surfaces in nano-and bio-tribology. *Materials Science and Engineering: R: Reports*, 58(3-5):162–193.

Parke, E. C. (2014). Experiments, simulations, and epistemic privilege. *Philosophy of Science*, 81(4):516–536.

Parker, W. S. (2009). Does matter really matter? computer simulations, experiments, and materiality. *Synthese*, 169(3):483–496.

Petersen, P. B. and Saykally, R. J. (2006). On the nature of ions at the liquid water surface. *Annual Review of Physical Chemistry*, 57:333–364.

Potochnik, A. (2017). *Idealization and the Aims of Science*. University of Chicago Press.

Reutlinger, A. (2018). Extending the counterfactual theory of explanation. In Reutlinger, A. and Saatsi, J., editors, *Explanation beyond Causation: Philosophical Perspectives on Non-Causal Explanations*, pages 74–95. Oxford University Press.

Rheinberger, H.-J. (2006). Schnittstellen. Instrumente und Objekte im experimentellen Kontext der Wissenschaften vom Leben. In Schramm, H., Schwarte, L., and Lazardzig, J., editors, *Instrumente in Kunst und Wissenschaft. Zur Architektonik kultureller Grenzen im 17. Jahrhundert*, volume 17, pages 1–20. De Gruyter.

Rice, C. (2015). Moving beyond causes: Optimality models and scientific explanation. *Noûs*, 49(3):589–615.

Rice, C. (2021). *Leveraging Distortions: Explanation, Idealization, and Universality in Science*. MIT Press.

Rohrlich, F. (1990). Computer simulation in the physical sciences. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2, pages 507–518. Cambridge University Press.

Ross, L. and Woodward, J. (2023). Causal Approaches to Scientific Explanation. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Saatsi, J. and Pexton, M. (2013). Reassessing woodward's account of explanation: Regularities, counterfactuals, and noncausal explanations. *Philosophy of Science*, 80(5):613–624.

Schweer, J., Hillerbrand, R., and Elstner, M. (fort). Simulation. In Gutmann, M., Wiegerling, K., and Rathgeber, B., editors, *Handbuch Technikphilosophie*. J. B. Metzler.

Soler, L., Zwart, S., Lynch, M., and Israel-Jost, V. (2014). *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*. Routledge.

Srivastava, A., Nagai, T., Srivastava, A., Miyashita, O., and Tama, F. (2018). Role of computational methods in going beyond x-ray crystallography to explore protein structure and dynamics. *International Journal of Molecular Sciences*, 19(11):3401.

Stöckler, M. (2000). On modeling and simulations as instruments for the study of complex systems. In Carrier, M., Massey, G. J., and Ruetsche, L., editors, *Science at Century's End: Philosophical Questions on the Progress and Limits of Science*, pages 355–373. University of Pittsburgh Press.

Stroscio, J. A. and Eigler, D. (1991). Atomic and molecular manipulation with the scanning tunneling microscope. *Science*, 254(5036):1319–1326.

Wang, Y., Ribeiro, J. M. L., and Tiwary, P. (2020). Machine learning approaches for analyzing and enhancing molecular dynamics simulations. *Current Opinion in Structural Biology*, 61:139–145.

Weirich, P. (2011). The explanatory power of models and simulations: A philosophical exploration. *Simulation & Gaming*, 42(2):155–176.

Weisberg, M. (2012). *Simulation and similarity: Using models to understand the world*. Oxford University Press.

Winsberg, E. (1999). Sanctioning models: The epistemology of simulation. *Science in context*, 12(2):275–292.

Winsberg, E. (2006). Handshaking your way to the top: Simulation at the nanoscale. *Philosophy of Science*, 73(5):582–594.

Winsberg, E. (2018). *Philosophy and climate science*. Cambridge University Press.

Winsberg, E. (2019). Computer Simulations in Science. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2019 edition.

Woodward, J. (1997). Explanation, invariance, and intervention. *Philosophy of Science*, 64(S4):S26–S41.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, 25:287–318.

Woodward, J. and Ross, L. (2021). Scientific Explanation. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edition.

Woody, A. I. (2015). Re-orienting discussions of scientific explanation: A functional perspective. *Studies in History and Philosophy of Science Part A*, 52:79–87.

Yang, W., Thordarson, P., Gooding, J. J., Ringer, S. P., and Braet, F. (2007). Carbon nanotubes for biological and biomedical applications. *Nanotechnology*, 18(41):412001.

Yarnell, A. (2004). Blockade in the Cell's Waterway. *Chem. Eng. News*, 82(4):42–44.

Ylikoski, P. and Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies*, 148(2):201–219.

Šešelja, D. (2023). Agent-Based Modeling in the Philosophy of Science. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2023 edition.

Scientific Explanations in the Wild: On Case Studies in the Philosophy of Science in Practice

Contents

Summary

In recent discussions, there is wide agreement that philosophy of science should be rooted in actual scientific practice. This trend is, perhaps, particularly evident in discussions around scientific explanations, where philosophers are moving away from idealized textbook examples to scrutinize real-world scientific explanations through an in-depth engagement with case studies from various fields. However, the use of case studies has elicited concerns regarding the risk of bias and overgeneralization. This paper addresses such concerns and underscores how case studies can play a valid and relevant role in philosophical examinations of science.

Building on a proposal by Hasok Chang, I suggest that rather than acting as pieces of evidence for inductive generalizations, one way in which case studies can figure in philosophical examinations is as resources for abstraction. Discussing a concrete example from molecular biology, I further spell out this idea and demonstrate how detailed analyses of patterns of scientific reasoning can offer nuanced insights into explanatory practices. I demonstrate how the suggested perspective shifts the focus away from searches for general theories of scientific explanation, and rather towards an examination of the reasoning patterns and tools that figure in certain explanatory practices. I conclude by outlining how the suggested perspective on case studies also contributes to a more comprehensive understanding of normativity within the philosophy of science in practice.

1 Introduction

"[I]f case studies are the paving stones, where does the path lead?"

– (Galison 2008, 120)

The question of how scientists explain is arguably keeping philosophers of science busy ever since. Traditionally, much of the philosophical literature on scientific explanation has focused on the physical sciences, and discussions have often centered around a few paradigmatic examples, be it that of flagpole shadow lengths, that of barometers and storms, that of men who take birth control pills, or that of pegs that fit through square but not round holes (Giere 2012, 63; Love 2015, 88).

Over the last years, however, more and more philosophers have promoted a philosophy of science that pays closer attention to actual scientific practice, and the recent literature on scientific explanation resonates with this trend. Concretely, philosophical engagements with scientific explanation have become more and more specialized, often turning away from idealized textbook examples, and instead zooming in on case studies from different scientific disciplines (Love, 2015: 88). Moreover, rather than being treated as individual achievements, explanations are increasingly perceived as a matter of social (often interdisciplinary and incremental) collaborative effort (Love 2015; Woody 2015), and it is recognized that they often involve the extensive use of sophisticated models or instruments – which themselves require special philosophical examination (e.g. Bokulich 2011; Bokulich 2014; Gelfert 2016).

Yet, while being widely supported in the philosophical community, the idea that philosophical engagement with scientific explanations should be rooted in actual scientific practice also comes with several challenges and open questions. In particular, the role of case studies requires further

clarification. Many of the concerns regarding philosophers' use of case studies are addressed by Pitt (2001). According to Pitt, philosophy faces a dilemma with case studies: on the one hand, if case studies are used to provide evidence or test certain philosophical claims, then it is unclear how one can avoid choosing them to one's convenience (Pitt 2001, 375). On the other hand, if one wants to start from 'neutral grounds' and use case studies to develop a philosophical claim or theory in the first place, then it is unclear how generalizations from individual cases are supposed to work (Pitt 2001, 373).

In this paper, I draw on various recent contributions in the philosophy of science in practice to respond to such concerns. I suggest that Pitt's dilemma is a false one because it rests on a problematic understanding of how case studies can figure in philosophical reasoning. Drawing on Chang (2012), I propose that rather than acting as pieces of evidence for inductive generalizations, case studies can play a legitimate and valuable role in philosophical examinations as resources for abstraction. To illustrate and further spell out this idea, I discuss a concrete example from molecular biology that concerns a scientific disagreement on how the biological function of certain channel proteins in human cells can be explained. In section 2, I introduce and briefly discuss the idea of practice-orientation in the contemporary philosophy of science. In section 3, I present Pitt's dilemma to give an idea of how philosophers' use of case studies has given rise to concerns. As illustration for my own response to the dilemma, section 4 introduces the example of the biological function of aquaporins, which are certain channel proteins sitting at the membranes of biological cells. Building on that, section 5 explores how case studies can figure in philosophical reasoning, and how there can be room for normativity in practice-oriented philosophy of science.

2 On practice-oriented philosophy of science

The concern of this paper is the role of case studies in practice-oriented philosophy of science. Hence, before turning to case studies, it is important to shed light on both the idea of 'practice-orientation', and the notion of 'practice' itself.

Despite having gained increased popularity in recent times, practice-orientation in the philosophy of science is certainly not a completely novel phenomenon (Ankeny et al. 2011; Soler et al. 2014). Oftentimes, the philosophy of late Wittgenstein with its idea of 'meaning as use' (Wittgenstein 1953) is regarded as an early proponent of practice-oriented philosophy in the 20th century (Soler et al. 2014, 4). Arguably, another important forerunner of the contemporary trend of practice-oriented philosophy of science was the movement of 'New Experimentalism', which originated in the 1980s (Woody 2014, 124). Criticizing the theory-centrism of traditional philoso-

phy of science, proponents of New Experimentalism, such as Hacking (1983) or Franklin (1989), stressed the diversity of experimental practices and their roles in the sciences (see also Soler et al., 2014: 7f.). Aiming at overcoming a narrow understanding of experimentation as theory testing (cf. Woody, 2014: 124), the New Experimentalists focused on detailed investigations of the role of instruments and measurements in experiments (e.g. Galison 1987), or of the social aspects of experimental activity, among other things (e.g. Knorr-Cetina 1999).

Another field with high relevance for – and intertwinement with – current approaches in practice-oriented philosophy of science is the integrated History and Philosophy of Science (HPS) which has its origins in the 1960s and in the 'historical turn' discussed in the context of Kuhn's *The Structure of Scientific Revolutions* (Kuhn 1962; Vos 2021). The reference to historical case studies has sparked a methodological controversy in the field of HPS, and many aspects of the respective discussions matter for how the role of case studies is considered in current debates in the practice-oriented philosophy of science, as I will point out in more detail in the next section. At the heart of the controversy is, as Chang puts it, that "HPS has witnessed too many hasty philosophical generalizations based on a small number of conveniently chosen case-studies" (Chang 2012, 109) and that this was perceived as problematic for both philosophy and history.

In its current form, the trend[1] of practice-oriented philosophy of science has emerged at least partly as a reaction to a perceived lack of concern about scientific practice in the contemporary mainstream Anglophone philosophy of science (Ankeny et al. 2011, 303). In their mission statement, the 'Society for Philosophy of Science in Practice' (SPSP), which was founded in 2006 by scholars promoting a practice-oriented approach philosophy of science, states that despite some exceptions, "the concern with practice has always been somewhat outside the mainstream of English-language philosophy of science." (Mission statement of the SPSP 2023). The declared aim of the SPSP is to "change this situation, through a conscious and organized programme of detailed and systematic study of scientific practice" (Mission statement of the SPSP 2023). Rather than dealing with "excessively idealized reconstructions" of science, the main idea of practice-oriented philosophy is to pay close attention to how scientists actually do science (Soler et al. 2014, 12).

Practice-orientation in the contemporary philosophy of science comes in different shapes. Some philosophers have adopted empirical methods to study scientific practice (e.g. Osbeck and Nersessian 2015) or even joined scientific research labs (Bursten 2020). Others have promoted a re-orientation of traditional discussions, e.g. on scientific explanation (Woody 2015) or reduction

[1]The increasing orientation towards scientific practice is sometimes alternatively referred to as a new practice 'turn' in the philosophy of science. Given that this term experiences a partly inflationary use and may evoke the idea that there is a clear 'before' and 'after' phase in the philosophical discourse (Soler et al. 2014, 2f.), I consider it clearer to speak of a 'trend' of practice-orientation in the philosophy of science.

(Chang 2011, Kaiser 2015), focusing less on the 'nature' of these things and more on the different shapes and roles of explanations or reductions in the sciences.

When speaking of a trend of practice-orientation in the philosophy of science, the notion of 'practice' requires clarification. As Ankeny et al. put it, "[p]ractice consists of organized or regulated activities aimed at the achievement of certain goals" (Ankeny et al. 2011, 304). As, for example, Chang (2011, 209) illustrates, scientists engage in many different activities, including describing, explaining, hypothesizing, testing, observing, measuring, classifying, representing, modeling, simulating, synthesizing, or analyzing – just to name a few (Chang 2011, 209).

Chang stresses that studying scientific 'practice' comes with a "change of focus from propositions to activities" (2011, 208), and that this change of focus means asking different questions. Undertaking the project of sketching a philosophical grammar[2] of scientific practice, Chang proposes that doing practice-oriented philosophy of science means transforming philosophical questions; turning away from questions such as 'what is the nature of x' or 'what is a correct account of x', and rather addressing questions like: "what purpose do activities aiming at x serve?" (Chang 2011, 216).

Consider the traditional topic of 'scientific reductionism'. Chang holds that instead of asking whether reductionism is true, it is more interesting to ask as to "what we do when we make reductions, and what purpose reductions serve" (Chang 2011, 216). This resonates with other proponents of practice-oriented philosophy of science, such as Kaiser, who endorses a bottom-up approach to the topic of reduction, assuming that reductions are quite common in the sciences, and aiming at identifying a sense of reduction that really captures scientists' practices (Kaiser 2015: 12f.). Another example of practice-oriented philosophy of science's shift in focus is Woody's functional account of explanation (e.g. Woody 2015). Rather than dealing with individual explanations as the results of scientific inquiry, Woody considers it more interesting and relevant to look at the roles that explanatory discourse play in science, and the aims it can serve.

In a nutshell, there are two particularly important (and intertwined) motifs in contemporary practice-oriented philosophy. First, a commonly shared idea is that practice-oriented philosophy of science should be empirically informed, i.e., closely informed by how scientists in fact do science. A second idea is that philosophers of science should not limit their inquiries to e.g. the relationship between theories and the world, or to an inquiry of the nature of scientific explanation, or to the results of scientific investigations; but instead pay attention to the processes that bring about these results, and the manifold activities in which scientists engage.

[2]'Grammar' understood in a Wittgensteinian sense, i.e., as concerning the implicit or explicit rules that need to be respected when successfully using concepts in a context of discourse (Chang 2011, 208).

3 Between bias and hasty generalizations? Pitt's dilemma

Both, the idea of being closely oriented towards how scientists actually do science, and the aim of paying particular attention to scientific activities naturally suggest that practice-oriented philosophy should zoom in on concrete case studies. Despite becoming more and more prevalent, philosophers' attention to concrete case studies has also sparked a host of worries. Chang (2012) diagnoses a "trouble" with case studies; Currie (2015) even states that there appears to be a "curse of the case study".

Being concerned with historical case studies in the HPS, Pitt claims that the use of case studies faces a dilemma, and he draws the puzzling conclusion that "even very good case studies do no philosophical work" (Pitt 2001, 373). Pitt considers two general ways in which case studies could figure in philosophical reasoning. Briefly put, his dilemma goes as follows: on the one hand, case studies may serve to *support* or *exemplify* a philosophical theory. This, however, seems hardly feasible without begging the question, as the selection and interpretation of case studies is inevitably prone to bias (Pitt 2001, 373). On the other hand, concrete case studies could be employed to develop a philosophical theory in the first place. According to Pitt, this, too, fails because it is not clear how generalizations from case studies are supposed to work (Pitt 2001, 373).

Other philosophers have criticized Pitt's dilemma as a false one (Burian, 2001; Currie, 2015). I agree, and part of my paper will be to illustrate how case studies can figure in philosophical engagement with scientific practice without culminating in either horn of the dilemma. Before doing so, I will first introduce the dilemma and illustrate how it resonates with other philosophical concerns about the use of case studies.

3.1 The first horn of the dilemma: case studies as a tribunal

Rather than engaging in armchair discussions about how science should ideally be done, practice-oriented philosophy suggests that philosophical engagement with science should be sensitive to how science is in fact done. One interpretation of this statement is that concrete case studies should be used to illuminate a given philosophical theory or to provide the 'data' against which philosophical reasoning and theorizing about science can be evaluated (Pitt 2001, 379f.). In other words, concrete case studies could possibly be used to serve as a tribunal for our philosophical theories about science, allowing us to assess whether these theories are sound. This view on the role of case studies could be called hypothetico-deductivist: philosophy presents the hypotheses and case studies are the tests (Kinzel 2015, 49). However, the first horn of Pitt's dilemma is that it

is unclear how case studies can fulfill this role.

Suppose we begin with a theory of scientific explanation and examine a case study to test it. As Pitt points out, it is questionable how this can be done without begging the question (Pitt 2001, 375). To actually learn something about scientific explanations, we should be confident, to some degree, that the case study is both 'scientific' and concerned with an 'explanation'. According to Pitt, it is not clear what the proper selection criteria are, i.e., how we can ensure that case studies are not chosen to our convenience; or that they are not already biased towards our aim of supporting (or rejecting) the theory in question (see Pitt 2001, 374f.).

A related problem is identified by Woody (2015) and Love (2015). Discussing her functional perspective on explanation, Woody notes that worries could arise regarding the adequacy conditions for explanatory success (Woody 2015, 83). If concrete explanations given by scientists are the hallmark of an understanding of what scientific explanations are or how they work, then how can we start with case studies without already having settled on criteria that allow us to assess as to when particular candidate explanations are successful?

In this sense, Pitt's original problem also turns out to be a problem of normativity: if our understanding of how scientific explanations work, or when they are successful, is obtained by looking at actual scientific cases, then how can this be done without circularity or without already having settled, before looking into practice, on certain criteria for explanatory success?

One could perhaps be inclined to accept that case-study based philosophical investigations of science cannot do more than offer detailed reconstructions of particular scientific argumentations. In a sense, this would certainly resonate with the idea that the practice-orientation in the philosophy of science marks a shift "from normative to descriptive perspectives on science" (Soler et al. 2014, 15). However, withdrawing from making any kinds of normative claims and restricting oneself to mere reconstructions of particular cases appears highly dissatisfying[3]. Even if one's goal is to develop one's philosophical account in close connection to how scientists themselves formulate their explanatory aims and successes, most philosophers of science would arguably insist that their engagement with scientific practice goes beyond a mere reconstruction of what, in a particular case, is being regarded as an explanation.

[3]Furthermore, as it has been pointed out by Kaiser, even a philosophy of science that would restrict itself to making only factual claims about what scientists do cannot let go of normativity entirely. The reason is that the development of factual claims requires making decisions about how to select, interpret and evaluate empirical information, e.g. depicting case studies in the first place – and these decisions make normative judgment inevitable (Kaiser 2019, 57).

3.2 The second horn of the dilemma: generalizations and the inductive leap

Instead of using case studies to exemplify or test one's philosophical claim, one could try to start from a blank sheet and look at case studies to *develop* a philosophical claim or theory in the first place. Since beginning with assumptions about how scientific explanations work and then subsequently testing these assumptions leads to problems, one could try to start with a question (e.g., "what is a scientific explanation?") and then proceed to an examination of case studies to answer it.

According to Pitt, this brings us to the second horn of the dilemma: "if one starts with a case study, it is not clear where to go from there—for it is unreasonable to generalize from one case or even two or three" (Pitt 2001, 373). In other words, if the aim is to make general claims about science, and if case studies are taken as the basis for such claims, it is not clear how one can justify jumping from statements about particular cases to more general conclusions. The reason is that science is, according to Pitt, in constant flux, making it hard to justify why an examination of a particular case study could yield any conclusions beyond the respective case under investigation (Pitt 2001, 381).

As Currie phrases this problem, "[...] case studies, by their nature, are peculiar and individual, but the generalizations philosophers seek are broad and unitary" (Currie 2015, 554). If philosophical reasoning is based on particular case studies, then how can it lead to any insights beyond these case studies?

Pitt is not the only one to identify this problem. In a likewise critical fashion, Chang (2012, 109) reports that the field of history and philosophy of science "has witnessed too many hasty philosophical generalizations based on a small number of conveniently chosen case-studies". Relatedly, Bishop and Trout (2002, 204) state that case studies are often cited by philosophers and historians of science to "propose general hypotheses about how science works" – and just as Pitt, they are highly doubtful that telling stories about particular scientific cases can serve to support more general claims. Quite the opposite, they characterize the narrative reference to particular case studies a deficient methodology and, looking at the prevalence of naturalism in contemporary philosophy of science, they claim that it is "[...] ironic that naturalistic philosophers – philosophers who are inclined to see no principled methodological gaps separating science and philosophy – employ a method for confirming generalizations that, from a scientific perspective, is highly unsophisticated" (Bishop and Trout 2002, 204).

Being concerned with explanatory understanding, Bishop and Trout even state that attempts to draw conclusions from particular scientific case studies can be epistemically deceptive and thus

detrimental rather than conducive to a more comprehensive picture of understanding in the sciences. While such examinations may spark the impression of having "grasped some deep truth about the nature of how science operates" (Bishop and Trout 2002, 204), it would be "[...] a serious mistake to suppose that such trappings of subjective judgment are a reliable sign of genuine understanding" (ibid.).

3.3 Pitt's Heracliteian outlook

Let us take stock. Pitt discusses two ways in which case studies may potentially figure in philosophical studies of science. Either the idea is to use case studies as a means for 'testing' certain philosophical claims or theories about science; or the idea is to use case studies to develop new philosophical ideas about science in the first place. In either way we run into problems. When using case studies as a 'tribunal' for our philosophical ideas, it is not clear how this can be done without circularity, or without risking that the selection or interpretation of case studies is biased. Alternatively, when trying to use case studies as a starting point for philosophical studies, it is not clear how generalizations from particular cases are supposed to work. In conclusion, we are left with the dilemma.

In a Heracliteian spirit, Pitt states that "[a]s philosophers we seek universals, but the only universal regarding science is change" (Pitt 2001, 374): using particular case studies as tests for philosophical claims seems prone to bias, and given the heterogeneity of science, it appears that we are not entitled to draw philosophically relevant conclusions from a few particular examples. In conclusion, case studies can, according to Pitt, at best serve as heuristics (Pitt 2001, 373).

Before developing my response to Pitt's dilemma, I will briefly introduce an example from molecular biology that concerns the surprising molecular transport selectivity of certain channel proteins in human cells. The example has sparked a lively debate in the sciences, and, for the remainder of the paper, it offers an opportunity to explore as to how philosophical conclusions from case studies may be drawn.

4 An example from molecular biology

Aquaporins are fascinating membrane proteins that play a crucial role for a plethora of electrochemical processes in the human body (de Groot et al. 2003, Hub et al. 2009). They allow for a rapid flow of water molecules through cell membranes, but strictly prevent protons from passing (de Groot et al. 2003, 279). On the one hand, maintaining proton gradients across membranes is critical for the synthesis of ATP, among other things (de Groot et al. 2003, 279). On the other hand,

the capacity of aquaporins to exclude protons is surprising, because protons are normally known to easily jump along the hydrogen bond networks of water molecules through a mechanism called Grotthuss mechanism (Yarnell 2004; de Groot et al. 2003).

In the molecular biology community, the remarkable selectivity of aquaporins has led to discussions, and various scientific groups using different methods and tools investigated how exactly aquaporins manage to perform their biological function (de Groot and Grubmüller 2005; Yarnell 2004).

Proteins consist of long chains of amino acids that fold up into specific three-dimensional structures. The three-dimensional shape of a protein is decisive for its function, and when being interested in explaining protein-involving biological processes, an important and typical task is to look at characteristic structural regions and examine their role for the respective function of the protein. In the case of aquaporins, an early idea was that a certain structural region near the center of the 'pore' which these proteins create causes the Grotthuss mechanism to be interrupted (Yarnell 2004). More precisely, scientists speculated that water molecules passing the pore constriction would form hydrogen bonds with a certain structural element of the protein, thereby being hindered from forming such bonds with adjacent water molecules and not allowing protons to hop along them (Murata et al. 2000, 604).

Disagreeing with this idea, other scientific groups using computer simulations argued that proton blockage is due to an electrostatic barrier created by a certain structural region near the pore center, rather than due to an interruption of hydrogen bond networks (de Groot et al. 2003).

Explicitly modeling proton transfer reactions and quantitatively comparing the energetic contributions of different structural regions, one group, for example, used computer simulations to trace how the energy of the system changes as a simulated water molecule passes through the aquaporin pore (de Groot et al. 2003). Surprisingly, they concluded that the re-orientation of water molecules at the pore constriction could only be a secondary factor at best (de Groot et al. 2003, 286). The reason is that the structural element which was previously held responsible for interrupting the Grotthuss-mechanism turned out to matter far less, energetically speaking, than originally assumed (de Groot et al. 2003, 286). Rather, the scientists found another electrostatic barrier with a comparably higher magnitude occurring near the center of the pore – at a point where another typical structural region is present. Thus, rather than interrupted bond networks, electrostatic forces induced by the presence of this other structural motif were identified as the main determining factor of why protons are excluded from passing through the pore (de Groot et al. 2003, 279).

Still, the discussion did not stop here. While agreeing with the finding that an electrostatic

barrier exists near the pore center, Burykin and Warshel (2003, 2004) opposed the idea of a certain motif being the main determinant for this barrier. Rather, they argued that the electrostatic barrier for proton transport were mainly due to a general loss of solvation energy occurring when a proton gets desolvated (Burykin and Warshel 2004, 43) and "pushed into an oily protein pore" (Yarnell 2004). Using molecular dynamics simulations, they artificially switched off the contributions of the structural motif that was previously assumed to be responsible for the electrostatic barrier. The reasoning behind this manipulation was that if the motif really had been a determining factor, removing it should show an effect regarding the occurrence of the behavior to be explained: "[...] in order for a given motif to be the reason for the barrier, it is essential that the removal of this motif will eliminate the barrier" (Burykin and Warshel 2004, 44). Given that the barrier remained largely unaffected, however, the motif could, according to Burykin and Warshel, not be responsible for proton blockage in aquaporins.

Again, Burykin's and Warshel's contribution was not the last word on the subject matter, and the discourse on aquaporin proton selectivity continued through a series of original research papers, reviews and overview articles (see e.g. Gonen and Walz 2006; Fu and Lu 2007). However, given space constraints, and given that the interest here lies less in tracing the history of the case and more in getting an idea of how scientists presented their arguments on the explanatory question at stake, I will end the brief reconstruction at this point and return to the philosophical discussion of Pitt's dilemma.

5 How case studies can figure in philosophical reasoning

Pitt's dilemma underscores several potential pitfalls associated with the use of case studies, including the risk of sample bias, the questionable capacity to support normative claims about science, and the failure to support generalizations. However, through a discussion of the aquaporin case study, the remainder of the paper aims to counter these criticisms and indicate that case studies can play a valid role in philosophical examinations of science.

I proceed in two steps. First, while acknowledging the diversity of scientific practices, I posit that recurring patterns within scientific practice plausibly exist that can yield philosophically relevant insights. This perspective challenges the idea that given the heterogeneity of science, particular case studies cannot lead to any significant conclusions.

Second, the paper contends that Pitt's dilemma stems from a misconceived perception of case studies as pieces of inductive evidence used to obtain or test generalizations. Drawing on a contribution by Chang and exemplifying the idea underlying his approach through the aquaporin ex-

ample, I propose a different view: case studies are not primarily pieces of evidence for inductive generalizations but can serve as resources for abstraction.[4]

5.1 Heterogeneity and patterns

To respond to the problem of heterogeneity, consider, as a starting point, Burian's (2001) approach to overcoming Pitt's dilemma. Burian argues that the dilemma is a false one because science is far less subject to 'Heracleitean flux' (Burian 2001, 398) than Pitt suggests. Even if change and heterogeneity are inherent to science, particular scientific investigations do not take place in a vacuum, or in isolation from each other. Rather, Burian suggests broadening one's lens and looking at scientific contributions from various groups or at different times.

One of Burian's ideas seems to be that even if, in a given situation, scientists start from diverging theoretical assumptions or adopt different methodologies for addressing a certain subject matter, there is usually some form of exchange and dialogue that can serve as a starting point for philosophical analysis. Concretely, given the fact that science is, to a significant extent, based on publications and concerned with broader problems within certain communities, it is usually possible to recover an account of 'larger issues' at stake (Burian 2001, 398).

As part of the emphasis on the wider context within which scientific problems are addressed, Burian suggests a broader philosophical treatment of 'case studies' (Burian 2001, 387). Rather than focusing on one particular scientific paper or an individual scientist's reasoning, a more promising approach is to trace a scientific approach or problem across different points in time, across different groups of scientists, or even across different disciplines and methodologies.[5] Concretely, Burian suggests that philosophers of science could "construct [...] studies that follow the evolution of the problem and of scientists' ways of dealing with it", or "set up comparative studies of approaches to a problem taken by workers in different laboratories or disciplines or by use of different tools and technologies" (Burian 2001, 387).

Even if the risk of sample bias or hasty generalizations cannot be escaped from entirely, giving up on universal theories and looking at the broader context in which a particular problem is being

[4] Even if the focus is on the role of case studies for abstraction here, I do not want to claim that this is the only alternative way in which case studies can figure in philosophical examinations of science. For further suggestions, see e.g. Currie (2015).

[5] As an illustration of a longitudinal study, Burian discusses the case of scientists in the 20th century who dealt with the localization and distribution of nucleic acids (Burian 2001, 389ff.). By paying close attention to the techniques and approaches used by these scientists, Burian argues that they reveal the prevalence of a particular mode of experimentation which he denotes as exploratory experimentation, and which, as he argues, would have remained unrecognized by what he considers to be 'top-down' approaches in the philosophy of science (Burian 2001, 391f.). Interestingly, while exploratory experimentation, according to Burian, turned out to be a quite ubiquitous mode of experimentation in various scientific disciplines, the conclusions drawn from case studies are certainly not straightforward generalizations. In fact, I think that Burian's example already shows that the status of the conclusions drawn from case seems quite different from what Pitt's framework suggests; and that recognizing this will be a further step in dissolving the dilemma in the next section.

addressed makes one less vulnerable to either threat.[6]

A similar direction is proposed by Currie (2015) who develops the idea that albeit being heterogeneous, the heterogeneity of science is still somewhat patterned; and that science displays a "patchy unity" (Currie 2015, 554). Like Burian, Currie suggests that it would be a mistake to treat particular scientific papers as 'stand-alone pieces' or to regard single papers or lines of reasoning as individual pieces of evidence for a certain philosophical position (Currie 2015, 562). Rather, Currie, too, stresses that even though there is a lot of 'wiggle room' for the contingent details of particular scientific approaches, this does not mean that there are no patterns in scientific practice (Currie 2015, 562).

I agree, particularly because scientific work is usually based on publications that cite and refer to one another, taking the contributions of others as a starting point for one's own position – or for distancing oneself from them.

To illustrate what it means to go beyond treating individual scientific contributions as stand-alone pieces, and to identify patterns in scientific discourse, let us return to the example of aquaporin selectivity. What was at stake in the case of aquaporins was an explanatory problem, namely as to why aquaporins keep protons out, while at the same time being highly permeable for water molecules. I suggest that the example of aquaporins allows us to obtain philosophically interesting conclusions about the topic of explanation – without having to start with strong ex ante commitments regarding how scientific explanations work – or with respect to what explanatory success means – and also without simply having to accept as to what some particular scientific group regards as a successful explanation.

Recall that as concerns aquaporin behavior, there was no consensus regarding why aquaporins exclude protons. While some scientists proposed that an interruption of the (Grotthuss) proton-hopping mechanism due to the presence of a certain structural motif might be the main reason for why aquaporins would not let protons pass, others suggested that an energetic barrier stemming from a different characteristic structural region of the protein were the main determining factor. Still other scientists argued that general solvation effects rather than the proposed structural region would be responsible for the observed energetic barrier that prevents protons from travelling through aquaporin pores. Considering that the different scientific groups not only arrived at different conclusions but also used different methods and models to reach their results, it is arguably difficult to provide any kind of quick answer regarding what counts as a successful explanation here. Luckily, this is not required to learn something philosophically interesting from the example.

[6] In Burian's words: "methodologically and epistemologically useful case studies need not be philosophically innocent and need not proceed to grand conclusions by induction from absurdly small samples." (Burian 2001, 388)

Even though scientists disagreed on what exactly explains the selective permeability of aquaporins, the ways in which results were defended and discussed show an interesting pattern: in all cases, explanatory efforts seem to have been concerned with what is commonly called structure-function relationships. Concretely, explanatory efforts involved discussing the impact (or lacking impact) of certain structural regions, and an evaluation of the extent to which the explanandum is sensitive to their presence[7]. Even if there was disagreement regarding the relevance of respectively different structural regions, this disagreement could only emerge against a shared background of assumptions, part of which is that explaining the selectivity of aquaporins involves examining the role of various structural regions to whose presence the explanandum is possibly sensitive.

Already at this point, a philosophically interesting observation can be obtained. If explaining aquaporin selectivity is a matter of assessing the impact of various structural regions of the protein, then this shows and rationalizes how computer simulations could act as particularly important tools for explanation. By allowing it to artificially switch the effects of a certain structural region on and off, computer simulations could help trace in detail as to how the presence of a particular structural protein region affects the explanandum. Thus, by emphasizing that explanatory efforts were directed at exploring the impact of various structural regions, we can appreciate how computer simulations could make important contributions to the sought explanation.

5.2 *What* can we learn from case studies?

A first step in responding to Pitt's dilemma was not to treat individual scientists' (or scientific groups') contributions as if they occurred in isolation. Instead, it was proposed that despite heterogeneity in methods used or in conclusions drawn, patterns can be identified. Even if scientists disagree about the phenomenon to be explained, and even if they employ different methods and models for their respective investigations, there plausibly exist certain communalities that create a basis for mutual exchange, and that make room for both agreement and disagreement in the first place.

Yet even though the identification of patterns alone already yields interesting findings, e.g. regarding the role of computer simulations for explaining aquaporin selectivity, such findings are still very 'local'. As it has been stated earlier, philosophers of science typically aim to obtain broader insights, i.e., insights that go beyond a particular scientific problem setting, or a particular scientific approach.

A response to Pitt's dilemma should therefore entail a closer look at the kinds of conclusions that

[7] Recall that in some of the simulation studies, such as that by Burykin and Warshel, such a 'sensitivity analysis' was performed by explicitly switching the contributions of certain structural regions on and off.

philosophical engagement with case studies can yield. In what follows, I show how conclusions can be obtained without being too narrowly restricted in scope, and while, at the same time, avoiding the problem of hasty generalizations. In line with Chang (2012), Currie (2015), and with what I take to be a motivating idea of Woody's account (2004), I suggest that the most important step is to question the idea that case studies should act as pieces of evidence for general statements. More precisely, we should question that the main purpose of case studies is to provide 'empirical data' that serves as an inductive basis for generalization.

The remainder of the paper suggests an alternative view by drawing on an idea endorsed by Chang. Looking at the use of historical examples in philosophy, Chang suggests regarding philosophical engagement with case studies less as a movement from the particular to the general, but more as a matter of moving from the concrete to the abstract (Chang 2012, 110).

As an illustration, Chang reflects on his philosophical engagement with the problem of temperature measurement (as in Chang, 2004). Inquiring the history of temperature measurement, Chang noted that historically, scientists faced a problem of circularity when trying to construct trustworthy measurement devices for temperature (Chang 2012, 112f.).

Basically, the problem was that in order to know how the expansion of some fluid could serve as a trustworthy vehicle to measure temperature, it was necessary to already know how temperature changes relate to the expansion of this fluid. Having identified this struggle, Chang extracts the following more abstract problem scheme which he calls the "nomic problem of measurement" (Chang 2012, 113):

> *(i) We want to measure quantity X;*
> *(ii) If quantity X is not directly observable, we infer it from another quantity Y, which is directly observable.*
> *(iii) For this inference we need a law that expresses X as a function of Y.*
> *(iv) But the form of this function cannot be discovered or tested empirically, because that would involve knowing the values of both Y and X, but X is the unknown variable that we are trying to measure.* (Chang 2012, 113)

Now interestingly, as Chang points out, it is plausible to assume that the identified problem scheme applies to many other contexts in which the measurement of an unknown quantity is at stake (Chang 2012, 113). That is, by looking at the historical case of temperature measurement, one can learn about a more abstract scheme that applies to other, similarly structured problems.

Furthermore, having carefully traced the history of the development of thermometers, Chang found that scientists dealt with the problem of nomic measurement not by starting form a firm

ground of evidence, but in a process of stepwise improvement and refinement. Chang suggests conceptualizing this mode of scientific work as "epistemic iteration" (2012, 115).

I think that the main essence of Chang's approach is that conclusions are drawn from the temperature example not by assessing as to how particular details of the temperature case resemble other cases of measurement and, based on that, by inferring more general claims; but by recognizing how the example gives a concrete instance[8] of a more abstract problem scheme that plausibly applies to other situations with a similar problem setting. Furthermore, based on his examination, Chang proposed a concept – that of epistemic iteration – for capturing scientists' way to deal with this type of problem[9].

Chang's idea of how case studies can help moving from the concrete to the abstract resonates with the case of aquaporins. Recall that scientists undertook, in a number of publications, efforts to explain the capacity of aquaporins to exclude protons. Rather than committing to the success of one particular explanation, we could find a few patterns in the presented lines of reasoning. For example, all of the discussed contributions were concerned with how the biological function of the protein depends on its structure; and all considered contributions dealt with an assessment of how various structural regions of the channel protein matter for its transport selectivity. Furthermore, computer simulations helped to systematically test how the presence or absence of certain structural regions affects the explanandum.

Like in Chang's example, we can use these observations to try identify the more abstract scheme of which these reasoning patterns are concretizations, leading to something along the following lines:

- In the considered example, obtaining an explanation involves examining how various factors potentially make a difference for the explanandum.
- In the considered example, searching for an explanation involves an examination of how the explanandum would have been different if certain factors (such as the energetic contributions of certain structural regions) had been different.
- What is at stake is to how the function of the protein depends on the energetic contributions of various component parts.

[8] I take Woody's conceptualization regarding the instantiation of what she calls 'explanatory structures' to underscore this direction (Woody 2004, 39ff.). Woody's term of 'explanatory structures' resonates with what I have earlier called a 'pattern', and similar to the examination presented here, Woody points out that starting from patterns or structures prevalent in certain scientific practices instead of starting from individual supposedly successful explanations makes one less vulnerable to the threat of drawing hasty inductive conclusions (cf. Woody 2004: 39).

[9] Chang suggest that just as in Burian's case (see footnote 5), such a concept would hardly have been obtainable from a standard philosophical approach to science (Chang 2012, 113).

The obtained scheme appears, of course, familiar, and is commonly captured under the header of a "counterfactual" or "difference-making" account of scientific explanation in the spirit of Woodward (2003). Furthermore, the idea that the sought explanation is somehow 'compositional' since it gives an account of how the explanandum depends on various components of the protein resonates with the idea of mechanistic explanation (see e.g. Bechtel and Abrahamsen 2005; Illari and Williamson 2012).

Of course, stating that the case of aquaporins can be regarded as an instance of a certain more abstract scheme is still an inductive and fallible move: it is certainly possible that the selection of scientific contributions is still biased, and when identifying patterns across these contributions as concretizations of a more abstract reasoning scheme, we can be mistaken. Notwithstanding, I contend that we are somewhat less vulnerable to the problem of drawing hasty conclusions, because particular scientific contributions do not serve a "confrontational role" regarding our philosophical reasoning (Schickore 2012; Kinzel 2015), and because we draw different conclusions than in what may be called an 'evidential support picture'. Rather than using our case study to answer, for example, what the nature of scientific explanation is; or how explanations generally work, we draw inferences about an abstract explanatory reasoning scheme.

The question of whether and to what extent this scheme occurs more widely in the sciences is a different question. My point is that in the context of aquaporins, the way in which scientists present their arguments resonates with the more abstract scheme captured by the counterfactual account of explanation, and recognizing that the search for an explanation involved testing how the explanandum is sensitive to a range of factors allows for a better understanding of the respective explanatory efforts (or when they go wrong), as well as for a more comprehensive view of the explanatory role of computer simulations (cf. Schweer and Elstner 2023). Rather than moving from particular explanations to an overarching theory about the requirements for successful explanation, we move from concrete patterns to more abstract schemes of explanatory reasoning.

Even though this approach does not give us generalizations in the sense of "all scientific explanations are counterfactual explanations", we can, of course, ask if the counterfactual scheme might reasonably be assumed to apply more widely in the sciences. Just as we have examined how the explanatory argumentation of scientists dealing with aquaporin behavior resonates with a more abstract reasoning scheme, we could ask if the problem of aquaporin selectivity itself has characteristics that can be recognized as instances of a more abstract class of explanatory problems. It seems, for example, that the case of aquaporins is a case in which the behavior of a complex system is at stake, i.e., a system with many (molecular) components that interact in complicated ways (cf. Schweer and Elstner 2023). Furthermore, rather than focusing on a single occurrence of proton

exclusion, explanatory efforts were directed towards understanding proton exclusion as a stable form of molecular behavior, i.e., a behavior that occurs frequently enough to play an important regulatory role in the biological functions of cells.

These more abstract aspects – that we are dealing with a complex system and with the occurrence of a stable pattern of molecular behavior – arguably apply to many other explanatory investigations in the sciences. When dealing with a system involving a plethora of components which could potentially play a role for why the system shows the behavior that it does; and when being interested in the occurrence of a stable form of molecular behavior, we may expect that we will be interested in finding dependence relations that hold under a range of variations in the examined systems. This does by no means prove that the counterfactual account of explanation applies generally, but it emphasizes that rather than being restricted to the particular problem under investigation, the aquaporin example offers lessons that are potentially relevant for other scientific situations with similar characteristics.

5.3 Normativity

Before I conclude, I briefly come back to the problem of normativity. It has earlier been suggested that the trouble with case studies – as sketched in Pitt's dilemma – is also a trouble with normativity, or with how philosophy of science can make evaluative judgements if it starts from how scientists actually do science. Relatedly, the problem was that if case studies are assumed to support a more general theory of explanation or explanatory success, then on what grounds can we know that the respectively cited case studies indeed reflect instances of successful explanations?

I suggest that once the idea is dismissed that case studies primarily act as pieces of evidence for philosophical generalizations, the issue of normativity can better be addressed. In the case of aquaporins, the suggestion was that – despite there being disagreement between the different scientific groups involved – it is possible to find common patterns in scientists' explanatory reasoning. Further, these patterns can be regarded as instances of a more abstract scheme of reasoning about difference-making factors. The individual candidate explanations given in particular papers or by particular scientific groups do not serve as evidence for somehow proving the adequacy of a more general theory of explanation; and to characterize the patterns of reasoning that can be found across various scientific contributions, we do not need to commit to the success of one particular candidate explanation.

At the same time, once we arrive at the idea that scientists' reasoning in the context of aquaporins adheres to a sort of 'difference-making scheme', it certainly becomes possible to make evalua-

tive judgements regarding the success of individual candidate explanations. In the context of the discussed case of aquaporins, for example, it seems that explanations are more or less successful depending on how they succeed to depict a number of factors to which the explanandum is sensitive; and if a certain paper shows, for example, that some other contribution failed to recognize a factor that matters for the occurrence of the explanandum, it seems legitimate to conclude that this other contribution fails to provide a satisfactory explanation.

This implies that even if no strong commitments need to be made regarding the success of particular candidate explanations, once a reasoning pattern has been identified that occurs across numerous contributions, one can formulate criteria for the success of individual candidate explanations. This may seem circular at first glance, but the idea is to evaluate scientists' explanatory success with the help of norms that are implicit to their own reasoning, as it is exerted across different scientific contributions. This does, at least, not mean that we must commit beforehand to normative judgements about the success of particular candidate explanations (see also Woody 2004, 39-40).

A thorough discussion of the role of normativity in the philosophy of science in practice is offered by Kaiser (2019). Kaiser argues that despite aiming for a reconstruction of how scientists actually do science, practice-oriented philosophy does not need (and cannot) move away from normativity entirely (Kaiser 2019, 49f.). Rather, doing philosophy of science in practice means moving away from a particular kind of normativity, namely "ex-cathedra metanormativity" (Kaiser 2019, 42f.). According to Kaiser, "[a] philosophical theory T about a feature or element of science E is metanormative iff T contains normative claims about E" (Kaiser 2019, 41).

'Ex-cathedra metanormativity' means that normative claims are being made regardless of how scientists themselves would conceptualize their aims and achievements (Kaiser 2019, 43) – which is obviously at odds with what philosophy of science in practice seems to be all concerned with. As an alternative, Kaiser suggests that besides involving methodological normativity and object normativity[10], philosophy of science can still be metanormative "if it contains only such normative claims about E that take into account, are drawn from, or are informed by factual claims about E" (Kaiser 2019, 44).

Kaiser only vaguely indicates as to what 'obtaining normative claims from factual statements' without running into the naturalistic fallacy might look like, but the approach suggested in this paper and the case of aquaporins may offer an answer. What makes a successful explanation is neither

[10]By object normativity, Kaiser refers to normativity that stems from the engagement with normative topics in the sciences, e.g. the role of values in science (Kaiser 2019, 51). By methodological normativity, Kaiser means the form of normativity that arises when issues such as the selection or interpretation of scientific examples are at stake. As Kaiser convincingly argues, methodological normativity cannot be escaped from (Kaiser 2019, 44ff.).

to be determined through a pre-existing 'ex-cathedra theory' of explanation; nor through simply accepting as to what particular scientists regard as a successful explanation of the phenomenon in question. Rather, it is the recognition of patterns in a particular scientific practice that can eventually inform us as to what the success conditions of explanations are in a certain problem context. Scientists participating in explanatory practices are usually engaging in discourse with others; and even if they disagree with one another, the structure of their exchange can reveal shared norms of reasoning that allow to infer regional standards for explanatory success.

6 Conclusion

Philosophy of science is becoming increasingly practice-oriented, but concerns have been raised regarding the justified use of case studies. In this paper, I have examined how case studies can figure in philosophical reasoning without leading to hasty generalizations or begging the question.

It is increasingly acknowledged that science is heterogeneous. At the same time, concerns have been expressed that this heterogeneity raises doubts whether particular case studies can yield any philosophically interesting insights at all. In a first step, responding to these concerns, I have, drawing on Burian and Currie, proposed that individual contributions by scientific groups should not be regarded as if they occurred in isolation, or in a vacuum. Using the example of aquaporin selectivity, it has been illustrated that even though scientific practice displays a diversity of methods and approaches, tracing lines of scientific reasoning across various individual contributions may allow one to recognize shared patterns of reasoning. These patterns can enable philosophically relevant insights, e.g., regarding the explanatory role of computer simulations.

In a second step, drawing on an approach by Chang, I have proposed that rather than regarding case studies as pieces of evidence for generalizations, an alternative way in which they can play a valid and relevant role in philosophical studies is as resources for abstraction. Even though the selection and philosophical interpretation of case studies certainly remains fallible, giving up on the idea that the main role of case studies is to provide evidence for generalizations means becoming less vulnerable to the problem of drawing big conclusions from a few examples.

As concerns normativity, the challenge for practice-oriented philosophy of science dealing with scientific explanations was to find a middle ground: to neither impose pre-existing ideas of how explanations work in science, nor to simply having to take for granted as to what individual scientists regard as an explanation. I have indicated a way in which philosophy of science can involve normative claims about explanations, while at the same time being sensitive to how scientists actually give explanations.

An important and natural consequence of the suggested alternative view, according to which the focus in on examining how case studies are concretizations of more abstract problems or reasoning schemes, is that the philosophical attention moves away from questions such as "what is the nature of scientific explanations?" or "how do explanations work generally?". Rather, the focus shifts towards an exploration of more 'regional' patterns of explanatory reasoning; and towards a recognition of the various reasoning schemes prevalent in the sciences, as well as to an understanding of the tools which are used in explanation-seeking contexts. While the suggested perspective on case studies may certainly be bad news for those interested in universal philosophical theories of scientific explanation, it hopefully contributes to showing a path for a philosophy of science that is both closely informed by scientific practice and avoids over-generalization. Lastly, the presented examination also underscores how scientific disagreements can be of particular interest to philosophical studies of scientific practice, as they make apparent as to what kind of things scientists agree on (or implicitly presuppose) even if they disagree on the subject matter they are concerned with.

Bibliography

Ankeny, R., Chang, H., Boumans, M., and Boon, M. (2011). Introduction: Philosophy of science in practice. *European Journal for Philosophy of Science*, 1:303–307.

Bechtel, W. and Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2):421–441.

Bishop, M. A. and Trout, J. (2002). 50 years of successful predictive modeling should be enough: Lessons for philosophy of science. *Philosophy of Science*, 69(S3):S197–S208.

Bokulich, A. (2011). How scientific models can explain. *Synthese*, 180:33–45.

Bokulich, A. (2014). How the tiger bush got its stripes: 'how possibly' vs. 'how actually' model explanations. *The Monist*, 97(3):321–338.

Burian, R. M. (2001). The dilemma of case studies resolved: The virtues of using case studies in the history and philosophy of science. *Perspectives on Science*, 9(4):383–404.

Bursten, J. R. (2020). Lab report: Lessons from a multi-year collaboration between nanoscience and philosophy of science. In *A Guide to Field Philosophy*, pages 35–47. Routledge.

Burykin, A. and Warshel, A. (2003). What really prevents proton transport through aquaporin? charge self-energy versus proton wire proposals. *Biophysical Journal*, 85(6):3696–3706.

Burykin, A. and Warshel, A. (2004). On the origin of the electrostatic barrier for proton transport in aquaporin. *FEBS Letters*, 570(1-3):41–46.

Chang, H. (2011). The philosophical grammar of scientific practice. *International Studies in the Philosophy of Science*, 25(3):205–221.

Chang, H. (2012). Beyond case-studies: History as philosophy. In Mauskopf, S. and Schmaltz, T., editors, *Integrating history and philosophy of science: Problems and prospects*, pages 109–124. Springer.

Currie, A. (2015). Philosophy of science and the curse of the case study. In Daly, C., editor, *The Palgrave Handbook of Philosophical Methods*, pages 553–572. Springer.

de Groot, B. L., Frigato, T., Helms, V., and Grubmüller, H. (2003). The mechanism of proton exclusion in the aquaporin-1 water channel. *Journal of Molecular Biology*, 333(2):279–293.

de Groot, B. L. and Grubmüller, H. (2005). The dynamics and energetics of water permeation and proton exclusion in aquaporins. *Current Opinion in Structural Biology*, 15(2):176–183.

Franklin, A. (1989). *The Neglect of Experiment*. Cambridge University Press.

Fu, D. and Lu, M. (2007). The structural basis of water permeation and proton exclusion in aquaporins. *Molecular Membrane Biology*, 24(5-6):366–374.

Galison, P. (1987). *How experiments end*, volume 88. University of Chicago Press Chicago.

Galison, P. (2008). Ten problems in history and philosophy of science. *Isis*, 99(1):111–124.

Gelfert, A. (2016). *How to do Science with Models: A Philosophical Primer*. Springer.

Giere, R. N. (2012). History and philosophy of science: Thirty-five years later. In Mauskopf, S. and Schmaltz, T., editors, *Integrating History and Philosophy of Science: Problems and Prospects*, pages 59–65. Springer.

Gonen, T. and Walz, T. (2006). The structure of aquaporins. *Quarterly Reviews of Biophysics*, 39(4):361–396.

Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge University Press.

Hub, J. S., Grubmüller, H., and De Groot, B. L. (2009). Dynamics and Energetics of Permeation through Aquaporins. What Do We Learn from Molecular Dynamics Simulations? In *Aquaporins*, pages 57–76. Springer.

Illari, P. M. and Williamson, J. (2012). What is a mechanism? thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2:119–135.

Kaiser, M. I. (2015). *Reductive Explanation in the Biological Sciences*. Springer.

Kaiser, M. I. (2019). Normativity in the philosophy of science. *Metaphilosophy*, 50(1-2):36–62.

Kinzel, K. (2015). Narrative and evidence. how can case studies from the history of science support claims in the philosophy of science? *Studies in History and Philosophy of Science Part A*, 49:48–57.

Knorr-Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

Love, A. C. (2015). Collaborative explanation, explanatory roles, and scientific explaining in practice. *Studies in History and Philosophy of Science Part A*, 52:88–94.

Mission statement of the SPSP (2023). Mission statement of the society for philosophy of science in practice. http://philosophy-science-practice.org/about/mission-statement/ (URL date 16.05.2023).

Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J. B., Engel, A., and Fujiyoshi, Y. (2000). Structural determinants of water permeation through aquaporin-1. *Nature*, 407(6804):599–605.

Osbeck, L. M. and Nersessian, N. J. (2015). Prolegomena to an empirical philosophy of science. In Wagenknecht, S., Nersessian, N., and Andersen, H., editors, *Empirical Philosophy of Science: Introducing Qualitative Methods into Philosophy of Science*, pages 13–35. Springer.

Pitt, J. C. (2001). The dilemma of case studies: Toward a heraclitian philosophy of science. *Perspectives on Science*, 9(4):373–382.

Schickore, J. (2012). Scientists' methods accounts: S. Weir Mitchell's research on the venom of poisonous snakes. In Mauskopf, S. and Schmaltz, T., editors, *Integrating History and Philosophy of Science: Problems and Prospects*, pages 141–161. Springer.

Schweer, J. and Elstner, M. (2023). Dealing with molecular complexity. atomistic computer simulations and scientific explanation. *Perspectives on Science*, 31(5):594–626.

Soler, L., Zwart, S., Lynch, M., and Israel-Jost, V. (2014). *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*. Routledge.

Vos, B. (2021). Integrated HPS? formal versus historical approaches to philosophy of science. *Synthese*, 199(5-6):14509–14533.

Wittgenstein, L. (1953). *Philosophical Investigations*. Wiley-Blackwell, New York, NY, USA.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Woody, A. (2014). Chemistry's periodic law: Rethinking representation and explanation after the turn to practice. In Soler, L., Zwart, S., Lynch, M., and Israel-Jost, V., editors, *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*, pages 123–150. Routledge.

Woody, A. I. (2004). Telltale signs: What common explanatory strategies in chemistry reveal about explanation itself. *Foundations of Chemistry*, 6(1):13–43.

Woody, A. I. (2015). Re-orienting discussions of scientific explanation: A functional perspective. *Studies in History and Philosophy of Science Part A*, 52:79–87.

Yarnell, A. (2004). Blockade in the Cell's Waterway. *Chem. Eng. News*, 82(4):42–44.

Dealing with Molecular Complexity: Atomistic Computer Simulations and Scientific Explanation

Contents

Summary

Explanation is commonly considered one of the central goals of science. Although computer simulations have become an important tool in many scientific areas, various philosophical concerns indicate that their explanatory power requires further scrutiny. We examine a case study in which atomistic simulations have been used to examine the factors responsible for the transport selectivity of certain channel proteins located at cell membranes. By elucidating how precisely atomistic simulations helped scientists draw inferences about the molecular system under investigation, we respond to some concerns regarding their explanatory power. We argue that

atomistic simulations can be tools for managing molecular complexity and for systematically assessing how the occurrence of the explanandum is sensitive to a range of factors.

1 Introduction

Computer simulations in a wide range of scientific fields have become a relevant (and often critical) means for studying the characteristics of complex systems under controlled conditions since they allow overcoming previous boundaries of analytical tractability or experimental accessibility. Besides offering the opportunity to make powerful predictions or explore and describe the properties of highly complex systems, computer simulations sometimes also play a crucial role in contexts in which the aim is to better explain or understand a given phenomenon.

As an example where computer simulations are seen as means to address explanatory questions, consider a case study from molecular biology, more specifically structural biology, that concerns the properties of certain channel proteins present at the membranes of biological cells; so-called "aquaporins." One remarkable property of aquaporins is that while they allow water molecules to quickly permeate cell membranes, at the same time they strictly prevent protons from passing. This selectivity is electrochemically important and yet surprising: usually, protons are known to be easily conducted by water molecules along hydrogen bond networks (Ko et al. 2008, p. 1442). At the same time, maintaining proton gradients across the membranes of biological cells is crucial inter alia for ATP synthesis - and thus pivotal for energy generation (de Groot et al. 2003, p. 279).

Although the characteristics of aquaporins have been extensively studied within recent years, the detailed mechanism of proton exclusion has long remained an open question (see, e.g., Yarnell 2004; de Groot and Grubmüller 2005; Ko et al. 2008). As several research groups report, it is precisely due to a lack of direct experimental access to the dynamics of water molecules and protons at the channel that so-called atomistic simulations, which allow taking into account inter- and intramolecular interactions, became so relevant for tackling the issue of aquaporin selectivity (de Groot and Grubmüller 2005, p. 179; see also Hub et al. 2009).

De Groot and Grubmüller (2005, p. 176) sketch the role of computer simulations in the context of research on aquaporins as follows:

> How efficient water permeation is achieved and how it is reconciled with the seemingly contradictory task of strict proton exclusion have been long-standing puzzles. Because neither the dynamics of the water molecules nor the mobility of protons inside the aquaporin channel could be experimentally accessed so far, several groups addressed this challenge using a variety of atomistic computer simulation methods.

However, although statements such as the aforementioned clearly suggest that simulation methods play a relevant role in explaining the phenomenon under investigation, the explanatory potential of computer simulations in general, and of atomistic simulations in particular, has, at the same time, not remained uncontested within philosophical debates. Discussing the case of an atomistic simulation study used to examine the characteristics of gold at the nanoscale, Lenhard, for example, raises the concern that "simulations squeeze out the consequences in an often unintelligible and opaque way" (2006, p. 612), and that they do not provide an "explanation in the usual sense" (2006, p. 609). The underlying worry seems to be that while simulations often successfully produce a certain behavior or particular phenomenon, at the same time they leave unclear how this phenomenon follows from the many details and from the modelling assumptions present in the simulation process.

Moreover, as Imbert points out, the widely discussed problem of explanatorily irrelevant details seems "especially acute" (2017, p. 754) in the case of computer simulations:

> Take a computer simulation that unfolds the detailed evolution of a system based on the description of its initial state and the laws governing it. Then all aspects of the computational model are actually used in the computational derivation. At the same time, all such aspects are not necessarily explanatorily relevant with respect to all facets of the computed behavior. (Imbert 2017, p. 754)

The reported difficulty may appear particularly striking in the case of atomistic simulations: Even if such simulations allow us to successfully derive the behavior of complex systems from a detailed description of the interactions between their constituents, it may seem a possible worry that it remains unclear as to why this particular behavior has emerged (or could emerge) from the underlying microscopic conditions. Put differently: Given that they are based on a description of large amounts of molecular level details, atomistic simulations may be suspected of leaving us lost in a sea of irrelevant microscopic information or of generating their results in a humanly unintelligible manner. Taking the case of computer simulations in quantum chemistry, Lenhard even states that "[c]omputational modeling faces a trade-off between explanation and prediction and enlarges predictive force at the cost of explanatory force" (2014, pp. 339-40).

Looking back to the case of aquaporins, there appears to be an interesting tension: on the one hand, the case of aquaporins shows that computer simulations are sometimes not only seen as being compatible with but even as central for the aim of obtaining an explanation. On the other hand, various worries indicate that relying on computer simulations could make it more difficult to obtain explanations or even implies a trade-off between explanatory and predictive force.

In this paper, we aim to address some of the philosophical concerns that have been expressed with respect to the explanatory potential of computer simulations. Taking the case of aquaporins as an example, we suggest that when trying to understand how computer simulations help scientists to obtain scientific explanations, we should pay close attention to concrete explanatory practices and to how precisely inferences about physical targets are drawn with the help of simulations. We outline that while the explanatory power of computer simulations will of course depend on the respective scientific context in which they are employed, the case of atomistic simulations and of aquaporins allows us to gain insight on a more general aspect that has so far not sufficiently been taken into account in philosophical debates. Rather than falling short of the capacity to distinguish between relevant and irrelevant details and thereby failing to meet common standards of scientific explanation, computer simulations can be tools to effectively manage complexity and help to systematically abstract from large amounts of explanatorily irrelevant details. In the case of aquaporins, atomistic simulations could grant access to a level of description at which the role of various factors that potentially figure in the explanation can be assessed. More specifically, they could help with the search for an explanation by enabling scientists to systematically examine how various structural contributors affect the occurrence of the explanandum.

We shall proceed as follows. In section 2, we provide an overview of recent philosophical discussions concerning the aforementioned problem of explanatory relevance. In section 3, we activate the endeavor of clarifying the explanatory role of computer simulations by outlining some of the philosophical worries that have been expressed with regard to their explanatory power. In section 4, we give a brief overview of the principles of molecular dynamics simulation which is a central type of atomistic simulation. While trying to avoid unnecessary technicality, we still consider it helpful for the purpose of our examination to offer some background about the kinds of questions these simulations usually serve to address and about how they are used to draw inferences about the properties of complex systems. Based on that, we take a look at the role of these simulations in the study case of aquaporin. In section 5, we discuss how atomistic simulations can help obtain scientific explanations. Section 6 takes a brief look at a second example in order to provide additional support for the suggested view on the explanatory role of atomistic simulations. Section 7 concludes and offers an outlook.

2 Scientific Explanation and Explanatory Relevance

A particularly prominent topic in recent philosophical literature on scientific explanation is the issue of explanatory relevance. A widespread demand is that any satisfactory account of scientific

explanation should say something about what is relevant for an explanation and what is not. In recent debates, explanatory relevance is commonly cashed out in terms of difference-making. One widely discussed account of difference-making has been proposed in the work of Woodward (2005). According to Woodward, explanations are answers to what-if-things-had-been-different questions. As he puts it: "An explanation ought to be such that it enables us to see what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways" (2005, p. 11). In this view, successfully explaining a phenomenon has to do with showing patterns of counterfactual dependence: explaining why an explanandum has occurred is to point out how the explanandum would have been different if the conditions that gave rise to it had been different (Woodward and Hitchcock 2003, p. 4). That is, only those factors for which the explanandum is a "sensitive" matter for an explanation.

Now in structural biology and related fields, explanations typically concern the way in which various components of a system and their interplay bring about a certain molecular behavior. That is, explanatory interests often seem to be concerned with what is broadly discussed under the header of "mechanistic explanation."[1] As Illari puts it: "In a very general way, finding mechanistic explanations involves finding and describing the phenomenon, and finding and describing the entities and activities, and their organization, by which the phenomenon is produced" (2013, p. 239). According to Craver (2006, p. 369), mechanistic explanations are "constitutive" in the sense that they explain the behavior of a mechanism that underlies a given phenomenon in terms of "the organized activities and interactions between its components." Relatedly, Bechtel and Abrahamsen (2005, p. 432) state that "[m]echanistic explanations [...] seek to identify component parts and operations of a mechanism." According to them, providing a mechanistic explanation of a system is essentially about decomposition into component parts and tracing how these parts and the operations they perform give rise to the explanandum (2005, p. 434).

Yet, even though mechanistic explanations seem to set a strong focus on decomposition and the organization of parts, within recent discussions it has increasingly been stressed that adding more details or decomposing a system in an as fine-grained fashion as possible does not always seem to lead to a better explanation. Levy and Bechtel, for example, express the concern that many accounts of mechanistic explanation have (implicitly or explicitly) regarded the "filling in of concrete details as a hallmark of explanatory progress" (2013, p. 258). Inspired by Strevens'

[1] We think that in the case of aquaporins, the kind of explanation that scientists seem to be looking for when examining the selective transport of molecules clearly involves an interest in a molecular mechanism. What is at stake is the way in which the structural parts of the protein as well as the interactions of these parts with water molecules at the pore give rise to a certain macromolecular function, namely, the selective transport of molecules. In this paper, however, we do not aim at discussing the scope and limits of mechanistic explanation in further detail. Instead, we focus on how one can respond to the philosophical concerns that have been expressed with regard to computer simulations' explanatory power.

account of explanation and difference-making (Strevens 2008), Levy and Bechtel stress that in order to obtain an explanation more details are not always better. Quite the contrary, they highlight that it is precisely the role of abstraction, i.e., the omission of details (Levy and Bechtel 2013, 242), to help outline those factors of a system that make a difference for its behavior: "oftentimes, omitting detail permits one to distinguish those underlying factors that matter from those that do not" (Levy and Bechtel 2013, p. 256). As they argue, abstraction helps identify a system's relevant causal organization and enables generalization (2013, p. 258).

As an illustration of why large amounts of details about the process that leads to the occurrence of an explanandum usually need to be omitted, consider an example by Robert Batterman. In his 2001 book, he discusses the example of steel struts that buckle after a critical load is placed on them. According to Batterman (2001, p. 9ff.), providing an explanation of this bending behavior cannot reasonably require us to give a full picture of the micro-constitution of particular struts and to trace the whole chain of molecular collisions that leads to their bending. If giving an explanation meant taking into account as many steps of the detailed molecular process as possible, then we could not make sense of the fact that many different struts which are composed of different materials and accordingly vary in their micro-constitution exert a similar behavior (Batterman 2001, p. 11). Put differently: to the extent to which we are interested not in explaining the bending of one particular steel strut but the behavior of steel struts in general, it seems that the central point is that they all exert a similar behavior after a critical load has been placed on them despite the details of the (causal) stories leading to their bending differ from case to case. Consequently it is, in Batterman's words, necessary to set aside most of these details as "explanatory noise" (Batterman 2001, p. 43). As he puts it:

> All of those details that may be relevant to the behavior of the particular strut don't serve to answer the question of why loaded struts in general behave the way that they do. Actually, what does the explaining is a systematic method for abstracting from these very details. (Batterman 2001, p. 13)

Coming back to the case of mechanistic explanation, Brigandt (2013, p. 480) has, for example, suggested that in order for a feature to be explanatorily relevant, a manipulation of that feature should affect the explanandum in a way such that neglecting or changing this feature has the consequence of not generating the explanandum anymore. If, for example, changing a certain component of a system has the effect that the explanandum no longer occurs, then this component should be cited in an explanation of the system. Vice versa, if changing the constituents of this component does not affect the occurrence of the explanandum, then the constituents are not explanatorily

relevant and should not be included in the explanation (Brigandt 2013, p. 488). Looking back to Bechtel's and Abrahamsen's comment on the role of abstraction, Brigandt (2013, p. 481) emphasizes that since abstraction contributes to the identification of difference-making features, it plays an important role in explanations that involve reference to mechanisms.

3 Computer Simulation and Scientific Explanation

In recent years, the question of how computer simulations help obtain scientific explanations or understanding has attracted increasing philosophical attention. While some have undertaken the issue of outlining more general characteristics of accounts of explanation for computer simulations (Weirich 2011; Durán 2017) others have investigated particular types of simulations and explored the way in which they provide explanations or enable understanding (Grüne-Yanoff 2009; Bechtel and Abrahamsen 2010; Marchionni and Ylikoski 2013; Parker 2014).

Another strand within the philosophical literature has addressed the question of whether certain characteristics of computer simulations – such as their epistemic opacity – pose a threat to the aim of generating explanations or obtaining scientific understanding (see, e.g., Humphreys 2004; Lenhard 2006; Kuorikoski 2012). According to Humphreys (2009, p. 619), "[the] computations involved in most simulations are so fast and so complex that no human or group of humans can in practice reproduce or understand the processes." In his (now widely adopted) terminology, the computational process between simulation model and output can remain epistemically opaque in the sense that it is not possible for an epistemic agent to know all epistemically relevant elements of this process (Humphreys 2009, p. 618). As he puts it, "[this] opacity can result in a loss of understanding because in most traditional static models our understanding is based upon the ability to decompose the process between model inputs and outputs into modular steps" (2004, p. 148).

A related concern has been stressed by Lenhard (2006; 2019). Considering that computer simulations often allow solving otherwise intractable models and enable scientists to study the behavior of highly complex systems in an explorative way, he suggests that they can aptly be characterized as means to circumvent a "complexity barrier." In cases in which complexity makes it unfeasible for scientists to straightforwardly extrapolate the behavior of a system from sets of known theoretical laws or assumptions, computer simulations often still provide an opportunity for control or prediction (Lenhard 2006, p. 609). At the same time, in spite of their many advantages, Lenhard states that the upcoming of simulation methods gave a new "twist" to the "epistemic opacity of nature and natural phenomena" (Lenhard 2006, p. 614). Drawing on Humphreys' conception of epistemic opacity, he emphatically states: "The models themselves, our own constructions, are epistemically

opaque!" (Lenhard 2006, p. 614). Discussing the case of an atomistic MD (Molecular Dynamics) simulation by means of which the surprising disposition of gold to form nanoscale wires under certain conditions had been predicted, Lenhard points out that despite having enabled scientists to observe surprising and unexpected behavior, the way in which the simulation reached its result stays somewhat unintelligible:

> The relation between the general Schrödinger equation and the golden wire remains opaque. Despite obviously being theory based, the simulation does not offer something like a theory-based insight into behavior. (Lenhard 2006, p. 609)

It seems that a central idea underlying both Lenhard's and Humphreys' worries regarding the opaque character of simulations is that even though the theoretical resources upon which concrete simulation models are constructed may be well established and validated, the relationship between these resources and the respective simulation output often still remains unintelligible.

Considering the case of *ab-initio* quantum chemistry, Lenhard argues that complex computational models often leave it unclear how the behavior of the model relates to the respective underlying assumptions:

> The fact that quantum chemists deal with a compound of principled theory, (autonomous) modeling steps, and adaptive loops shows that quantum theory plays a decisive role as a part of the compound, but at the same time it shows that epistemic opacity looks inevitable. It becomes difficult to attribute model performance or certain parts of model performance to certain model assumptions. (Lenhard 2014, p. 355)

Adopting Batterman's terminology, Jebeile (2018, p. 218) suggests using the term "explanatory noise" in order to refer to the opacity that concerns the connection between simulation output and the components of the underlying simulation model. According to her, there is the challenge of identifying those details of the simulation that are relevant for an explanation – and especially in cases where the behavior of complex systems is being studied – vast amounts of such details are taken into account (Jebeile 2018, p. 218). Against this backdrop, computer simulations seem inherently plagued with the problem of identifying explanatorily irrelevant information. As she puts it:

> The more one encounters explanatory noise, the more difficult it is for a cognitively unaided human to grasp relations of explanatory relevance between the inputs and the outputs of simulations. (Jebeile 2018, p. 218)

Finally, going in a similar direction, Imbert has pointed out that computer simulations are based on "informationally replete descriptions of their target systems" and thus "raise the suspicion of being computational arguments that contain many irrelevancies, and therefore of being poor explanations – even when they are not" (2017, p. 754).

4 Unveiling Molecular Mechanisms: The Puzzle of Aquaporin Selectivity

To adequately capture the role of computer simulations for scientific explanation, it is essential to look at the broader research situation in which their use is embedded. In contexts such as that of aquaporin research, simulations are typically preceded and complemented by various experimental or theoretical studies. Such studies create a point of reference and form a background of expectations which shape scientists' explanatory interests. Hence, let us now for a moment set aside the philosophical discussions pertaining to the role of computer simulations in the context of scientific explanation and become familiar with the study case of aquaporins as well as with the scientific research field within which it is placed.

As noted earlier, the regulated and rapid flow of water through cell membranes is crucial for a variety of physiological processes in the human body. While it is generally known that water can pass through lipid bilayers present at the membrane of cells, laboratory observations have indicated that the water permeability of certain tissues is significantly higher than that found for other cell types, suggesting that specific transmembrane pathways through which water can pass more rapidly play a relevant role for osmoregulation (see Borgnia et al. 1999, pp. 426–27).

And indeed, in the late 1980s, scientists discovered certain kinds of channel proteins located at the membrane of cells since known as aquaporins. By enabling a fast and highly selective transport of molecules through cell membranes, aquaporins play a crucial role in many different physiological processes ranging from protection against osmotic shock to urine concentration regulation or maintenance of brain function (Janosi and Ceccarelli 2013; Day et al. 2014). Even though much effort has been devoted to studying the characteristics of aquaporins, their capacity to let water molecules enter the cell while blocking the passage of protons has, as has been indicated, for a long time left scientists with open questions.

Exploring how the function of macromolecules (e.g., their capacity to catalyze biochemical reactions in cells or contribute to the formation of larger structures such as muscles or bones) is determined by their structural features is at the heart of many research studies belonging to a field

at the interface of physics, chemistry, and biology called structural biology. Liljas et al. characterize structural biology as "the science that tries to make the sub-cellular and molecular objects of biology visible and understandable" (2016, p. 7). As we will see, the case of aquaporin proton selectivity can be regarded as paradigmatic for the kind of research questions addressed within the field of structural biology: the aim is precisely to show how their capacity for selective molecule transport rests on circumstances that are due to certain structural features of the macromolecule as well as due to the interactions between the protons and these features. It is this broader research context that one should have in mind when assessing the explanatory role of computer simulations in the case of aquaporin selectivity.

4.1 Before Simulations: Structural Analyses

Generally, the structure of macromolecules such as aquaporins can be examined by means of crystallographic analysis or NMR (nuclear magnetic resonance) spectroscopy. A common aim of structural analyses is the identification of characteristic structural motifs that give rise to the overall shape of the respective protein. For example, one very common motif is the so-called alpha-helix—a screw-like looking secondary structural element that is mainly stabilized by hydrogen bonds. Crystallographic studies of aquaporins have suggested an X-shaped protein structure that is constituted by an arrangement of six membrane-spanning alpha helices with two additional short helical segments located near the pore region (Mitsuoka et al. 1999). Near the pore center, two NPA motifs are present, i.e., structural domains each constituted by the amino acid sequence asparagine-proline-alanine (de Groot et al. 2003, p. 280).

As regards proton exclusion, an early suspicion was that the passage of protons is hindered by an interruption of hydrogen bond networks between water molecules along which protons could otherwise have travelled. More precisely, the idea was that asparagine amino acids at the NPA region cause water molecules to be set apart and thus prevent them from creating hydrogen bonds with adjacent water molecules that would have allowed for proton conduction Murata et al. 2000, p. 604; see also de Groot et al. 2003, p. 280).

Whereas this early suspicion has later been called into question, the indicated reasoning already provides a good illustration of how research about the structure of proteins is often guided by the unveiling of a molecular mechanism and the aim of learning how molecular structure determines molecular function: what is of interest is whether and how the peculiar arrangement of amino acid sequences at the pore region poses an orientational restriction at water molecules that pass the channel.

4.2 Atomistic Molecular Dynamics Simulations and Their Role in Explaining Aquaporin Selectivity

Now that the general line of research on aquaporin proton blockage has become clearer, let us illuminate the role that simulations played in the context of aquaporin research. Consider that the aforementioned early idea according to which proton exclusion is due to the isolation of water molecules that prevent protons from hopping along hydrogen bonds was, in a sense, based on indirect reasoning (de Groot et al. 2003; Yarnell 2004). Rather than directly tracing the molecular dynamics at the pore region, the suspicion that an orientational restriction is posed on water molecules at the pore was based on static information about the shape of aquaporins (Hub et al. 2009, p. 58).

With this in mind, the motivation for using MD simulations in the context of explaining proton exclusion becomes clearer. Taking into account atomic-level information and tracing the dynamics of the system, such simulations seem to come with the promise of giving a more direct account of what happens at the protein pore and thus better illuminate the mechanism that prevents protons from permeating through aquaporins. However, before discussing the proposed explanation for proton blockage, it is helpful to give a brief overview over the general characteristics of the (here) most relevant class of atomistic simulations, namely molecular dynamics (MD) simulations. We suggest that by taking a closer look at the way in which scientists typically draw inferences with the help of such simulations, their explanatory role can be better understood.

On MD Simulations.

Molecular dynamics approaches are concerned with the examination of the movement, deformation and interactions of molecules over time (Leimkuhler and Matthews 2015, p. 1). MD simulations are used in many application fields ranging from, e.g., solid state physics and polymer physics to organic chemistry or biophysical chemistry. In all these fields, a typical aim that MD simulations help to pursue is to model the time-dependent evolution of systems and to learn, e.g., about the size and shape of molecules, about how the system behaves under pressure, about how molecules interact with each other, or about which molecular conformation a system will likely adopt (Leimkuhler and Matthews 2015, p. 1).

Classical atomistic MD simulations are based on the idea that the state of a microscopically described system at a given point in time is determined by the momenta and positions of all of its particles. Knowing the particles' coordinates and momenta at a certain time point and specifying the energy of the system, one can determine the state of the system at a subsequent time point

by calculating the forces acting on the particles as derivatives of the energy with respect to the particles' coordinates and solving the classical equations of motion.[2] By repeating this procedure and thereby "updating" the state of the system according to the respectively calculated forces acting on the particles at a given time step, MD simulations allow us to obtain so-called trajectories, i.e., paths of movement that show the dynamic evolution of all atoms in a system in time.

One central challenge, however, is the determination of a suitable energy function. Generally, the behavior of systems at an atomic resolution is governed by the fundamental laws of quantum mechanics, i.e., for non- relativistic problems, the solutions of the Schrödinger equation give the information needed for describing their properties. At this level, the elementary particles are the atomic nuclei and the surrounding electrons which give rise to all the different bonding patterns present in condensed matter. Unfortunately, however, the computational effort required for a quantum mechanical description is prohibitive for the treatment of many systems of interest.[3] Therefore, approximate descriptions of the energy of given systems in the shape of so-called force-fields are commonly used. Force-fields are parametrized functions that consist of various terms describing the relevant kinds of interactions to which the atoms of a given system are subject.[4]

The basic idea is that forces between atoms are induced by different sorts of intra- and intermolecular interactions (as shown schematically for water in 3.1). The interaction energy terms that are commonly used in MD approaches take very simple functional forms. For example, chemical bonds are modelled by using Hooke's law (i.e., by treating the atoms as if there were 'springs' placed between them). Besides bonding interactions, non-bonded interactions need to be modelled. They are commonly described in terms of Coulomb-interactions between the charged atoms, supplemented by simple energy terms for the so-called van der Waals (VdW) interactions, which are essential ingredients for the structure formation of all soft-matter systems.[5]

Now, whereas a description of the interactions between the microscopic constituents of systems is at the heart of molecular dynamics methods, it is important to stress that it is in many cases not the detailed particular path of molecular movement but larger-scale quantities that scientists are interested in when running a MD simulation. As, for example, Hoover puts it:

> The two main goals of microscopic molecular dynamics calculations are to simulate and to understand the macroscopic behavior in microscopic terms. We wish to under-

[2] According to Newton's second law, the force F is related to the particles' masses m and positions r as follows: $F = m\ddot{r}$.

[3] Only small systems containing several hundreds of atoms and timescales of several pico-seconds can be investigated.

[4] The appropriate parametrization of force fields is based on quantum mechanical calculations as well as on experimental (e.g., spectroscopic) data (Durrant and McCammon 2011, p. 2). The appropriateness of the parametrization crucially determines the success of a MD simulation. Consequently, the improvement of force fields constitutes an ongoing effort in the respective scientific communities.

[5] Soft matter systems are actually defined as being bonded through weak forces such as hydrogen bonding (Coulomb forces) and VdW interactions.

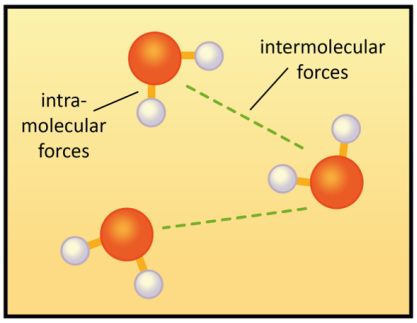

Figure 3.1: Forces induced by different kinds of molecular interactions (here schematically shown for water molecules).

> stand the way in which the relatively complicated microscopic many-body dynamics gives rise to the relatively simple macroscopic few-variable behavior described by phenomenological thermodynamics and hydrodynamics. (Hoover 1986, p. 42)

Thus, a typical aim is not to learn about individual trajectories, but to use trajectorial information in order to draw inferences about fewer larger scale variables that describe relevant properties of the respective systems. Even more, it is often clear from the very beginning that individual calculated trajectories do not necessarily have direct physical meaning. Consider, for example, the following statement by Leimkuhler and Matthews:

> The complex nature of the force fields involved and the large size of typical molecular systems mean that molecular dynamics is almost always chaotic. The changes in a molecule that occur over time are important, but these should be understood in terms of changes in averaged quantities, structural forms, or families of 'nearby' structures. Molecular dynamics relies on time-stepping to compute successive snapshots, but these are often used for sampling a probability distribution, or else a number of evolving paths are averaged to describe the likely sequence of changes that would be observed in a typical evolution of the molecule. (2015, p. 1)

Thus, rather than focusing on one detailed path, it is a common case that several different trajectories are used to compute probability distributions (from which free energy differences can be determined). Still, the question of why such probability distributions are of scientific interest arises. As has been shown, the overarching research questions tackled by means of MD simulations often concern the ways in which the microscopic details of a system determine the occurrence of larger scale events or phenomena: Which conformation will a certain molecular system adopt under specific circumstances? Why will it typically take that conformation? How is the particular

structure of some protein responsible for its capacity to catalyze a certain chemical reaction?

Whereas questions of the first kind concern the structural aspects of molecular systems, the latter question is about how this structure pertains to a molecular function—as has been illustrated earlier. Obviously, providing a detailed microscopic trajectory of the respective system does in none of these cases suffice for giving an explanation. Even if, for example, a simulation of aquaporins shows the detailed path of proton movement and reveals that they fail to enter the cell, then this does not yet explain as to why this behavior occurs. Rather, the so-called thermodynamic free energy F of a system constitutes a relevant quantity of interest.[6] This is because according to the laws of thermodynamics, systems will reach an equilibrium state at the point of minimal F: If a system is disturbed by an external perturbation, it will react such that it will move towards the minimum of the free energy, as schematically shown in Fig. 3.2.

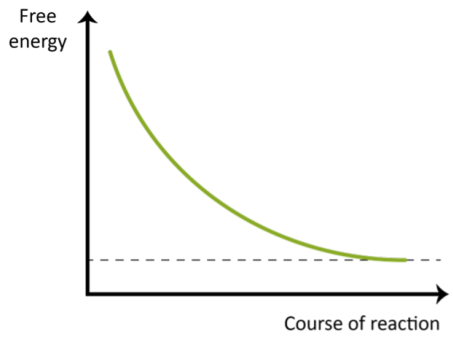

Figure 3.2: Free energy as a function of the course of a chemical reaction.

Thus, stable molecular structures or ensembles of conformations are characterized by a minimum of the free energy. If a scientific question is concerned with the identification of the structure of biological macromolecules, the guiding idea is the search for the minimum of F.[7]

If one addresses questions that concern the function of molecules, a typical aim is to determine the free energy of a system along a so-called reaction coordinate, i.e., a coordinate that displays the progress of a certain chemical reaction (e.g., the folding of a protein—or, in the case of aquaporin, the transport of water molecules through the pore). Reaction coordinates are a key for the analysis of chemical events in molecular systems as they allow systematic describing of certain types of molecular events by providing a projection onto a certain subset of molecular degrees of freedom

[6] This holds for systems where the volume is kept constant. For systems where the pressure is constant, the volume can change and the so called free enthalpy G is the relevant property. This means that G can play an equivalent role to F in a given simulation setup, depending on the system under investigation. In the remainder of this section, we will for illustration focus on F.

[7] And as we will see, this can easily be done using MD simulations. Moreover, note that in cases in which the aim is not only to predict but also to explain the emergence of a certain structural conformation, we may not only want to identify the minimum of the free energy but also analyze how different energetic factors or molecular interactions contribute to it. We will briefly outline such a case at a later point.

while averaging over all other, irrelevant degrees.[8] In the context of MD, focusing on the free energy of a system can, e.g., serve to quantify processes such as the transport of ions through membranes, electron transfer, or solvation of small molecules (Meirovitch et al. 2009). Moreover, if, for example, the calculated free energy pathway along a given reaction coordinate shows the occurrence of a so-called energetic barrier at some point, one can take the precise location of this barrier as a starting point in order to draw inferences about the factors that determine its occurrence – as will become clearer when discussing the case of aquaporins in the next section.

For now, however, the remaining elephant in the room is: How do molecular dynamics methods that are based on a description of a system's micro-constitution allow learning about thermodynamic properties of molecular systems such as changes of its free energy?

What makes it possible to establish bridges between a microscopic description of systems and a specification of their macroscopic properties within MD approaches are the principles of statistical thermodynamics. Consider that the microstate of a system consisting of N particles with three degrees of freedom is given as a point in a 6N-dimensional space constituted by the positions and momenta of all N particles. As has already been indicated, many different microstates of a system can correspond to one and the same macrostate (i.e., a state for which thermodynamic properties such as temperature, entropy etc. are specified). Call a collection of microstates that all correspond to the same macrostate an "ensemble." The energy E of a given atomistically-described system can be specified as a function $E(x, p)$ of its microstate. Taking into account fluctuation effects relevant at the microscopic level, statistical thermodynamics associates the thermodynamic internal energy U (i.e., the total energy contained in a system) with the system's average energy $\langle E \rangle$. Now to determine $\langle E \rangle$, one can proceed either by tracing the time evolution of the system and averaging over all reached microstates (which is, unsurprisingly, computationally expensive) or by considering the mean of the energy of all microstates k of an ensemble weighted by the probability of their occurrence:

$$U = \langle E \rangle = \sum p_k E_k \tag{3.1}$$

Here, the important point is that connecting a microscopic and a macroscopic view depends on referring to ensembles that can be treated statistically: The internal energy of a system is spelled out as the sum of microstate energies that are each assigned a certain probability. Boltzmann's famous formula

[8] It is, however, not always possible to find such coordinates. In some cases, the x-axis denoting the progress of a reaction remains abstract. Reactions such as the folding of a protein or enzyme catalysis are located in high-dimensional phase spaces that can include many irrelevant coordinates and make the search of a meaningful reaction coordinate very difficult (McGibbon et al. 2017, p. 146).

$$S = k \ln W \tag{3.2}$$

relates the number W of possible microstates to the entropy S of the thermodynamic system—and, based on this, a connection to the free energy can easily be drawn.

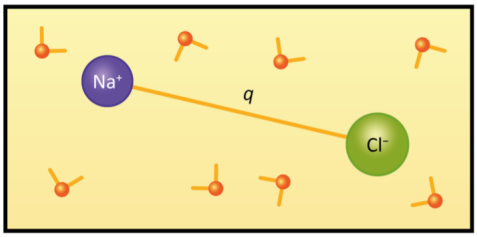

Figure 3.3: Na^+ and Cl^- in solution.

The free energy of a system is specified as its internal energy minus the product of its absolute temperature (T) and entropy[9]:

$$F = U - TS \tag{3.3}$$

As an illustration of how bridges between a microscopic and thermodynamic description of a system can be drawn, let us consider a simple problem of two ions, Na+ an Cl- in solution, as shown in Figure 3.3. These ions will move freely in the water solvent, but because of the mutual attraction due to Coulomb-forces, in simulations one will find a range of preferred distances, i.e., a range of most probable distances.

MD simulations allow determinations of the distances for different snapshots of the trajectories and obtaining a probability distribution of the distances, i.e., one can determine the relative probability $P(q_A)$ to find a certain distance (q_A).[10] By means of statistical thermodynamics, it is then possible to derive from these probabilities differences in free energy as follows:

$$F(q_A) - F(q_B) = -kT \ln \frac{P(q_A)}{P(q_B)} \tag{3.4}$$

Thus, rather than focusing on a single trajectory of the ions, often a large amount of trajectories serves as a basis in order to calculate the probabilities for certain distances. That is, from counting how often certain molecular conformations are met during a MD simulation, free-energy differences can be derived.

[9]The relationship between the free enthalpy (cf. footnote 6) and the free energy now goes as follows: $G = F + pV$. Note that the enthalpy is defined as $H = U + pV$.

[10]In this case, the distances q are taken as so-called reaction coordinates.

In the MD community, two major efforts concern (a) the improvement of force fields and (b) the development of enhanced sampling methods that enable scientists to compute free energy differences along a given reaction coordinate. We will in the following see how this latter aspect guides how MD simulations help tackle explanatory questions.

The Search for an Explanation of Aquaporin Selectivity.

After having elucidated how MD simulations can act as bridging tools between the microscopic and macroscopic properties of a system, let us finally take a closer look at their role in explaining the proton exclusion of aquaporins. Departing from the earlier mentioned initial idea of interrupted hydrogen bonds being responsible for proton blockage, different simulation studies that provided a more direct and dynamic picture of what happens at the protein pore have been conducted lately. In what follows, we do not aim at providing an exhaustive explanatory story of aquaporin selectivity. Rather, we zoom in on two exemplary studies and shed light on the role that computer simulations have played for the search of an explanation. We outline how these studies were guided by the aim to evaluate the role that different structural components of the molecule play for the observed behavior.

An early simulation study conducted by de Groot et al. (2003) has suggested that an electrostatic field at the pore region is mainly responsible for the reported selectivity of aquaporins. In their study, the researchers make use of a simulation technique that includes stochastically-described quantum mechanical proton hopping events in an otherwise conventional MD environment (de Groot et al. 2003, p. 289). As they point out, this approach makes it possible to explicitly take into account proton transfer reactions between H_2O molecules in the pore region and hence allows more direct tracing of the motion of protons at the channel (de Groot et al. 2003, p. 289).

Having performed simulation runs with protons placed at different initial positions in the pore, de Groot et al. found that most protons indeed fail to pass the earlier-mentioned NPA region near the pore center of the aquaporin molecule (de Groot et al. 2003, p. 281). Yet, why is the passage of protons blocked near the NPA region? Remember that when being interested in the factors responsible for a certain molecular function (here: proton exclusion), the free energy of a system along a given coordinate usually becomes a central property of interest. Given the non-equilibrium character of the proton trajectories, de Groot et al. did a so-called maximum likelihood calculation in order to obtain a free energy profile that realistically reflects the dynamics of protons at the pore (de Groot et al. 2003, p. 282).[11]

[11] Without going into the specifics of this method, it is interesting to note that – as has been indicated earlier – individual trajectories are from the very beginning not necessarily regarded as being physically meaningful but, quite in contrast, need to be treated statistically. As de Groot et al. themselves put it:

We have shown that when analyzing the function of macromolecules, a common aim is to examine how the free energy of a system changes along a specified reaction coordinate and to correlate observed energy barriers with structural motifs of the system. Hence, having done calculations of free energy changes along the protein pore axis, de Groot et al. (2003, p. 284) report the occurrence of two energetic barriers, a dominant one being located close to the aforementioned NPA region and a smaller one located near the so-called ar/R region.[12]

Interestingly, De Groot et al.'s simulations suggest that an interruption of the hydrogen-bonded water networks which has previously been assumed to be responsible for proton selectivity can be found mostly near the ar/R region rather than at the NPA region (de Groot et al. 2003, p. 284). Given that the energetic barrier at the ar/R region is much smaller compared to that at the NPA region, De Groot et al. infer that an interruption of hydrogen bonds is not the main factor for proton exclusion.

Being interested in the origin of the dominant energetic barrier near the NPA region, the scientists proceeded by comparing their calculated proton free energy profile to an estimate of the mean electrostatic potential at the pore region (de Groot et al. 2003, p. 284). They point out that interestingly, the observed maxima and minima of both the electrostatic potential profile and the free energy profile correspond with regard to their respective locations at the pore axis. Against this backdrop, they suggest that rather than being mainly due to an interruption of hydrogen bond chains, the proton exclusion mechanism in aquaporins is dominantly sparked by electrostatic interactions between the structural components of the protein at the NPA region, i.e., mainly by the macro dipoles induced by the earlier mentioned two helical loop segments (de Groot et al. 2003, p. 284). They infer from their calculations that even if the breaking of hydrogen bonds may still play a role for isolating channel-passing water molecules, the primary determinant of aquaporin proton exclusion is the effect of an electrostatic field erected mainly by two macrodipoles at the helical segments at the pore. What prevents protons from passing through the aquaporin channel is their electrostatic interactions with the aquaporin molecule at the pore center. The originally suggested interruption of hydrogen bond networks is only a secondary factor (de Groot et al. 2003, p. 288).

A further simulation study that additionally took into consideration the environment of the membrane has been conducted by Burykin and Warshel (2003, 2004). Whereas their study results agree with de Groot et al. in that a barrier present at the channel region is central for proton ex-

> "Note that the obtained profile does not necessarily reflect the detailed microscopic character of the underlying dynamics that occur in the individual trajectories (e.g., the instantaneous proton translocation during a hop event), but rather captures the ensemble behavior of the collection of trajectories, independent of the underlying microscopic mechanism." (2003, p. 292)

[12] Which is the aromatic-arginine region.

clusion, they came to different conclusions with regard to the origins of the barrier. According to Burykin and Warshel (2004, p. 45), a certain structural motif can only be taken to be responsible for the electrostatic barrier if removing that motif significantly affects the occurrence or magnitude of that barrier. Within their approach, they inter alia focus on showing that the NPA motif is not a difference-making factor in this sense. As a starting point, Burykin and Warshel ran MD simulations to generate multiple configurations of the protein for the relevant charged and uncharged states (2004, p. 42). In a second step, they used these configurations to calculate so-called electrostatic free energy profiles (i.e., profiles considering the contributions by the solvation energy). The obtained profiles showed that—in accordance with the earlier approach by De Groot et al.—a barrier is located at the protein channel.

Yet, Burykin and Warshel proceeded by analyzing each of the different energetic contributions to these profiles. Interestingly, they found that the observed barrier is mainly due to those contributions to the free energy that reflect the general desolvation penalty which occurs as the proton tries to enter the protein channel (Burykin and Warshel 2004, p. 43). Furthermore, they argue that even if the effects of the earlier mentioned helix dipoles are set to zero within their calculations, the observed electrostatic barrier is still high (Burykin and Warshel 2004, p. 45). Thus, rather than being due to the dipoles induced by the NPA motif of aquaporins, the scientists suggest that the exclusion of protons is mainly due to the general energetic costs of moving the proton charge from water to protein channel regions (Burykin and Warshel 2004, p. 45).[13]

In the last years, further simulation studies have provided additional support for Burykin and Warshel's explanatory approach (see Hoffmann 2006). However, rather than delving deeper into the details and differences of these studies, let us in the following shed light on the general way in which computer simulations were seen as means for explaining aquaporin selectivity.

5 The Explanatory Power of Atomistic Simulations

Remember that an initial concern was that the often-attested threat of explanatorily irrelevant details is particularly serious for the case of atomistic simulations. In a nutshell: Given the sheer amount of microscopic details and derivational steps, how does giving an account of the atomic level properties of a system help explain a certain phenomenon? So far, we have advanced the idea that the search for an explanation of aquaporin selectivity was decisively guided by the identification of structural features of aquaporins that are responsible for the molecular function in question. As

[13] As the scientists themselves (2004, p. 45f.) indicate, it can thus also be the lack of specific structural components that lead to the required polarity for proton transfer that can be crucial for the occurrence of the phenomenon in question.

regards the discussed simulation studies, both illustrate that despite starting with a description of the microscopic setup of the system under investigation, it is a thermodynamic property (i.e., the free energy) that becomes the central dimension of analysis for assessing the factors that potentially figure in an explanation.

One central merit of MD simulations in the context of macromolecular research clearly lies in their capacity to offer a starting point for deriving free energy profiles based on microscopic information. What is of interest is not the precise evolvement of individual proton paths but the extent to which the blockage of protons at the pore center is generally determined by certain effects or kinds of interactions that are given rise to by the structure of the protein or of the protein together with its surrounding. MD simulations set the stage for the application of statistical methods that in turn allow to calculate free energy profiles and to correlate the location of observed energy barriers with structural features of the system. Statistical thermodynamics provides a means to connect the microscopic features of a system with certain macroscopic properties. In this sense, statistical thermodynamic principles in the context of MD simulations give a tool for abstraction. Consider again Levy and Bechtel:

> An abstract description includes only some of what could, in principle, be said about its subject matter. It leaves matters open, in certain respects. A simple case would be specifying that a certain property falls within a range [...] while not indicating the exact value (which may, in principle, be known). More interesting cases involve depicting global features without specifying how they are realized. Thus, one can specify the average of some property, its variance, or some aspect of its distribution, without giving details about the individuals thus distributed or averaged over. (2013, p. 242)

Despite starting with a description of the dynamic evolution of a 6N dimensional system consisting of particles with certain momenta and positions, the explanatory merit of the performed simulations certainly did not consist in providing an as detailed as possible account of these trajectories. On the contrary, remember that connecting a microscopic description with a macroscopic view requires referring to ensembles (rather than particular microstates) that can be treated statistically.[14] Based on MD simulations and the specification of a reaction coordinate, free energy differences can be derived by counting how often certain conformations occur. Following Levy and Bechtel, this can be seen as a systematic way of neglecting details about the constituents of a system, or in Batterman's words, as a way to get rid of explanatory noise: Whereas the atoms of a system served as a starting point, one can now move from a description of their momenta and

[14] For abstraction in the context of statistical mechanics, see also Robertson (2020).

positions to a description of properties of the overall system such as its free energy along a certain reaction coordinate. Rather than being plagued with explanatory noise, MD simulation studies seem to offer an opportunity precisely to manage or get rid of such noise.

Making use of statistical mechanical methods, atomistic simulations provide resources to abstract from and systematically omit large amounts of microscopic details. While taking atomic-level information as a starting point, they can act as tools to draw inferences about larger-scale, e.g., thermodynamic properties of a system. Put differently, they grant access to a coarser level of description at which fewer variables that specify the potentially explanatorily relevant properties of complex systems can be taken account of.

In the case of aquaporins, computer simulations help with a struggle that is typical for many complex molecular systems: On the one hand, what happens at the atomic level is considered to play an important role for the behavior to be explained; on the other hand, in order to compare the relative effect of various structural or energetic factors on the explanandum, large amounts of atomic details need to be systematically eliminated, which is done mathematically by averaging within the framework of statistical mechanics. Moreover, in order to appropriately assess the explanatory role of computer simulations, it is essential to see how the formulation of explanatory interests and hypotheses occurs against a background of previous theoretical or experimental studies. Such studies guide the specification of a level of description at which the variables deemed adequate for an explanation are located and inspire the speculation about factors that could potentially figure in an explanation. That is, the research context of structural biology already equips scientists with a host of factors that could be relevant for an explanation of aquaporins – and the explanatory role of computer simulations was precisely to help weighing and comparing the extent to which the explanandum is sensitive to these potential contributors.[15]

With this in mind, it is interesting to come back to Lenhard's concern according to which simulations, despite being theory-based, leave it unclear how the underlying theory yields the respectively simulated behavior of a system. In the case of the golden nanowire, the expressed worry was that the relationship between the general Schrödinger equation and the formation of gold wires stays opaque. In a certain sense, we agree: In the case of molecular systems with many degrees of freedom, the structural equilibrium of molecular systems is determined by the delicate interplay of, e.g., hydrogen bonds, van der Waals interactions, chemical bonds, etc. In most applications, it is unfeasible to extrapolate from theoretical equations how the interplay of the manifold physical interactions between the constituents of molecular systems (as described by the potential mentioned earlier) leads to a stable structure of these systems. Here, computer simulations are, just

[15]We thank an anonymous reviewer for helping us make our argument clearer.

as Lenhard has highlighted, certainly a means for overcoming the complexity barrier because they make it possible to study these interactions in the first place.

Yet, when it comes to the explanatory power of simulations, we think it is important to stress that the Schrödinger equation—having the status of an universal axiom—is necessarily abstract and can of course not be applied directly to given phenomena. Rather, one needs to solve the Schrödinger equation for concrete potentials that pertain to specific molecular systems—or, equivalently, solve Newton's equations with these respective potentials. And to formulate such potentials, investing and specifying particular models such as, e.g., the harmonic oscillator is essential and needed for describing problems such as elasticity or molecular vibrations. If we want to explain phenomena that pertain to certain molecular properties, it is not the mere form of the Schrödinger equation but such models and the specific potentials that become the central dimension of analysis.[16]

And it is precisely because most potential for real materials are too complex to be solved analytically that computer simulations play such a relevant scientific role. As we have illustrated in our examination, they offer an opportunity for exploring, e.g., how structural or functional features of the system under investigation stem from the contributions of different kinds of molecular interactions or motifs to the respective potential.

While simulations may not render intelligible how precisely the respective phenomena in question follows from general theoretical assumptions, they help assess how the explanandum is sensitive to a range of factors. A molecular structure, as we have sketched earlier, is governed by the minimum of the free energy—and the location of this minimum may indeed not be foreseeable from the form of the invested potentials. Yet, simulations precisely give a tool at hand that allows weighing the relative contributions of the factors which lead the respective system into this free-energy minimum.

Lastly, consider that even for ab-initio quantum methods there seems no straightforward tradeoff between predictive capacity and explanatory force. Using so-called energy decomposition methods, one can analyze as to how different kinds of energy components (such as electrostatic, polarization, or charge transfer components) contribute to the overall interaction energy of a given system (Phipps et al. 2015). If their explanatory potential of simulations is at stake, one should assess it not so much in terms of how simulation results relate to theoretical assumptions, but – just as in the case of aquaporins – in terms of their capacity to contribute to a systematic disentanglement and analysis of how different factors of a system give rise to its overall behavior.

[16] In physics or physical chemistry textbooks, several simple potential forms are commonly introduced as prototypical solutions which can be applied real world problems—like the thermodynamic and spectroscopic behavior of simple two-atomic gases (such as O_2, CO_2) etc.) or solids. Certain aspects of these systems can be explained by referring to their atomic structure and the atomic interactions potentials by connecting structural features of the investigated system with characteristics of the invested models.

6 Brief Outlook at a Second Example: Ion Surface Preference

To further illustrate our point about how atomistic simulations can be tools for weighing and comparing various potential explanatory contributors, let us briefly consider a second example. Traditionally, scientists have considered it "conventional wisdom" that no ions are present at the air-solution interface of aqueous solutions (Jungwirth and Tobias 2002, p. 6361). However, more recent studies have shown that in contrast to this assumption and in contrast to the scientifically widely acknowledged general disposition of ions to be repelled from air-water interfaces, certain larger ions, namely large halide ions, are preferably solvated at the surface rather than inside small water droplets (Petersen and Saykally 2006; Caleman et al. 2011).[17]

Similar to the case of aquaporins, atomistic simulations have been brought into play in order to explain this surprising behavior. A general difference is, however, that in case of halide ions, the interest does not concern a certain (macro-)molecular function but the question of why the investigated molecular system as whole will adopt a certain kind of structural arrangement (namely one in which the ion is present at the droplet surface). As we have pointed out earlier, the location of the ion within the system in equilibrium is determined by the minimum of free energy.

In a MD-simulation study, Caleman et al. (2011) calculated the free energy along a reaction coordinate r which, in this case, describes the distance between the respective large ions and the droplets center of mass. As expected, it could be seen that all of the investigated large halide ions (Cl-, Br- and I-) possessed a local minimum in their free energy profile near the droplet surface.[18]

In order to assess the energetic factors responsible for the structural arrangement observed with regard to the solvation of large halide ions, Caleman et al. proceeded by disassembling the free energy changes in their enthalpic and entropic components[19]. They found that the locations of the energetic minima for halide ions are mainly determined by the respective enthalpic contributions (Caleman et al. 2011, p. 6839) – which is an interesting result as it shows that it is not primarily the entropy that is responsible for the movement of ions towards the surface, but the mutual interactions between the atoms.

Now Caleman et al. further decomposed the obtained enthalpic profiles into energetic contributions stemming from interactions between different water molecules on the one hand and from interactions between water molecules and ions on the other hand. This was done by considering

[17] This surface preference is of importance when examining how aerosols present in the Earth's atmosphere chemically interact with other atmospheric constituents (Caleman et al. 2011, p. 6838).

[18] Note that this is, of course, not yet an explanation of ion surface preference but rather confirms that experimental observation can successfully be reproduced. Had the simulation revealed by the energetic minimum in the bulk rather than at the surface of the respective droplets, then this would question the validity of the simulation model.

[19] More specifically, they consider changes in the free enthalpy G which relate to changes in enthalpy H and entropy S as follows: $\Delta G = \Delta H - T\Delta S$ (Caleman et al. 2011, p. 6839).

time-averaged interaction potentials (Caleman et al. 2011, pp. 6839–40):

$$\Delta H(r) = \langle \Delta V_{water-water}(r) \rangle + \langle \Delta V_{water-ion}(r) \rangle \tag{3.5}$$

As the scientists conclude, the interaction energy profiles allow to conclude that it is precisely the interplay of both water-water as well as ion-water interactions that is responsible for the observed surface preference (Caleman et al. 2011, p. 6840). The more the ions migrate toward the surface, the more do the water-water interactions (whose energy is mainly determined by the average number of water-water hydrogen bonds [see Hub et al. 2012, p. 9543]) become energetically favorable. Only at the point where the ions start to leave the droplet, the loss of ion-water interactions countervails the gain in water-water interactions so that the enthalpy increases again (Caleman et al. 2011, p. 6839). Thus, the large ions are drawn to the droplet surface because their presence in the bulk disturbs energetically favorable water- water interactions. Only when coming close to the surface of the droplet, the loss of water-ion interaction becomes energetically decisive for keeping them at the surface region. Accordingly, the interplay of both types of interactions is the central key driver for their location near the surface.

Now coming back to the question of how computer simulations play a role in providing scientific explanations, it is again not the precise molecular trajectory of the respective ions that does the explanatory work. Rather than being interested in all elements of the detailed path of movement that brought the ions to the surface, the crucial step was to again calculate free energy paths along a reaction coordinate and, here, to decompose the calculated free energy changes into their energetic components in order to determine the energetic factors responsible for the observed surface preference. The situation is one in which one has already observed a certain stable molecular structure (i.e., one in which the ions are at the surface) described by the minimum of F and in which the question is as to why the system adopts this structure.

Once more, it can be seen that by addressing the issue by means of a MD simulation and statistical methods, the scientists were able to consider averaged quantities and to systematically get rid of large amounts of molecular-level information and draw systematic inferences about thermodynamic properties such as the entropy, enthalpy, and free energy. Moreover, the case of ion surface preference illustrates how the systematic examination of energy components contributes to the investigation of potential difference-making factors. In their performed simulation studies, the scientists explicitly examined the possible impact of factors such as polarization or ion size on surface preference by varying their simulation input and analyzing the respective effects.[20] Again,

[20]For example, they ran simulations with inversely charged hypothetical ions and observed that no surface preference could be found. From this, they concluded that it is not the specific combination of polarizability and ion size that alone determines the surface preference of large halide anions (Caleman et al. 2011, p. 6840).

the explanatory role of simulations can best be captured in terms of their capacity to compare the contributions of different (energetic) factors and to assess as to how they contribute to the system's overall behavior.

7 Conclusion

In this paper, we have suggested that albeit MD simulations operate at a microscopic level of description, it is usually not primarily a reference to these details or to particular trajectories which does the explanatory work. Remember again Batterman's example of bending steel struts: a description of single detailed trajectories remains explanatorily inconclusive in so far as it does not help clarify why they show this typical bending behavior. Similarly, in the discussed cases, tracing how particular protons move through aquaporin pores or tracing how halide ions collide with water molecules until they come to the droplet's surface does not suffice to explain as to why the respective behavior can be observed. Even more, each particular trajectory will usually tell a slightly different story. Explanations are rather given at a thermodynamic level to which a connection is made with the help of statistical methods: in the case of the halide ions, the insight that the average energy is favorable at the droplet surface can only be drawn based on an analysis of an ensemble of trajectories.

As we have illustrated in the case of aquaporins, MD simulations can help with analyzing the extent to which various factors make a difference for the occurrence of the respective explanandum. Even if (as, e.g., in the case of the emerging nanowire discussed by Lenhard) no straightforward explanation of a certain surprising behavior could be given, we think that one should not conclude that relying on computer simulations means trading predictive force for explanatory force. The second discussed case about ion surface preference can, in particular, be seen as a counterexample here. Just as in the case of the observed gold nanowire, the behavior was (compared to what has been expected based on a macroscopic description or previous knowledge of the system) certainly surprising. This is a very common situation—which in fact nicely illustrates why computer simulations can be such a valuable scientific tool: it is likely that for phenomena whose behavior is governed by events occurring at the nanoscale, factors that are beyond the scope of a macroscopic perspective will become relevant. In some sense, it is precisely because phenomenological approaches are expected not to be fully applicable at the nanoscale that taking an atomistic perspective becomes so important. In the scientific community, atomistic simulations are commonly regarded as computational microscopes since they help unravel the (sometimes hitherto inaccessible) details that give rise to macroscopically observable phenomena. Even if, in some cases, no

explanation can be acquired, this does not imply that relying on computer simulations generally comes at the cost of explanatory power. In contrast, both the case of aquaporin selectivity and ion surface preference show that computer simulations can be relevant for explanatory purposes not despite that they are themselves quite complex tools and involve many derivational steps but precisely because they provide resources for systematically averaging over and abstracting from many details. Based on this, they allow accessing a level of description at which the factors that potentially figure in an explanation can be compared and assessed. In the case of aquaporins, atomistic simulations have contributed to the elucidation of a molecular mechanism by helping scientists to quantitatively assess as to how different regions of the channel protein contribute or fail to contribute to its biological function.

Our conclusion also relates to the more general point of how explanations can be obtained in cases in which scientists deal with highly complex systems. As, e.g., Imbert points out, explanations in such sciences may generally be more difficult to obtain due to the complex nature of the objects under investigation and the amount and character of the data that needs to be collected and interpreted (Imbert 2017, p. 753).

Finally, we think that no general answer to the question of how computer simulations provide explanations can reasonably be expected. In contrast, considering that computer simulations rely on various methods and are used in a vast variety of application fields, understanding their explanatory role will require us to take a closer look at these fields and to shed light on how precisely inferences about the respective targets are drawn by means of them. Explanatory interests are formulated against the background of existing knowledge, previous experiments, and theory. In the case of aquaporins, potential explanatory contributors can (and have been) identified even before the use of simulations. Rather than pointing to general features of computer simulations in order to render their explanatory role intelligible, it is fruitful to investigate how precisely they support scientific reasoning in concrete explanatory practices. In the case of aquaporins, their capacity to weight and compare the impact of potential contributors can be regarded as their main explanatory contribution. While the impact of these potential contributors is affected by what happens at the atomic level, their assessment requires moving to a coarser level of description with fewer variables. Atomistic simulations offer resources to filter micro-level information and draw systematic inferences about thermodynamic properties.

While stressing the need to focus more closely on concrete explanatory practices, we still contend the investigated cases point to a more general lesson, namely, that in fields that deal with complex systems, explanations may not only be more difficult to obtain, but simulations can be particularly useful tools as they enable a systematic reduction of degrees of freedom and allow us

to effectively assess and weigh features with regard to their effect on the explanandum.

Bibliography

Batterman, R. W. (2001). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. New York: Oxford University Press.

Bechtel, W. and Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2):421–441.

Bechtel, W. and Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3):321–333.

Borgnia, M., Nielsen, S., Engel, A., and Agre, P. (1999). Cellular and molecular biology of the aquaporin water channels. *Annual review of biochemistry*, 68(1):425–458.

Brigandt, I. (2013). Systems biology and the integration of mechanistic explanation and mathematical explanation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4):477–492.

Burykin, A. and Warshel, A. (2003). What really prevents proton transport through aquaporin? charge self-energy versus proton wire proposals. *Biophysical Journal*, 85(6):3696–3706.

Burykin, A. and Warshel, A. (2004). On the origin of the electrostatic barrier for proton transport in aquaporin. *FEBS Letters*, 570(1-3):41–46.

Caleman, C., Hub, J. S., van Maaren, P. J., and van der Spoel, D. (2011). Atomistic simulation of ion solvation in water explains surface preference of halides. *Proceedings of the National Academy of Sciences*, 108(17):6838–6842.

Craver, C. F. (2006). When mechanistic models explain. *Synthese*, 153(3):355–376.

Day, R. E., Kitchen, P., Owen, D. S., Bland, C., Marshall, L., Conner, A. C., Bill, R. M., and Conner, M. T. (2014). Human aquaporins: regulators of transcellular water flow. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1840(5):1492–1506.

de Groot, B. L., Frigato, T., Helms, V., and Grubmüller, H. (2003). The mechanism of proton exclusion in the aquaporin-1 water channel. *Journal of Molecular Biology*, 333(2):279–293.

de Groot, B. L. and Grubmüller, H. (2005). The dynamics and energetics of water permeation and proton exclusion in aquaporins. *Current Opinion in Structural Biology*, 15(2):176–183.

Durán, J. M. (2017). Varying the explanatory span: scientific explanation for computer simulations. *International Studies in the Philosophy of Science*, 31(1):27–45.

Durrant, J. D. and McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC biology*, 9(1):1–9.

Grüne-Yanoff, T. (2009). The explanatory potential of artificial societies. *Synthese*, 169(3):539–555.

Hoffmann, M. (2006). *Investigation of Bioenergetical Processes with Multiple-length-scale Approaches*. PhD thesis, Paderborn, Univ., Diss.

Hoover, W. G. (1986). *Molecular Dynamics*, volume 258 of *Lecture Notes in Physics*. Springer-Verlag.

Hub, J. S., Caleman, C., and van der Spoel, D. (2012). Organic molecules on the surface of water droplets–an energetic perspective. *Physical Chemistry Chemical Physics*, 14(27):9537–9545.

Hub, J. S., Grubmüller, H., and De Groot, B. L. (2009). Dynamics and Energetics of Permeation through Aquaporins. What Do We Learn from Molecular Dynamics Simulations? In *Aquaporins*, pages 57–76. Springer.

Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press.

Humphreys, P. (2009). The philosophical novelty of computer simulation methods. *Synthese*, 169(3):615–626.

Illari, P. (2013). Mechanistic explanation: Integrating the ontic and epistemic. *Erkenntnis*, 78:237–255.

Imbert, C. (2017). Computer simulations and computational models in science. In Magnani, L. and Bertolotti, T., editors, *Springer Handbook of Model-Based Science*, pages 735–781. Springer.

Janosi, L. and Ceccarelli, M. (2013). The gating mechanism of the human aquaporin 5 revealed by molecular dynamics simulations. *PLoS One*, 8(4):e59897.

Jebeile, J. (2018). Explaining with simulations: Why visual representations matter. *Perspectives on Science*, 26(2):213–238.

Jungwirth, P. and Tobias, D. J. (2002). Ions at the air/water interface. *The Journal of Physical Chemistry B*, 106(25):6361–6373.

Ko, Y. J., Huh, J., and Jo, W. H. (2008). Ion exclusion mechanism in aquaporin at an atomistic level. *Proteins: Structure, Function, and Bioinformatics*, 70(4):1442–1450.

Kuorikoski, J. (2012). Simulation and the sense of understanding. In Humphreys, P. and Imbert, C., editors, *Models, Simulations, and Representations*, pages 250–273. London: Routledge.

Leimkuhler, B. and Matthews, C. (2015). *Molecular Dynamics With Deterministic and Stochastic Numerical Methods*, volume 39 of *Interdisciplinary Applied Mathematics*. Springer.

Lenhard, J. (2006). Surprised by a nanowire: Simulation, control, and understanding. *Philosophy of Science*, 73(5):605–616.

Lenhard, J. (2014). Autonomy and automation: computational modeling, reduction, and explanation in quantum chemistry. *The Monist*, 97(3):339–358.

Lenhard, J. (2019). *Calculated Surprises: A Philosophy of Computer Simulation*. Oxford University Press.

Levy, A. and Bechtel, W. (2013). Abstraction and the organization of mechanisms. *Philosophy of science*, 80(2):241–261.

Liljas, A., Liljas, L., Lindblom, G., Nissen, P., Kjeldgaard, M., and Ash, M.-r. (2016). *Textbook of Structural Biology*, volume 8. World Scientific.

Marchionni, C. and Ylikoski, P. (2013). Generative explanation and individualism in agent-based simulation. *Philosophy of the Social Sciences*, 43(3):323–340.

McGibbon, R. T., Husic, B. E., and Pande, V. S. (2017). Identification of simple reaction coordinates from complex dynamics. *The Journal of Chemical Physics*, 146(4).

Meirovitch, H., Cheluvaraja, S., and White, R. P. (2009). Methods for calculating the entropy and free energy and their application to problems involving protein flexibility and ligand binding. *Current Protein and Peptide Science*, 10(3):229–243.

Mitsuoka, K., Murata, K., Walz, T., Hirai, T., Agre, P., Heymann, J. B., Engel, A., and Fujiyoshi, Y. (1999). The structure of aquaporin-1 at 4.5-å resolution reveals short α-helices in the center of the monomer. *Journal of Structural Biology*, 128(1):34–43.

Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J. B., Engel, A., and Fujiyoshi, Y. (2000). Structural determinants of water permeation through aquaporin-1. *Nature*, 407(6804):599–605.

Parker, W. S. (2014). Simulation and understanding in the study of weather and climate. *Perspectives on Science*, 22(3):336–356.

Petersen, P. B. and Saykally, R. J. (2006). On the nature of ions at the liquid water surface. *Annual Review of Physical Chemistry*, 57:333–364.

Phipps, M. J., Fox, T., Tautermann, C. S., and Skylaris, C.-K. (2015). Energy decomposition analysis approaches and their evaluation on prototypical protein–drug interaction patterns. *Chemical Society Reviews*, 44(10):3177–3211.

Robertson, K. (2020). Asymmetry, abstraction, and autonomy: Justifying coarse-graining in statistical mechanics. *The British Journal for the Philosophy of Science*.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.

Weirich, P. (2011). The explanatory power of models and simulations: A philosophical exploration. *Simulation & Gaming*, 42(2):155–176.

Woodward, J. (2005). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Woodward, J. and Hitchcock, C. (2003). Explanatory generalizations, part I: A counterfactual account. *Noûs*, 37(1):1–24.

Yarnell, A. (2004). Blockade in the Cell's Waterway. *Chem. Eng. News*, 82(4):42–44.

Computer Simulations, Scientific Explanation and Difference-Making: Lessons from Simulations of Ions in Aerosols

Contents

Summary

Obtaining scientific explanations can be a demanding task – especially if the phenomena under investigation are very complex. Given that computer simulations have become ubiquitous tools for studying the properties or behavior of complex systems, it is important to scrutinize their

role in the acquisition of scientific explanations. Aiming at a more comprehensive understanding of how computer simulations contribute to scientific explanations of complex systems, this paper looks at a case study from physical chemistry and investigates how atomistic simulations have concretely been employed by scientists to support explanatory reasoning about the surprising behavior of ions in atmospheric aerosols.

We suggest that in the examined case, just as in many other contexts that concern complex systems, scientific explanation is a matter of depicting various factors that make a difference for the explanandum. We argue that computer simulations can be powerful tools for explanation because they can contribute to the assessment of difference-making factors in at least three relevant ways. First, they allow scientists to perform a broad range of manipulations in a highly controlled way. Second, they can offer epistemic access to the 'right' variables at relevant levels of description. Third, they can provide a framework for quantitative analysis.

1 Introduction

Scientific explanation is a pertinent and controversial topic in the philosophy of science. Although the search for explanations is widely taken to be at the heart of science, it is often considered a challenging task – especially if the investigated phenomena are very complex and difficult or impossible to tackle by means of traditional scientific approaches. The nanosciences are a particularly illustrative example: At the nanoscale, the behavior of matter often deviates surprisingly from what would be expected on a macroscopic scale, leading to interesting scientific surprises and sparking a plethora of explanatory questions. Experiments to study such behavior are, however, often difficult to perform and phenomenological theories abstracting from molecular details often turn out to lead to inaccurate predictions.

Over the last decades, computer simulations have emerged as a powerful tool to investigate the behavior and properties of complex systems – in the nanosciences and elsewhere. From a philosophical perspective, given their prevalence in and relevance to scientific research, it is important to understand how they contribute to the search for scientific explanations.

In this paper, we look at the case of atomistic simulations as one important class of computer simulations, and examine how they can contribute to scientific explanations. We do so by paying close attention to how such simulations have been used by scientists in an actual explanatory setting. Concretely, we focus on an illustrative case from physical chemistry that concerns a nanoscale phenomenon relevant for chemical reactions occurring in the atmosphere, namely the surprising presence of certain large ions near the surface of small aerosol droplets. Considering that these ions potentially interact with climatically-relevant reactive gases such as ozone in the troposphere

(Jungwirth and Tobias 2002, 6363), various research groups undertook efforts to explain this phenomenon.

In what follows, we reconstruct and analyze how atomistic simulations were involved in the search for an explanation of the surprising presence of certain ions near the surface of aerosol droplets. The example of halide ions is particularly suited for examining the role of atomistic simulations for scientific explanations for two reasons: First, it deals – as we will see – with a common type of scientific questions about complex systems; second, it offers insight into a typical manner in which atomistic simulations are used by scientists to tackle explanatory questions.

We propose that, just as in many other scientific contexts, explanatory interests in the discussed case can best be reconstructed and understood from within a difference-making account of explanation in the tradition of Woodward (2003), and as it has been further developed for explanations that involve (computational) models by Bokulich (e.g. 2011, 2017). Starting from the idea that computer simulations can be important tools for assessing how the explanandum is sensitive to a range of factors (see e.g. Bechtel and Abrahamsen 2010, Ylikoski 2014, Schweer and Elstner 2023), this paper aims to provide a more comprehensive picture of the ways in which simulations can help with the disentanglement of difference-making factors, thereby offering a better understanding of their explanatory power.

Concretely, by drawing on the case study of ions in aerosol droplets, we propose that atomistic simulations can contribute to the assessment of difference-making factors in at least three relevant ways. First, atomistic simulations help with explanations because they can be tools for effectively manipulating systems in a highly controlled manner and in ways that would otherwise often remain unfeasible. As the examined case study will illustrate, simulations can be a powerful means for 'isolating' particular features of a system and artificially manipulating them in a way such that scientists can systematically assess the degree to which the explanandum is sensitive to these manipulations. Even in situations in which systems could be addressed by means other than simulation, simulation methods can come with the benefit of allowing for more fine-grained or tractable manipulations.

Second, atomistic simulations can grant epistemic access to the 'right variables'. That is, they can provide the resources for systematically getting rid of vast amounts of explanatorily irrelevant micro-level details of complex systems while at the same time capturing factors that cannot be accessed through the lens of macro-level approaches. Various philosophers have recently discussed how dealing with complex systems requires systematic abstraction, a reduction of degrees of freedom or coarse-graining (see e.g. Woodward 2016; Batterman and Green 2021). At the same time, it has been outlined that particular attention should be devoted to factors that figure not just on the *highest*, but on intermediary scales between micro and macro (Batterman 2021). We argue that

atomistic simulations can be powerful tools precisely for accessing such factors.

Third, and relatedly, atomistic simulations provide the resources for a quantitative treatment of complex systems. This matters for explanation because in the case of complex systems, it is often expected that the delicate interplay of various factors gives rise to the phenomenon or behavior to be explained. Hence, computer simulations contribute to scientific explanations by helping scientists to more precisely estimate and compare the degree to which various factors affect the occurrence of the explanandum.

Despite mainly focusing on atomistic (molecular dynamics) simulations, the case study inspires a broader outlook on challenges regarding scientific explanations in the context of complex systems, and on the role of computer simulations in providing these explanations in particular.

2 Scientific explanations of complex systems

Examining the ways in which atomistic simulations contribute to scientific explanations presupposes a reflection on what scientific explanation amounts to.

To get a grip on the explanatory questions at stake in the case of ion surface preference, it is useful to shed light on the scientific background against which explanatory interests have been formulated, as well as on some typical characteristics of explanations that concern complex systems[1]. Based on our reconstruction, we propose that a difference-making account of explanation in the tradition of Woodward (2003), and as further developed by Bokulich (2011), offers a fruitful framework for examining the contributions that atomistic simulations can make to scientific explanations.

2.1 Contrasts, complexity, and stability

Explanation-seeking questions often arise in a context of scientific surprise, i.e., in situations in which the behavior of a system is found to deviate from our expectations, or from what previous studies or examinations of similar systems have suggested. The case of halide surface preference offers a good example: What originally gave rise to an examination of ions at droplet surfaces was a more profound explanatory puzzle in atmospheric science, namely the surprising occurrence of certain chemical reactions at air/water interfaces in atmospheric aerosols (Tobias et al. 2013).

[1]Drawing on a characterization offered by Simon (1969, 86 as cited in Stöckler 2000, 363-364), we use 'complex' in a broad sense to refer to systems made up of a high number of parts that interact in non-simple ways. Of course, the suggested view on explanations may also fit for many contexts in which non-complex systems are considered. We simply suggest here that since explanatory questions about complex systems often concern the occurrence of stable or robust patterns or structures, an interest in invariant or robust dependence relation arises almost naturally. Also, we do not want to imply that the systems addressed by computer simulations *necessarily* need to be complex. However, we believe that computer simulations are *typically* used in situations in which the behavior or properties of complex systems are at stake.

Traditionally, it had been assumed that no atomic ions are present at interface regions between the air and aqueous solutions (Jungwirth and Tobias 2002). This assumption was backed up by considerations from within a classical continuum picture according to which water is treated as an idealized continuum medium with a certain dielectric constant and ions as 'hard spheres' with a uniform surface charge (Tobias et al. 2013, 344). In this picture, what possibly explains the lack of ions near the surface is that when approaching the interface region between water and air, the 'hard sphere ions' are repelled from the surface by their image charge on the air side of the air/water interface (Tobias et al. 2013, 344).

Calling the traditional picture in question, however, a host of climatically-relevant chemical reactions had been observed by atmospheric scientists to occur on sea-salt particles, ocean surfaces and on aerosols in laboratory contexts (Petersen and Saykally 2006, 333f.). Attempts to rationalize the occurrence of these reactions led to the suspicion that certain ions might be present at water/air interfaces after all and influence the reactivity of atmospheric aerosols (ibid.). Confirming such speculations, a number of recent experimental and computational studies have indicated that indeed, certain polarizable[2] anions are preferably located near the surface rather than in the inside of aerosol droplets (Jungwirth and Tobias 2002, Tobias et al. 2013).

This background shows that the explanatory setting in which the behavior of large halides is addressed is a *contrastive* one[3]: Given traditional assumptions about ions at air/water interfaces, the reported tendency of certain large polarizable ions to move to the surface of small water droplets raises explanatory questions because other atomic ions do *not* exert this behavior, and because traditionally, ions were generally believed to be located in bulk water rather than near air/water interfaces.

Another noteworthy fact about the case of halide surface preference is that explanatory interests concern a *complex* system, i.e., a system involving a large number of components exerting complicated interactions. It is increasingly pointed out that traditional accounts of explanation, such as a covering-law or causal-mechanical account, fall short of capturing how explanations of such systems work: While consisting of a plethora of details, inspecting all such details is unfeasible – and most of these details typically do not matter for an explanation. Giving a detailed story of the causal processes and interactions that lead to a certain explanandum, or deducing how they follow

[2]Polarizability describes the likeliness of an ion to acquire an electric dipole moment when exposed to an external electric field.

[3]In the philosophical literature on explanation, there is a debate on whether *all* explanations can plausibly be reconstructed as being in some sense contrastive (e.g. Lipton 1990; Hitchcock 1999; Ylikoski 2007). Rather than taking a stance on this issue, we suggest for the purpose of our paper that recognizing how explanatory reasoning is shaped by a particular background of expectations and contrasting previous results helps get a grip on the character of the explanation at stake. As Ylikoski and Kuorikoski suggest, paying attention to the 'contrastive settings' of explanatory examinations helps to "make the explanation more explicit in an analytically fruitful manner" (Ylikoski and Kuorikoski 2010).

from certain laws, is typically not only difficult but also strikingly inadequate (see e.g. Woodward 2003, 354; Bokulich 2011; Ross and Woodward 2023). According to Batterman, this is especially the case when being concerned with universal phenomena or the occurrence of stable structures or patterns (Batterman 2001, 2019, 2021).

The latter certainly applies to our example. Although the surface preference of large halide ions can be expected to result from fine-tuned molecular interactions, it is hard to see how giving an as-detailed-as-possible account of molecular details or citing the entire chain of molecular interactions leading to a particular ion's migration to the air/water interface could possibly lead to a satisfactory explanation. What is of interest is why certain *types* of molecular systems – those consisting of large halide ions in droplets surrounded by air – generally tend to adopt a certain structural arrangement; namely one in which the ion is located near the interface between air and water.

That is, explanatory interests concern the emergence of a stable structural arrangement and the factors based on which the occurrence of this arrangement is to be *expected*. While certain molecular details likely matter, the observed behavior is considered to be robust against certain changes in the micro-constitution of particular droplets (such as, e.g. the droplet size, or, to some extent, the temperature) and it is assumed that it occurs frequently enough to supposedly have a remarkable effect on the reactivity of atmospheric aerosols.

2.2 Difference-making

The previous section has highlighted two points: that (1) the explanatory setting in which questions about halide ions in aerosols have arisen is one in which the behavior of a system turned out to be different from what might have been expected, and (2), rather than being about the occurrence of a particular event, explanatory questions concern the expectedness of a particular behavior or structure, given that certain properties are present.

These points arguably apply to many explanatory contexts in the sciences[4], and they provide an idea of what the sought explanation in cases such as that of halide ions will look like: when explaining why the behavior of these ions deviates from most other ions, we seem to be interested in depicting a number of factors upon which the behavior of these ions depends. And since we are interested in the expectedness of a stable molecular arrangement, it seems natural to ask specifically for those dependence relations that hold under a range of variations in the examined system.

[4]In other words, explanatory settings are often contrastive, and a typical aim is to explain as to why a certain behavior or structure can be *expected* to occur. Consider, as another example, the case of enzyme biocatalysts, i.e., enzymes that catalyze chemical reactions. When explaining their capacity to act as catalysts, we are asking why chemical reactions occur faster in the presence of these enzymes, *as opposed to them being absent*, and we will be interested in depicting a number of factors (here, among other things, structural components of the molecules) that matter for why they will reliably speed up the respective reactions.

Many contemporary accounts of scientific explanation agree that scientific explanation is a matter of outlining various factors that make a difference for the explanandum to occur (e.g. Woodward 2003; Strevens 2008; Rice 2021). More specifically, it has been suggested that scientific explanations are concerned with counterfactual conditionals (Woodward and Hitchcock 2003; Woodward 2003; Ylikoski and Kuorikoski 2010; Bokulich 2011; Rice 2021). The central idea of what can be captured under the header of a 'counterfactual account' of explanation is that scientific explanation amounts to depicting a number of factors to which the occurrence of the explanandum is sensitive. Concretely, the counterfactual account essentially regards explanations as exhibitions of systematic patterns of counterfactual dependence (Woodward 2003, 191). As Woodward puts it, "[an] explanation must enable us to see what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways" (ibid., 11). In his famous wording, explanations are answers to what-if-things-had-been-different questions.

We suggest that the explanation sought by scientists in the case of halide surface preference can best be captured and rationalized from within a counterfactual, or difference-making, account of explanation: when explaining the behavior of these ions, we are looking for relations of counterfactual dependence that are sufficiently robust to capture the stable preference of halides to migrate to the surface of aerosol droplets. More specifically, what will be of interest is how the expected behavior of these ions depends on various factors, e.g. on properties that these ions have, as well as on the types of molecular interactions that occur between the respective ion and its surrounding.

2.3 Explaining with (computational) models

Before dealing with the concrete contributions that computer simulations – and, more specifically, atomistic simulations – can make for the assessment of difference-making factors, it is worthwhile to briefly address the preliminary question of how it is *possible* for them to support the disentanglement of dependence relations in real-world targets in the first place.

Rather than directly intervening on or interacting with the target system, drawing inferences from computer simulations requires us to work with a computationally-implemented *model* of the target system. In the case of halide ions, we need a model containing information about the molecules of the system (i.e., the respective ion and surrounding water molecules) and their interactions, as well as a starting geometry and appropriate boundary conditions.

A comprehensive account of how scientific models can contribute to the assessment of difference-making factors is offered by Bokulich (2011, 2014, 2017). For the context of our paper, we suggest that Bokulich's account offers a fruitful starting point for understanding how explanations given

with the help of computer simulations draw on scientific models.

Bokulich formulates three key features of 'model explanations'. The first feature of model explanations is that the explanans makes reference to a scientific model (Bokulich 2011, 39). As Bokulich puts it:

> "Model-based explanations (or model explanations, for short) are explanations in which the explanans appeal to certain properties or behaviors observed in an idealized model or computer simulation as part of an explanation for why the (typically real-world) explanandum phenomenon exhibits the features that it does." (Bokulich 2017, 104)

The second feature of model explanations is that models should 'reproduce' the relevant properties of the target system and offer information about how the target system would have been different if certain aspects of the model had been different (Bokulich 2011, 39). Bokulich here draws on Woodward's counterfactual framework[5] and suggests that the counterfactual structure of the model should in relevant ways be isomorphic to the counterfactual structure of the target system (Bokulich 2011, 39).

Let us briefly clarify as to what this means in the case of complex molecular models. For problems such as the surface preference of halide ions, the computational model that underlies a simulation takes the form of a so-called 'force field', together with a starting configuration of the atoms in the system, and appropriate boundary conditions. Force fields are parameterized potential energy functions that provide an estimate mathematical description of the relevant types of interactions between the atoms of a systems. Each type of interaction that is considered relevant for a given problem (think, for example, of Van-der-Waals interactions or bond stretching energies) should be reflected in a term of a given force field.

That is, in order for the counterfactual structure of the computational model to be 'isomorphic' to the target system in relevant regards, the parameterized energy terms should adequately reflect the relevant types of interactions between the atoms in the target; such that in a simulation context, the model can be worked with as a 'stand in' for the target.[6]

[5] At the same time, Bokulich departs from Woodward by dropping the interventionist underpinnings of his account. In Woodward's causal-interventionist framework, the antecedents of counterfactuals are made true by so-called interventions, heuristically to be understood as manipulations that could potentially be performed in an idealized experimental setting (Woodward 1997, 29). Interventionism restricts Woodward's account to *causal* explanations. Bokulich suggests letting go of interventionism and proposes that the counterfactual dependence relations which figure in model explanations can have different origins, allowing for a distinction between different *types* of model-based explanation – causal explanation being just one of them (Bokulich 2011, 40). While exceeding the scope of this paper, it would be an interesting topic for future work to scrutinize in more detail how computer simulations support such different types of counterfactual explanations.

[6] Yet is is important to note that in the case of force-field models (other than may be the case in certain highly idealized models), those elements of the model which are supposed to capture the counterfactual structure of the system are not necessarily *identical* to the candidate difference-making factors to be cited in explanations given with the help of such models. Even if the structure of the target is 'preserved' in relevant regards by a computational model by means of appropriate parameterized energy terms, the relevant difference-making factors will need to be extracted from such models by drawing

This also brings us to Bokulich's third feature of model explanations, namely that they require a *justificatory step* in which the domain of applicability of the model, as well as its representational scope are assessed (ibid.). In other words, there needs to be justification that the model indeed reflects the relevant features of the target system to be explained (Bokulich 2008, 226). According to Bokulich, the required justification is typically based either on theory that 'backs up' the model's domain of applicability, or on empirical considerations, or a combination of both (Bokulich 2011, 39).

In the simulation of molecular systems – such as ions in aerosols – the required justification indeed typically involves both theoretical and empirical reasoning. The computational modeling of molecular systems looks back to a decades-long history of development of numerical methods, as well as the careful calibration and stepwise improvement of computational models in light of theoretical predictions and experimental data. The parameterization of force fields, for example, can be based on calculations from quantum mechanical theory, but it can also be a matter of fitting of the model with experimentally-accessible properties, often in a 'hand'-tuned and incremental way. While, of course, activities such as the choice of the functional form of a molecular model or its parametrization can potentially introduce errors and hence cause distortions or unintended computational artefacts (see also Ylikoski 2014, 329), this should not cast general doubt on the explanatory power of computer simulations, but simply underlines that they are fallible — just as other tools can be (Zednik 2015).

3 The case of halide surface preference

Aiming at offering insight into the concrete contributions that computer simulations can make to the assessment of difference-making factors, let us now turn to our case study and examine how concretely atomistic simulations have been involved in the search for an explanation of ion surface preference. Note that a considerable number of simulation studies using different models and methods have been performed in recent times, and they have partly come to different conclusions with regard to their explanation of the inspected phenomenon (Tobias et al. 2013). The aim of the following investigation is not to clarify as to what eventually *explains* halide surface preference, but – taking one particular simulation method and study as a starting point – to shed light on *how* atomistic simulations were employed in the *search* for an explanation of halide surface preference, thereby providing insight into the systematic explanatory contributions that such simulations can

systematic inferences with simulations. In a sense, this 'extraction' is precisely what our paper aims to give a more detailed picture of.

make.

3.1 A brief introduction to molecular dynamics simulation

The simulations used in the case of halide ions are molecular dynamics simulations, a common atomistic simulation method. To understand the scientific setting in which explanatory questions about halide ions in aerosols were addressed, it is helpful to briefly shed light on how such simulations 'work' and how they are used concretely as tools in explanatory scientific practices.

We have earlier said that the use of simulations in cases such as our example is based on having crafted a 'force field' for the system, and that force fields reflect the types of interactions between the atoms or molecules in the system, thereby offering a description of its energy. Given that information about the energy can be used to learn about the forces acting on the atoms of the system, the essential idea of atomistic molecular dynamics simulations is to numerically solve Newton's equation of motion for the atoms of the system, thereby producing atomic trajectories of the system in time.

Bridging between micro and macro

One important goal of molecular dynamics simulations is to use a description of micro-level motion and interactions to obtain information about the energetic, structural or kinetic properties of a system that undergoes a certain chemical reaction. The underlying theoretical framework for this purpose is statistical mechanics. More concretely, the idea of statistical mechanics is to bridge between a *micro-* and *macrostate* description of a system. Whereas the former is given in terms of the positions and velocities of the particles in it, the latter describes thermodynamic properties of a system such as its temperature. Many microstates can correspond to one and the same macrostate.

When using statistical mechanics, the idea is to employ equations by means of which one can *bridge* from a description of particle velocities and positions to a thermodynamic description. Enabling a computation of atomic trajectories, MD simulations provide the resources for drawing inferences about larger-scale, e.g. thermodynamic, quantities of a system.

A particularly important thermodynamic quantity is the free energy[7] G. Free energy changes along the course of a chemical reaction inform us when a given system is stable or whether a certain reaction runs spontaneously; and free energy differences are used to address a plethora of chemical problems, ranging from molecular conformational preferences over binding constants to absorption coefficients, just to name a few.

[7] In the following, we will focus on the Gibbs free energy G that applies to systems with constant pressure and temperature. For systems at constant volume and temperature, there is the Helmholtz free energy A.

Having at hand trajectorial data obtained by molecular dynamics simulation runs and using statistical mechanical methods, free energy differences can be computed. As an illustration, consider the change of a protein in solution from an unfolded to a folded state, as shown in Figure 4.1.

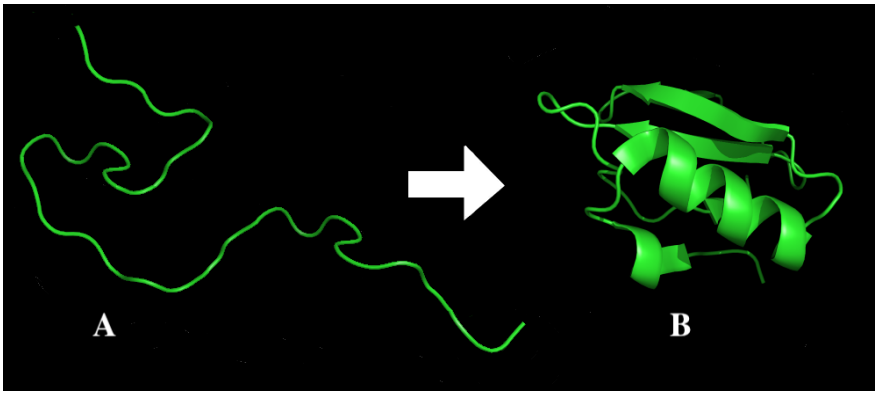

Figure 4.1: Process of protein folding. A denotes the unfolded and B the folded state. Adapted from DrKjaergaard, public domain, via Wikimedia Commons.

Using different snapshots of the trajectories, MD simulation data provide the resources to count how often the protein adopts a folded (B) or unfolded (A) conformation, respectively. Based on this, a probability $P(X)$ of finding the system in a certain state X can be assigned. The principles of statistical mechanics then make it possible to obtain free energy differences between two given states of a system[8]:

$$\Delta G = G(B) - G(A) = -kT \ln \frac{P(B)}{P(A)} \tag{4.1}$$

Accordingly, by using statistical mechanics, MD simulations can be employed as *bridging tools* between a description of molecular configurations at the microscopic scale and a description of thermodynamic quantities.

Learning about energetic factors

We have earlier characterized ion surface preference in droplets as a problem that concerns the occurrence of a stable molecular arrangement: The explanandum concerns the tendency of certain ions – when being placed in small water droplets – to migrate to and eventually stay near the surface of these droplets.

When analyzing the occurrence of stable molecular structures or patterns, free energy differences are commonly investigated – and the employment of statistical mechanical principles to derive free energy differences is what makes MD simulations a powerful tool for investigating them.

[8]In the case of proteins, it is important to note that one is dealing with a so-called *rare* event which requires appropriate sampling methods to make sure that the relevant parts of phase space are included (i.e. the relevant molecular conformations are 'visited') to accurately capture the transition path between the states.

In this regard, the case of ions in aerosol droplets resembles the aforementioned example of protein folding. In both cases, we are concerned with the course of certain chemical reaction that leads to a certain stable molecular arrangement: In the case of ions, their movement from the droplet interior towards the surface can be regarded as a reaction occurring along a reaction coordinate capturing the path of ion movement, and molecular dynamics can be used to analyze free energy differences along this coordinate.

Typical questions arising when analyzing such reactions concern the energetic factors that determine the shape of the respective free energy curve of a system along the reaction coordinate. In particular, a common question is as to how free energy curves are shaped by entropic and enthalpic contributions, respectively. Gibbs free energy changes relate to changes in enthalpy ΔH and changes in entropy ΔS as follows:

$$\Delta G = \Delta H - T\Delta S \tag{4.2}$$

Many biologically relevant chemical reactions are 'entropy-driven', meaning that their course is mainly determined by the increase of entropy. In molecular systems, an important interpretation of a system's entropy is given in terms of conformational freedom: Roughly put, an increase in conformational entropy means that more 'arrangements' of the molecules in system become possible. Consider the earlier-introduced case of proteins as an example: If a protein folds, the protein itself typically loses entropy because a folded state restricts the number of possible conformations. Enthalpic changes, on the other hand, correspond to the effects that inter- and intramolecular interactions have on the energetics of the system.

If, for example, a reaction leads to the formation of new non-covalent bonds, this will correspond to a release of heat energy and, thus, a negative enthalpy change. In the example of protein folding, enthalpy changes could be due to the formation of hydrogen bonds or the presence of Van-der-Waals interactions. In explanatory contexts, deriving free energy differences can be important because they are the possible starting point for a more detailed analysis of how different energetic factors and interactions between the molecules in a system affect molecular behavior.

3.2 Simulation study by Caleman et al.

Having set the stage, let us now turn to how atomistic simulations have concretely been employed to support the search for an explanation of halide surface preference. In what follows, we will take a closer look at a MD simulation study conducted by Caleman et al. in 2011. Again, given that our focus is to offer an examination of how simulations help with scientific explanations, we do not

aim to scientifically evaluate as to what eventually explains the surface preference of large halide ions, but to offer an analysis of how simulations were used in the search for this explanation.

A number of factors come into question for playing a role in explaining the surface preference of large halides. First of all, halide anions are known to be more polarizable than alkali cations. This means that – compared to the cations – they are more prone to having the distribution of their charge affected (i.e. their electrons distorted from their original arrangement) when being exposed to an electric field. Second, they obviously differ from the cations in the sign of their charge. Third, compared to the small halide anion fluoride that does *not* show surface preference, Cl^-, Br^- and I^- are larger.

It seems reasonable to expect that all three properties somehow play a role for the tendency of Cl^-, Br^- and I^- to be located the air/water interface of droplets. What is open, however, is a clarification of the relative extent to which the different properties determine the surface preference of large halides; and, more specifically, the reason *why* these properties bring about the observed behavior.

Using force-field based atomistic MD simulations, Caleman et al. simulated the behavior of various ions in small water droplets. More specifically, having selected a number of initial ion positions in droplets and using energy minimization, Caleman et al. ran various simulations each tracing the respective ion on its way from the droplet center to the surface region. From the obtained trajectories, free energy profiles were derived (Caleman et al. 2011).

Besides examining the free energy profiles of these large negatively charged ions (anions), Caleman et al. also ran simulations for the small halide anion fluoride (F^-) and for the alkali cations Li^+, Na^+, K^+, Rb^+ and Cs^+.

Based on their simulations, Caleman et al. computed free energy profiles and found that – as could have been expected[9] – the free energy profiles of those ions exhibiting surface preference possess a minimum near the surface (Caleman et al. 2011). In contrast, the free energy profiles of both the alkali cations and the small halide ion fluoride were found to increase monotonically as the ions move closer to the surface, suggesting that these ions are energetically favored in the bulk, as opposed to exhibiting surface preference (Caleman et al. 2011).

The role of various ion properties

Before Caleman et al.'s study, various groups had suggested that polarizability might be the main determining factor for the surface preference of large halide ions (see Petersen and Saykally 2006).

[9] That is because free energy minima indicate the point at which a system is in chemical equilibrium. Had the minimum been elsewhere, this would have cast doubt on the validity of the simulation.

More precisely, one idea was that the potential at the air/water surface induces a dipole in polarizable ions that eventually pulls halide anions to the surface. Caleman et al. explicitly address this idea in their simulation study. Inter alia, they examined for all ions how the ion dipole changes as a function of the position of the ion in the droplet: if the polarizability of large halide ions mainly determined their surface preference, one would expect that those ions which exert surface preference are most polarizable. Interestingly, the researchers report that while the dipole of the ions indeed increases slightly as they approach the surface, this increase is most pronounced for the small halide ion fluoride that does not show surface preference (Caleman et al. 2011). Hence, polarizability alone does not suffice to distinguish those ions that show surface preference from others.

A subsequent hypothesis tested by Caleman et al. was that for an ion to exert surface preference, it needs to be both polarizable *and* large: Even if fluoride is the most polarizable ion, it is comparably smaller than those ions that exert surface preference. Vice versa, all those ions that show surface preference indeed have in common that they are both polarizable *and* large. If nothing but the combination of size and polarizabilty mattered, surface preference should also occur for fictitious ions that have same polarizability and size as the halide ions but an artificially inverted charge sign. Testing this, Caleman et al. computed free energy profiles for 'hypothetical cations' whose size and polarizability is identical to those of Cl^-, Br^- and I^- but whose charge is artificially inverted.

If polarizability in combination with ion size were sufficient for explaining surface preference, then surface preference would also occur for hypothetical cations that have exactly the same properties as the large halide ions but only differ in the sign of their charge (Caleman et al. 2011). However, the obtained free energy profiles did not show a minimum near the droplet surface. Hence, Caleman et al. conclude that the combination of polarizability and ion size alone cannot sufficiently explain surface preference. Vice versa, they infer that the sign of the charge matters (ibid.).

Structural rearrangements and shift towards an energetic perspective

Since attempts to explain the surface preference of halide ions by referring solely to one or more of their properties remain unsatisfactory, Caleman et al. proceed with an analysis of the thermodynamics of the system. As has been sketched earlier, reactions are commonly analzyed in terms of how they are driven by enthalpy or entropy, respectively; and free energy profiles obtained by means of atomistic simulations can be a powerful starting point for such an analysis.

Recall that the free energy of a system can inform us about the point at which a system is at chemical equilibrium. Having computed free energy profiles of all ions as they migrate to the

surface, Caleman et al. decomposed the obtained free energy profiles into contributions stemming from enthalpic forces on the one hand and entropic forces on the other hand (see Equation 4.2). It was found that for all halide ions, the entropy decreases as they move towards the surface, meaning that if nothing else were to play a role, the entropic contributions would make it energetically preferable for halide ions to stay in the bulk of aerosol droplets rather than migrate to the surface.

In the case of the *large* halide ions, however, favorable enthalpic changes make a dominant contribution to the location of the energetic minimum near the surface. This means that in the case of the large halide ions, unfavorable entropic changes only partly counterweight the favorable free enthalpy near the surface. In other words, contributions stemming from molecular interactions in the system rather than entropic effects make the surface region energetically favorable for these ions (Caleman et al. 2011, 6839).

Even though entropic factors alone would make it favorable for the halide ions to remain in the bulk, the relative impact of enthalpic contributions outweighs these factors in the case of large halide ions (Caleman et al. 2011, 6840). In the case of the small halide ion fluoride, in contrast, the enthalpic changes are less pronounced so that they are overcompensated for by entropy reduction (ibid.).

A possible explanation for how the enthalpic changes bring about the surface preference large halide ions is that their presence in bulk water 'disturbs' the preferred local structure[10] of surrounding water molecules.

However, having employed a certain method[11] to quantitatively describe how the local structure of fluids is affected by the presence of a given particle, Caleman et al. found in their simulation study that even though the orientation of water molecules to a respective ion seems to be somewhat sensitive to whether the ion is near the surface or in the bulk of a given droplet, the effects on the local structure are similar for *all* ions – i.e., they are not specific for those ions that show surface preference. Hence, changes in the local structure of water molecules do not seem to be the main factor for why large halide ions show surface preference.

[10] 'Structure' here refers to the local arrangement of molecules in a given fluid (e.g., the arrangement of molecules of a solvent relative to a reference point or the arrangement of molecules in a solvent in which a solute is placed). As e.g. Roberts et al. (2009) put it: "Water is highly structured because of its ability to form up to four hydrogen bonds, resulting in a tetrahedral network of molecules."

[11] A common method to quantitatively describe the structure of fluids is the use of certain correlation functions – called radial distribution functions (RDFs). More specifically, RDFs are a means to statistically specify the structure of fluids at molecular length scales by determining the probability of finding a particle at a certain distance relative to another reference particle.

Disentangling the impact of two types of molecular interactions

In a subsequent step, Caleman et al. proceeded with a more detailed analysis of how enthalpic changes affect ion movement in droplets. Note that in the examined case, changes in enthalpy can be interpreted as being composed of two kinds of molecular interactions, namely those between water molecules and the ion on the one hand and those between water molecules and other water molecules on the other hand.

Even if effects on the local arrangement of water molecules are not the key factor for the surface preference of large halides, it may still be the case that their presence more subtly affects how the different kinds of molecules in the systems interact. Caleman et al. thus further decomposed the enthalpic profiles of all ions into the contributions stemming from the water-water interactions and water-ion interactions, respectively. This was facilitated by interpreting enthalpy changes ΔH along the reaction coordinate r as the sum of time-averaged[12] water-water and water-ion interaction potentials V (Caleman et al. 2011, 6839):

$$\Delta H(r) = \langle \Delta V_{water-water}(r) \rangle + \langle \Delta V_{ion-water}(r) \rangle \tag{4.3}$$

The obtained profiles show that water-water interaction energies decrease as the ions approach the surface. At the same time, water-ion interaction energies increase. As Caleman et al. (2011, 6840) argue, the obtained profiles suggest that the partial desolvation of large halide ions enables more favorable water-water interactions. Only when arriving at the proximity of the droplet surface, losses in ion-water interactions start to become energetically more dominant over gains in favorable water-water interactions, preventing the ions from 'leaving' the droplet entirely (Caleman et al. 2011, 6839). Vice versa, their results suggest that alkali cations are energetically favored in the bulk because their desolvation hardly affects water-water interactions (ibid.).

Overall, this suggests that taken together, properties such as polarizability, charge sign and ion size certainly *do* matter, but getting a grip on their significance to halide surface preference requires to acknowledge how the presence of an ion having such properties energetically influences its surrounding. In other words, according to Caleman et al, large halide ions migrate towards the surface of droplets because a particular factor becomes more pronounced than for other ions, namely how their presence negatively affects how water molecules exert interactions with other water molecules.

This is different in the case of other ions whose presence hardly affects interactions between water molecules and makes them prone to stay in the bulk. Hence, large halide ions differ from

[12]The square brackets denote the averages over the MD trajectories.

other ions in that enthalpic effects are specifically pronounced for these ions, more precisely the effects by water-water-interactions. A disturbance of the latter is what eventually makes a difference for them be energetically preferred near the surface.

4 Three explanatory contributions of atomistic simulations

The case of halide surface preference has given an impression of how atomistic smulations can help tackle explanatory questions by supporting the assessment of difference-making factors.

In the remainder of the paper, we aim to offer a more fine-grained analysis of their explanatory contributions. More specifically, the idea is to show that not only can such simulations take part in an assessment of difference-making factors, but that there are reasons for why they are often particularly valuable tools for doing so. In other words, by drawing on the counterfactual account, the remainder of the paper discusses concrete regards in which atomistic simulations can contribute to explanations, thereby also giving a more comprehensive account of their explanatory power.

Concretely, we suggest that atomistic simulations can (1) make it possible to perform a broad range of (sometimes actually counterfactual) manipulations in a highly controlled way, that they (2) can allow to access to the 'right' variables which code for the stability or expectedness of the explanandum, and that (3) they can help with a quantitative assessment of difference-making factors, thereby allowing to estimate in more detail the extent to which the explanandum is sensitive to changes in these factors.

4.1 Tools for controlled manipulation

In the counterfactual account, explanation is a matter of examining how the explanandum would have been different if various factors had been different. In other words, when trying to explain the behavior or properties of a system, we are interested examining how variations in certain features of the system would affect the explanandum.

When dealing with systems that are sufficiently complex, performing this task can, however, become notoriously difficult. If the behavior of a system depends on the intricate interplay and interdependency of a large number of components, tracking the effect of (hypothetical) changes on particular factors is often hardly feasible or impossible – especially when dealing with scales that are difficult to tackle by traditional experimental approaches.

As an illustrative example, consider how scientists describe the limits of traditional structure-based methods as concerns studies of the behavior of macromolecules, pointing to the relevance of simulation approaches:

> "An atomic-level structure is tremendously helpful and typically generates substantial insight about how the biomolecule functions. The atoms in a biomolecule are in constant motion, however, and both molecular function and molecular interactions depend on the dynamics of the molecules involved. One would like not just a static snapshot but the ability to watch these biomolecules in action, to perturb them at the atomic level, and to see how they respond." (Hollingsworth and Dror 2018)

Atomistic simulations can be tools to perform such pertubations. We suggest that the first way in which such simulations can make an important contribution to the assessment of difference-makers is by increasing the number of possible manipulations in the first place, and by acting as powerful 'tracking tools' for the effect of these manipulations.

This point is recognized in the philosophical literature beyond atomistic simulation. According to Ylikoski, computer simulations can contribute to explanation by increasing the number of possible counterfactual interventions in the first place (Ylikoski 2014). Discussing the case of agent based simulations in the social sciences, Ylikoski (2014, 331) states that with such simulations, "[...] we can systematically study how the changes in the assumptions change the outcomes."

Similarly, dealing with dynamic mechanistic explanations in cognitive science, Bechtel and Abrahamsen (2010) name the expansion of parameter value spaces as one important regard in which the explanatory power of simulations can be rendered intelligible. As they put it: "A model provides a means of exploring a much larger space of parameter values than would be feasible experimentally, and thereby of projecting how the actual mechanism would behave under a variety of conditions" (Bechtel and Abrahamsen 2010, 325).

The case of ion surface preference further underlines this idea and stresses the high degree of control that simulations enable over such manipulations. Recall that as part of their search for an explanation, Caleman et al. performed interventions on what they called "hypothetical" (Caleman et al. 2011, 6840) ions. The explanatory finding that ion size together with polarizability is not sufficient for surface preference was gained by examining whether the explanandum still holds for these artificial ions with the polarizability and size kept fixed, but with an inverted charge sign.

By means of their simulation, the researchers could intervene on the charge by setting it to a counterfactual value – from negative to positive – while keeping everything else fixed. Thus, simulations can enable a particularly high degree of control over the interventions to which the explanandum might be sensitive: they make it possible to isolate and manipulate one particular variable (charge sign), while at the same time keeping other variables invariant.

Moreover, note that the operation with fictitious systems offers a chance to explicitly examine

counterfactual scenarios that would otherwise not be available for actual manipulation. Not only do simulations allow us to infer the effects of such purely hypothetical interventions, they make such interventions actually possible in the first place.[13]

4.2 Granting epistemic access to the 'right' variables

The previous section has shown how atomistic simulations can extend the scope of possible explanatory investigations by allowing it to perform a wide range of highly controlled manipulations.

It is arguably of no surprise that the simulations used to study phenomena such as ions in aerosols are often called 'computational microscopes': They allow scientists to 'zoom in' and get a detailed picture of micro-level interactions. As regards the behavior of halide ions, a challenge was that the surface of droplets has the thickness of just a few molecules – and given the inhomogeneity and the disordered character of the interface region, it was previously hardly feasible to conduct experimental studies (Jungwirth and Tobias 2002; Garrett 2004). Atomistic simulations could help overcome such experimental boundaries, and – as the previous section has emphasized – they enabled scientists to perform highly controlled interventions on very small scales.

However, this is only one part of the story of their important explanatory role. Recall that the problem of halide surface preference is a problem about the occurrence of a *stable* molecular structure: It descibes a tendency of such ions, when being present in small droplets, to reliably migrate to the surface of these droplets. While some atomic details certainly matter for explaining such structures, others will not. Rather, the phenomenon becomes interesting in the first place because it remains stable under a range of possible variations in the microdetails of the system, and because it occurs frequently enough to potentially matter for chemical reactions in atmospheric aerosols.

As has been argued, an as detailed as possible story about the exact chain of molecular interactions leading to an ion's surface preference will never lead to a satisfactory explanation because it would contain an enormous amount of irrelevant details, and because it would fail to account for the stability of the phenomenon.

In recent discussions, it has been stressed that while a description that remains purely at the micro-level usually fails to account for the stability, or robustness of a phenomenon, determining the appropriate level to capture such variables is a non-trivial task (Woodward 2010, Batterman and Green 2021, Batterman 2021).

[13] There are many other interesting counterfactual interventions that can be performed by means of (atomistic) simulation. One can use atomistic simulations to artificially alter the strength of molecular bonds, the magnitude of electrostatic interactions or the sign of charges. Furthermore, simulations can be used to study scenarios that violate or by design manipulate the laws of nature (Humphreys 2004, 116).

Woodward, for example, emphasizes that the choice of an appropriate level of description can usually not be done in an a priori fashion. Rather, careful consideration is necessary in order to figure out the right level of description, i.e., a level specific enough to include explanatorily relevant details but coarse enough to neglect irrelevant detail:

> "The [...] idea that I want to defend involves the claim that, depending on the details of the case, [...] explanation can be either inappropriately broad or general, including irrelevant detail, or overly narrow, failing to include relevant detail. Thus, which level (or levels) is (are) most appropriate will be in large part an empirical, rather than a priori matter." (Woodward 2010, 296-297)

The case of halide surface preference illustrates this point: Citing the detailed path of molecular interactions will remain unsatisfactory, but so, too, will a too coarse-grained account. Remember that traditionally, the properties of solution surfaces were studied using continuum models that treat the solvent as a polarizable continuum rather than a set of individual molecules that exert particular kinds of interactions. Whereas such models come with the advantage of being applicable to large systems that would otherwise be difficult to trace, they neglect the precise molecular details of the system, such as the asymmetric solvation of positive and negative charges in water (Petersen and Saykally 2006, 344). Although relying on theoretical continuum models can certainly be useful if the aim is to investigate electrostatic interactions over larger molecular distances, these models reach their limits in situations in which molecular structure or the effects of short-range interactions matter (ibid.).

As concerns halide ions, simulations showed that the charge sign of these ions matters for the occurrence of surface preference, and that a model accounting for the asymmetrical distribution of charges in water molecules is needed. To the extent that continuum models fail to incorporate such details, they principally occur unsuitable for providing an explanation of the observed halide surface preference, as they do not offer sufficient access to detail (Caleman et al. 2011, 6840). Vice versa, one of the merits of atomistic simulations in tackling problems such as halide surface preference is that they enable access to variables such as the charge sign of an ion which figure at a finer-grained molecular level.

Still, while the charge of ions certainly is a local small-scale variable, some of the other variables that matter for the explanation of halide surface preference are not, especially not those accounting for the stability of the phenomenon. The variables which 'code'[14] for the stability of surface preference of large halides are thermodynamic ones, first of all the free energy – as it pro-

[14] We adopt this terminology from Batterman and Green 2021.

vides information about the point at which a system is in chemical equilibrium. By employing statistical-mechanical methods, Caleman et al. could use atomistic simulations as tools to bridge from 'small-scale' trajectorial information to such thermodynamic variables.

More specifically, by disentangling the (thermodynamic) free energy profiles into their respective enthalpic and entropic contributions[15], it became possible to obtain an even more detailed picture of how different energetic factors affect the explanandum.

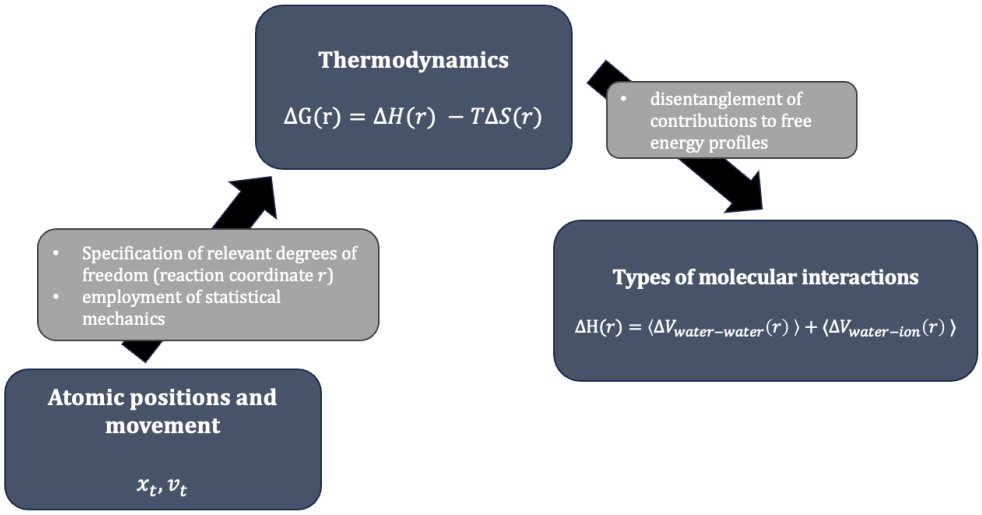

Figure 4.2: Molecular dynamics simulations as tools for accessing different levels of description.

And this is not even yet the full story. Rather than just assessing how changes in the enthalpy and entropy along the reaction coordinate matter for the explanandum, respectively; Caleman et al. further used this energetic information to establish a bridge *back* to variables described in the language of molecules and their behavior – namely the water-water and water-ion interactions.[16]. The interesting point here, however, is that, given that the focus is now on *types* of interactions, this should *not* be seen as a move back to particular molecules and their properties, but as a move to an *in-between* scale at which *averages* of molecular interactions are addressed.

Dealing with averages of molecular interactions importantly differs from dealing with detailed causal chains of molecular interactions: What is at stake now is how different *types* of molecular interactions contribute to a chemical reaction leading to a particular larger-scale molecular arrangement – the stability of which is accounted for in *thermodynamic terms*; and already reflected in the equation according to which these averages are derived. By considering averages along a specific reaction coordinate, the degrees of freedom have already been reduced to a limited number of

[15] According to Equation 4.2.
[16] Which together account for the enthalpy of the system, as expressed in Equation 3.2

relevant degrees; and the expectedness of the phenomenon is already *coded for* in the relationship that the different types of molecular interaction have to the free energy changes along this reaction coordinate.

Ultimately, the peculiar behavior of halide ions was explained by showing how the presence of these ions near the surface subtly affected a particular *type* of interactions occurring in the molecular system, namely water-water interactions, which made it energetically unfavorable for the ions to be in the bulk of droplets. On the one hand, some molecular variables *did* play an important role in the explanatory story provided by Caleman et al. – such as the charge sign which turned out to be an important difference-making factor that cannot not be abstracted away. On the other hand, much of the explanatory work relied on *filtering* from molecular-level details and on establishing a systematic connection between molecular details and thermodynamic variables, as well as on a connection between thermodynamic variables and variables capturing certain *types* of molecular interactions (see Figure 4.2).

Even if certain factors at a molecular level matter, fully capturing their explanatory significance required it to understand how they affect thermodynamic properties figuring at higher scales that code for the expectedness – or stability – of the explanandum. The energetic factors were then further analyzed to connect back to types of molecular interactions, which in turn helped illuminate the factors that are responsible for this expectedness. For all these steps, the relevant resources were provided by atomistic simulations.

That is, even though the variables which capture the stability or robustness of the explanandum typically *depend* on small-scale details, some of them are not themselves located on these smaller scales, but on *intermediary* scales between small and large. By employing statistical-mechanical methods, atomistic simulations turned out be powerful 'bridging' tools for explaining complex systems such as halide surface preference, as they allow both a systematic abstraction from atomic details and access to variables figuring at higher scales (cf. also Schweer and Elstner 2023). Besides merely extending traditional boundaries of scale and thereby increasing the number of systems that can be scientifically examined, atomistic simulations could explicitly grant access to variables that code for the stability or robustness of complex phenomena.

4.3 Providing a framework for quantitative analysis

We have outlined how, in the case of halide ions in aerosols, atomistic simulations could enable a more detailed account of the structural peculiarities at the interface region, and of the complex interplay of various energetic factors that give rise to the observed ion behavior.

Even before simulations had entered the stage, it was of course to be expected that the behavior of ions – just as in the case of other complex systems – is determined not just by one or two major factors but by the fine-tuned interplay of various contributors. Thus, explanatory interests typically concern the precise *extent* to which each factor matters. As Caleman et al. (2011, 6838) state, progress in force fields that underlie atomistic simulations "finally allows establishment of the surface preference for all halide and alkali ions quantitatively."

Our examination has shown that a major part of what made atomistic simulations so explanatorily important was that they enable a quantitative comparison of the effect of the different energetic factors and underlying molecular interactions for surface preference. Large halide ions prefer to be solvated near the surface of droplets because their properties are such that the gain in water-water interactions *quantitatively* dominates over the energetic effects of a loss in water-ion interactions. While it is clear that the free energy minimum of the ions' path towards the surface is changed by the effect of various energetic contributions, insight into how sensitive the explanandum is to these contributions is to be acquired by means of quantitative comparison. Such a detailed comparison is hardly possible by means other than simulation and further stresses their explanatory power.

Thus, besides granting access to explanatory-relevant variables, and besides making it feasible to perform manipulations on them, atomistic simulations can contribute to explanations by helping analyze the *size* of the contributions of different types of molecular interactions to the free energy profiles.

5 Conclusion

In this paper, we have examined how atomistic simulations contribute to scientific explanations. We have drawn on a counterfactual account of scientific explanation in the tradition of Woodward (2003), and as it has been further developed for explanations with the help of models by Bokulich (2011). As we suggest, such an account provides a realistic picture of the characteristics of many explanations in science, and it can accommodate various challenges that occur when the behavior of complex systems is at stake.

Concretely, we have outlined three ways in which atomistic simulations contribute to scientific explanations, namely by (1) extending the scope of possible explanatory investigations and enabling scientists to effectively perform and track manipulations that would otherwise oftentimes remain unfeasible, by (2) offering epistemic access to variables at relevant levels of description, and by (3) offering a framework for quantitative analysis.

By focusing on how atomistic simulations could concretely support explanatory reasoning in

scientific practice, our examination contributes to a more comprehensive understanding of their explanatory merits. Although our examination was centered around a concrete example from physical chemistry and a particular type of simulation (atomistic molecular dynamics), much of what we have shown may apply to other cases that are concerned with complex systems: Computer simulations can be tools for systematically assessing how the explanandum is sensitive to a range of variations in the respective system – often in ways that would otherwise remain unfeasible; they can help systematically draw inferences about variables that code for the stability or robustness of the behavior of systems; and they can help determine in a fine-grained manner the relative importance of particular variables.

Bibliography

Batterman, R. W. (2001). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. New York: Oxford University Press.

Batterman, R. W. (2019). Universality and RG explanations. *Perspectives on Science*, 27(1):26–47.

Batterman, R. W. (2021). *A Middle Way: A Non-Fundamental Approach to Many-Body Physics*. Oxford University Press.

Batterman, R. W. and Green, S. (2021). Steel and bone: mesoscale modeling and middle-out strategies in physics and biology. *Synthese*, 199(1-2):1159–1184.

Bechtel, W. and Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3):321–333.

Bokulich, A. (2008). Can classical structures explain quantum phenomena? *The British Journal for the Philosophy of Science*, 59(2):217–235.

Bokulich, A. (2011). How scientific models can explain. *Synthese*, 180:33–45.

Bokulich, A. (2014). How the tiger bush got its stripes: 'how possibly' vs. 'how actually' model explanations. *The Monist*, 97(3):321–338.

Bokulich, A. (2017). Models and explanation. In Magnani, L. and Bertolotti, T., editors, *Springer Handbook of Model-Based Science*, pages 103–118. Springer.

Caleman, C., Hub, J. S., van Maaren, P. J., and van der Spoel, D. (2011). Atomistic simulation of ion solvation in water explains surface preference of halides. *Proceedings of the National Academy of Sciences*, 108(17):6838–6842.

Garrett, B. C. (2004). Ions at the air/water interface. *Science*, 303(5661):1146–1147.

Hitchcock, C. (1999). Contrastive explanation and the demons of determinism. *The British Journal for the Philosophy of Science*, 50(4):585–612.

Hollingsworth, S. A. and Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143.

Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press.

Jungwirth, P. and Tobias, D. J. (2002). Ions at the air/water interface. *The Journal of Physical Chemistry B*, 106(25):6361–6373.

Lipton, P. (1990). Contrastive explanation. *Royal Institute of Philosophy Supplements*, 27:247–266.

Petersen, P. B. and Saykally, R. J. (2006). On the nature of ions at the liquid water surface. *Annual Review of Physical Chemistry*, 57:333–364.

Rice, C. (2021). *Leveraging Distortions: Explanation, Idealization, and Universality in Science*. MIT Press.

Roberts, S. T., Ramasesha, K., and Tokmakoff, A. (2009). Structural rearrangements in water viewed through two-dimensional infrared spectroscopy. *Accounts of Chemical Research*, 42(9):1239–1249.

Ross, L. and Woodward, J. (2023). Causal Approaches to Scientific Explanation. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Schweer, J. and Elstner, M. (2023). Dealing with molecular complexity. atomistic computer simulations and scientific explanation. *Perspectives on Science*, 31(5):594–626.

Stöckler, M. (2000). On modeling and simulations as instruments for the study of complex systems. In Carrier, M., Massey, G. J., and Ruetsche, L., editors, *Science at Century's End: Philosophical Questions on the Progress and Limits of Science*, pages 355–373. University of Pittsburgh Press.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.

Tobias, D. J., Stern, A. C., Baer, M. D., Levin, Y., and Mundy, C. J. (2013). Simulation and theory of ions at atmospherically relevant aqueous liquid-air interfaces. *Annual Review of Physical Chemistry*, 64:339–359.

Woodward, J. (1997). Explanation, invariance, and intervention. *Philosophy of Science*, 64(S4):S26–S41.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, 25:287–318.

Woodward, J. (2016). The problem of variable choice. *Synthese*, 193:1047–1072.

Woodward, J. and Hitchcock, C. (2003). Explanatory generalizations, part I: A counterfactual account. *Noûs*, 37(1):1–24.

Ylikoski, P. (2007). The idea of contrastive explanandum. In Persson, J. and Ylikoski, P., editors, *Rethinking Explanation*, pages 27–42. Springer.

Ylikoski, P. (2014). Agent-based simulation and sociological understanding. *Perspectives on Science*, 22(3):318–335.

Ylikoski, P. and Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies*, 148(2):201–219.

Zednik, C. (2015). Heuristics, Descriptions, and the Scope of Mechanistic Explanation. In Braillard, P.-A. and Malaterre, C., editors, *Explanation in Biology*, volume 11, pages 295–318. Springer Netherlands, Dordrecht. Series Title: History, Philosophy and Theory of the Life Sciences.

Multiscale Simulations, Reduction, and the Mediatory Role of Models

Contents

Summary

Multiscale simulations have become important means for studying scientific problems across multiple time and length scales. In the philosophy of science, the role of such simulations in tackling multiscale problems has sparked considerable interest.

Several authors have argued that the success of multiscale simulations challenges a traditional reductionist view of how connections between different theories and their models work. Winsberg and Bursten, for example, point out that in the case of multiscale simulations, bridging

between different levels of description turns out to be a matter of empirical and physical (Winsberg) or even conceptual (Bursten) reasoning rather than logical or mereological insight. While being sympathetic to these suggestions, I propose that the way in which different levels of description are combined in multiscale simulations deserves further attention. In this chapter, I sketch a number of recent philosophical contributions to the question of how different levels of description can be combined in multiscale simulations and I offer connecting points for further discussions. Concretely, by drawing on an approach famously articulated by Morgan and Morrison, I examine how models in multiscale simulation contexts act as partly autonomous mediators between theories and the world. Based on that, I explore how recognizing the mediatory role of models may enhance an understanding of how different levels of description can be integrated and 'work together' in multiscale simulation frameworks.

1 Introduction

When corrosion occurs in metals, several scientifically interesting processes occur at different time and length scales: Whereas electron transfer events at the surface area take place at a time span of a few femtoseconds, the building up of oxide layers occurs within several minutes (Raabe et al. 2009). Finally, at a macroscopic scale, the corrosion and decomposition of materials is affected by environmental factors such as rainfall, and it can take several decades or more (Raabe et al. 2009; Cole and Marney 2012). Metal corrosion gives an example of a multiscale problem, i.e., a problem that asks us to pay attention to processes occurring and interacting at multiple scales.

Over the last years, sophisticated multiscale simulations have led to advances in the scientific study of multiscale problems. Rather than looking at just one level of description, multiscale simulations enable scientists to study how processes occurring at lower levels of description affect those at higher levels of description, and vice versa. A particularly exciting type of multiscale simulations are 'concurrent' simulations, i.e., simulations in which the respectively investigated problem is divided into different domains at multiple levels of description which are simultaneously modeled and coupled to one another.

In the philosophy of science, the use of multiscale simulations has sparked discussions on the possible relationships between different scales or levels of description. On the one hand, it seems widely agreed on that the success of multiscale simulations is at odds with a naïve reductionist understanding of how different scientific levels of description can be connected to each other (cf. Winsberg 2006a; Winsberg 2010; Green and Batterman 2017; Bursten 2018; Batterman 2021; Batterman and Green 2021). Rather than simply deriving higher-scale descriptions from bottom up, multiscale simulations rely on sophisticated methods and strategies that establish connections

between different levels of description in a fine-tuned and often multi-directional way.

On the other hand, even if a naïve reductionist view does not do justice to how different levels of description are connected to each other, the question remains of how the working together of different levels of description in multiscale simulation contexts can better be understood. In some sense, the success of multiscale simulations poses an interesting puzzle: On the one hand, it seems inadequate to understand the working together of different levels of description in a rather naïve reductionist fashion, i.e., by assuming that descriptions at higher levels could simply be derived from descriptions at lower levels. On the other hand, scientists' increasing use of multiscale simulations underscores that rather being largely separate from one another, the close combination of different levels of description seems not only possible but integral for tackling multiscale problems.

This puzzle becomes even more striking if we consider a diagnosis by Eric Winsberg who states that "parallel multiscale models [...] are sewn together using specially constructed algorithms to mediate between otherwise incompatible frameworks"(Winsberg 2010, 73) or that multiscale models are "are – all at once – models of an inconsistent set of laws" (Winsberg 2006a, 593). How can multiscale simulations integrate otherwise incompatible frameworks, and how can they rely on models from 'inconsistent laws'?

Taking this puzzle serious, the ways in which different levels of description can be combined in multiscale simulations deserve further philosophical clarification. In this chapter, I draw on a seminal approach by Mary Morgan and Margaret Morrison who consider models as partly autonomous mediators between theories and the world. I examine how the 'models as mediators' approach resonates with the context of multiscale simulation. Concretely, I suggest that recognizing how models in multiscale simulation contexts are mediators between theories, other models and the world can contribute to a better understanding of the communication between different levels of description in multiscale simulation contexts.

To begin, I will, in section 2, briefly sketch how philosophers of science have taken multiscale problems and the success of multiscale simulations as an occasion to criticize a traditional (and rather crude) reductionist understanding of how different levels of description can work together. In section 3, I will look in more detail at a concrete example, namely that of a multiscale simulation about crack propagation in solid materials. The example has been discussed extensively by Winsberg and Bursten. Besides criticizing a reductionist view of how different levels of description can work together, Winsberg and Bursten have also drawn positive lessons from the example of crack propagation, namely that successful communication between different levels of description is a matter of empirical and physical (Winsberg) or even conceptual (Bursten) strategies rather than logical or mereological insight. In section 4, I discuss these positive suggestions by Winsberg and

Bursten and propose that an interesting question deserves further attention – namely what the use of these strategies tells us about the character of models in the context of multiscale simulations. I introduce Morgan's and Morrison's 'models as mediators' approach and explore how considering models in multiscale simulation contexts as mediators can contribute to an understanding how different levels of description can work together in multiscale simulation contexts. Section 5 concludes and offers an outlook.

2 Challenging a Traditional Reductionist View

According to a traditional reductionist view, the behavior or properties of systems[1] at higher scales – be it organisms in biology or materials in materials science – should ideally be derivable from properties of these systems at small scales. Being concerned with the properties of materials across scales, Green and Batterman state that

> " [i]deally, from a reductionist perspective, one would like to be able to determine the values for the material parameters at a more fundamental scale and model the system "bottom-up" in one coherent model capturing the 'sum of the parts'." (Green and Batterman 2017, 25)

Opposing this reductionist perspective, Green and Batterman argue that upper-scale properties normally cannot simply be obtained in a bottom-up fashion.

2.1 Arriving at Wrong Parameter Estimates

As an example, they consider the case of steel. According to Green and Batterman, the dynamics of steel at higher scales cannot be obtained from structural models at the microscale (Green and Batterman 2017, 32). The idea is the following: At a molecular scale, one can think of steel as consisting of atoms that are arranged in a regular lattice-like fashion (Green and Batterman 2017, 22). Now if we were to assume that steel is nothing but a uniform lattice of atoms, and if we tried to directly infer from this lattice structure the properties of steel at higher scales, we would arrive at incorrect estimations of these properties (Green and Batterman 2017, 32). In another paper dealing with the epistemic significance of constraints and parameters at intermediate scales, the two authors present their point in more detail:

[1] The discussion frequently makes use of the notion of 'systems' which is, however, not epistemically innocent. When speaking of 'systems of scientific interest' or of 'physical systems', we address the objects of investigation in a particular way; we presuppose having an idea of what is part of the 'system' and what is not, and we specify *how* the object of investigation can be addressed. I think that this indicates that from the beginning on, possible suggestions regarding how bridges between levels at which these 'systems' are described presuppose a reflection on what features or aspects of the world ought to be captured, as I will outline later.

> "A reductionist/fundamentalist perspective [...] would insist that since steel is composed of nothing but its atomic constituents, a detailed microscale description should suffice for the prediction of the macroscale properties of steel, including values for material parameters. This is not the case. A model at the lowest scale (just as an upper-scale continuum model) would not be able to see the microstructures—the voids, cracks, grain boundaries, etc.—that exist at intermediate scales. Accordingly, a direct bottom-up approach to steel would arrive at wrong estimates for the elastic properties of steel, because the relevant parameters cannot be estimated at the smallest scale [...]." (Batterman and Green 2021, 1166)

Thus, the idea is that when proceeding in a reductionist fashion, we would be unable to take account of the relevant features of steel at an intermediate scale, and since these features matter for the overall behavior of steel, a direct bottom-up approach would eventually not be able to get the properties of steel at higher scales right.

An even finer-grained illustration of why a straightforward bottom-up approach seems unsuited for describing or explaining the behavior of steel is offered by Wilson (2017, 221ff.). Concretely, Wilson invites us to consider the hysteresis of steel. Think of a train that runs over a steel railroad track. After some time of being exposed to stress by trains that run over it, the steel track will eventually become brittle. However, Wilson suggests that if we tried to explain the overall brittleness of steel tracks from bottom up without accounting for irregularities in the atomic lattice which show on intermediate scales, we would overestimate its sensitivity to stress and assume that it is more prone to deformation than it really is (cf. Wilson 2017, 221).

Like Green and Batterman, Wilson holds that the reason for this overestimation is that an important feature of steel only becomes 'visible' at an intermediate scale between the molecular and the macroscopic scale, namely the presence of dislocations. Dislocations are slight irregularities in the atomic lattice structure that accumulate at an intermediate scale between the atomic-lattice scale and the macroscopic scale and 'protect' individual molecular bonds from forces exerted by the imposed stress (Wilson 2017, 221). As Wilson argues, "we cannot develop an adequate account of rail hysteresis working upwards from the molecular scale in a naïve manner" (Wilson 2017, 221). Only if dislocation effects at intermediate scales are considered, an adequate description of the brittleness of steel can be obtained. Furthermore, Wilson suggests that steel hysteresis more generally can only be appropriately explained once the capacity of dislocations to dissipate energy from exerted stress is taken account of.

Thus, a shared diagnosis by Green, Batterman and Wilson seems to be that when trying to

explain why steel at higher scales exerts the behavior that it does, we cannot proceed in a straightforward bottom-up fashion, because doing so would lead to wrong estimates regarding the behavior of the material under stress.

Wilson demonstrates in further detail how communication between the different sub-models in a multiscale simulation context regarding steel hysteresis could look like (Wilson 2017, 221ff.). According to Wilson, a possible computational architecture integrating descriptions of steel at various scales would contain sophisticated feedback loops between the respectively involved models. Rather than simply transferring information upwards, models at different scales could dynamically exchange information and trace how processes occurring on the respective scales affect each other: At a macroscopic scale, we could model how stresses from the load of trains distribute across the steel bar. Then, we could 'zoom in' on the regions at which these stresses have particularly significant effects and investigate on smaller scales how the structure of the material eventually changes in response to the higher-scale load. If we notice, for example, that the drift[2] of dislocations at an intermediate scale between micro and macro reaches a certain threshold, this information could be passed back to the higher scale in terms of a report of decreased fracture strength (Wilson 2017, 224).

2.2 Questioning the 'Superiority' of Models at Lower Levels of Description

According to Batterman, Green and Wilson, the behavior of steel cannot be derived in a straightforward fashion from a molecular-level model that describes steel as a regular lattice of atoms and ignores characteristic structures and heterogeneities of the material at intermediate levels. Franklin, however, has noted that Batterman, Green and Wilson paint a rather simplistic picture of the reductionist (Franklin, forth): It is not clear why someone striving for a bottom-up explanation of steel behavior should simply grant the material is quite homogeneous at an intermediate level, and in a certain sense, dislocations may even be 'present' at a lower level of description, as they correspond to slight local disturbances of the lattice arrangement that travel through the material (Franklin, forth).[3]

Yet, even if the picture of the reductionist painted by Batterman, Green and Wilson is a rather simplistic one, it is arguably not their main concern to argue against someone who would really defend that the behavior of steel can be explained in a straightforward fashion from bottom-up. Rather, I assume that the main target of the presented criticisms by Batterman, Green and Wilson is

[2] If dislocations eventually drift after a series of loading, this will, at some point, decrease their capacity to protect molecular bonds from being affected by applied stress (Wilson 2017, 224).

[3] Franklin actually develops an argument according to which the behavior of steel on macroscopic scales *can* indeed be explained based on factors that figure at the molecular scale, see Franklin (forth).

a widespread intuition according to which descriptions at lower levels of description are 'privileged' over or 'superior' to those at higher levels of description (cf. Batterman 2021). Even if there is agreement that obtaining bottom-up descriptions or explanations is in many cases unfeasible, descriptions at lower levels would, according to this intuition, under ideal circumstances, still be considered as more complete or privileged over higher level descriptions.

As I have indicated, all three authors oppose this intuition by emphasizing the epistemic and methodological relevance of descriptions at intermediate scales or levels of description. In their 2020 article, Batterman and Green offer a detailed discussion of how the reference to factors at higher levels (particularly those figuring at intermediate levels between micro and macro) is often not primarily a matter of computational convenience or mathematical feasibility, but crucial for capturing the relevant functional or organizational features of complex systems (see Batterman and Green 2021). As they put it:

> "A common challenge for multi-scale modeling is to understand the relationship between continuum dynamics at macroscale and material-specific parameters that code for complex and heterogenous structures at lower scales. Mesoscale structures often provide a means for determining such parameters." (Batterman and Green 2021, 1181)

That is, since parameters described at intermediate scales often 'code' for relevant properties of complex systems, it seems not unproblematic to say that descriptions at small scales are superior to or privileged over descriptions at higher scales.

I agree, and while I think that the criticisms do not necessarily rule out that there can, in some cases, be viable bottom-up modeling strategies, the authors emphasize that higher levels of description often serve a different purpose rather than just offering simplifications from lower scales. The perspective suggested in the remainder of this chapter will underline this idea.

3 Learning from Multiscale Simulations

So far, an impression has been given of how philosophers of science have taken multiscale models as an impetus to criticize a traditional reductionist understanding of how descriptions at different levels are related to each other. In what follows, I will look at an example from materials science that has been discussed in more detail by Winsberg (2006a, 2010) and Bursten (2018). Besides providing further support for the aforementioned criticism of reductionism, Winsberg and Bursten also make positive suggestions regarding how the combination of different levels of description can better be conceptualized, and I here take their contributions as a starting point for exploring the

mediatory role of models in contexts of multiscale simulation.

3.1 The Case of Silicon Cracks

The example both Winsberg and Bursten are concerned with is that of a multiscale simulation of emerging cracks in silicon, developed and presented by Abraham, Broughton et al. in the late 1990s. The silicon crack study was designed specifically to explore a particular computational methodology, namely the concurrent coupling of various length scales in a multiscale simulation framework (Abraham et al. 1998, 783). Being interested in studying how cracks propagate through silicon, the scientists separated the material into three spatial regions, addressing each with a different modeling approach and linking them by means of tailored algorithms.

The Three Sub-Models

In the region where the tip of the emerging crack in silicon is located, a semi-empirical tight binding (TB) approach was used (Abraham et al. 1998, 784; Broughton et al. 1999, 2392). TB is a quantum mechanical method which provides a simplified picture of electron motion in solids and which describes a material's electronic structure. The TB method is itself based on another quantum mechanical approximation, namely the Born-Oppenheimer approximation (Winsberg 2006a, 588). In the case of silicon cracks, the TB model could take account of the breaking of bonds in the area of the emerging crack (Broughton et al. 1999, 2392).

The region surrounding the crack tip was modeled by means of a molecular dynamics (MD) approach. The central ingredient to a conventional MD simulation is an interatomic potential that captures the relevant kinds of interatomic interactions in a system and allows it to trace the motion of the atoms of the system (Abraham et al. 1998, 784).

Note that the choice and parametrization of the respectively used potential depends on the kind of material one deals with. In the context of the silicon crack simulation, the chosen potential was the Stillinger-Weber potential (Abraham et al. 1998, 784), which is a typical potential for certain covalently bonded systems, specifically silicon. Starting with a description of the interactions between the atoms, MD simulations allow it to calculate the forces acting on the atoms and, based on that, to numerically solve Newton's equations of motion for the atoms of a given system, thereby generating the evolution of the system in time. In the case of the silicon crack simulation, such an approach can account for how atoms rearrange themselves as the crack propagates through the material, and it can capture the effect of thermal fluctuations and pressure waves stemming from bond rupturing or the defective region (Broughton et al. 1999, 2392).

To describe the region farer from the tip, the finite element (FE) method was used which is based on continuum mechanics and linear elastic theory. In elastic theory, materials are addressed as a continuum media which are subject to mechanical stress or strain; and the FE method approximates the energetic properties of a given continuous material by treating the material as a 'mesh' consisting of a finite amount of mesh points whose distribution changes in response to transitions in displacement and strain (Bursten 2018, 159). By means of the FE method, one can construct a Hamiltonian approximating the elastic behavior of the material in the region far from the crack.

Handshake between FE and MD

To model the interface region between the FE and MD region, Abraham, Broughton et al. defined a plane that separates both domains from one another (Broughton et al. 1999, 2398). Essentially, the bridge between the two regions was built by establishing a "one-to-one mapping of a mesh point to an atom site" (Broughton et al. 1999, 2396). Technically, this is possible because in FE theory, the orientation of mesh elements can be varied and here adjusted so as to match with the atoms of the MD region (Bursten 2018, 160). The energetics at the interface region are then mathematically accounted for two times, one time treating the interface as consisting of FE mesh points and one time treating it as consisting of MD atom sites. The average of the obtained result gives the energy of the system at the interface region (see Winsberg 2006a, 589). The proposed "50/50" handshaking Hamiltonian can capture both types of contributions because it is assumed that atoms and mesh points can be treated interchangeably: "The SW energy formulation that concentrates upon atomic coordinates [...] and the FE energy formulation that concentrates upon displacements [...] can be used throughout the interface because of the indistinguishability of what are atoms or mesh points" (Broughton et al. 1999, 2396).

Handshake between MD and TB

In the interface region between MD and TB, no "50/50" Hamiltonian like that in the FE/MD region is possible (Broughton et al. 1999, 2398). The reason is that the TB Hamiltonian does not calculate the energy locally and does not offer a straightforward way of assigning energy to individual interatomic bonds (Winsberg 2006a, 590). The scientists therefore used a trick: They cut off 'dangling' bonds by changing the properties of the atoms that 'sit' at the edge of the QM region such that they behave like hybrid monovalent 'hydrogenic' silicon atoms (Broughton et al. 1999, 2398; Winsberg 2006a, 590).

These 'silogens', as they named them, have the same bond energy as a single silicon-silicon

bond and the same longitudinal force constant as silicon, but the electronic symmetry of hydrogen (Broughton et al. 1999, 2398). Since silogens do not energetically interact with their neighbors, the energetic contributions in the TB region can now be localized (Winsberg 2006a, 590). To bridge between both regions, the bonds to an atom at the TB-side of the interface region can be obtained from the TB Hamiltonian and the bonds on the MD-side can be obtained from the Stillinger-Weber potential (Abraham et al. 1998, 785).

3.2 Philosophical Discussion

According to Winsberg, the silicon crack simulation appears "to be at odds with some basic philosophical intuitions about the relationships between different theories and between theories and their models" (Winsberg 2006a, 590-91). Concretely, Winsberg suggests that the scale-bridging methodology in the silicon crack simulation puts under pressure a traditional reductionist idea according to which the bridging between different levels of description is a matter of logic or mereology. As Winsberg puts it:

> "One issue that has received perennial attention from philosophers of science is that of the relationship between different levels of description. Traditionally, the focus of this inquiry has been debate about whether or not, and to what extent or in what respect, laws or theories at higher levels of description are reducible to those at a lower level. Underlying all of this debate, I believe, has been a common intuition: the basis for understanding interlevel interaction – to the extent that it is possible – is just applied mereology. In other words, to the extent that the literature in philosophy of science about levels of description has focused on whether and how one level is reducible to another it has implicitly assumed that the only interesting possible relationships are logical ones – that is, inter-theoretic relationships that flow logically from the mereological relationships between the entities posited in the two levels. But if methods that are anything like those described above [in the crack example, J.S] become accepted as successful in nanoscale modeling, that intuition is likely to come under pressure." (Winsberg 2010, 84)

Taking up Broughton, Abraham et al.'s own wording, Winsberg contends that finding the appropriate bridge between two levels of description requires 'physical insight' – rather than logical or mereological insight (Winsberg 2006a, 592).

As Winsberg puts it, the developed relationships might well be "suggested" by their mereology, but they are certainly not determined by it. Rather, "whether or not the relationships employed by

Abraham and his group will turn out to be the correct ones is an empirical/physical question and not a logical/mereological one" (Winsberg 2006a, 592).

Bursten goes one step further and argues that the example of silicon cracks illustrates how bridging between different levels of description predominantly depends on the employment of *conceptual* strategies rather than on logical, mereological or even physical reasoning. Concretely, she demonstrates that in the example of the silicon crack simulation, two different conceptual scale-bridging strategies were employed.

The first one, which she calls the "MNRF" (manipulating non-representational features) strategy, figures at the FE/MD handshake. According to Bursten, the non-representational feature that is being manipulated in the FE/MD handshake is the mesh points (Bursten 2018, 160). The idea is the following: Aligning the mesh points with the size of the atoms in the MD region is viable because the regions where mesh points figure is the region where continuum mechanics applies. In continuum mechanics, materials are treated as continuous media rather than as consisting of atoms. Hence, even though mesh points are a way of 'applying' principles from continuum mechanics within a computational modeling framework, it is clear that the mesh points themselves do not represent any actual feature of the material (cf. Bursten 2018, 160). This allows it to manipulate them such that one can build a bridge between two conceptually distinct descriptions of the material. In Bursten's own words:

> "The problem that this handshake faced was that there was a fundamental conceptual mismatch between the FE and MD component models, namely that one was about continuously distributed matter and the other was about a discrete set of atoms in a lattice. The solution to this problem, the MNRF strategy, was to use this mismatch to the model's advantage, by moving around the mesh points in the FE model so that they aligned with the lattice spacing of the MD model." (Bursten 2018, 164)

The other handshake occurring at the interface between the TB and MD region depends on what Bursten denotes as the construction of a "contrived, un-physical fiction" (Bursten 2018, 161). The creation of 'silogens' allows it to get a localized description of the energy in the TB region where energetic interactions are otherwise delocalized. Bursten emphasizes that silogens are neither idealizations of realistic atoms, nor straightforward abstractions (Bursten 2018, 165). That is because they could not be de-idealized into a more realistic entity (Bursten 2018, 161, footnote).

While Bursten agrees with Winsberg that physics play an important and constraining role in the mediation between the different scales in the silicon crack simulation, she contends, however, that this is not the most crucial point:

> "There is certainly significant back-and-forth between the models, but it is not a straightforward exchange of energy or other strictly empirical determinables; it is an exchange of ideas constrained and moderated by the computational, mathematical, and yes, also, physical limits of the component models' frameworks. This exchange certainly requires more than mere logic, or mereology. But it also requires more than just an understanding of the physics, since mesh points and silogens are not physical objects." (Bursten 2018, 163)

Bursten's main point seems to be that while the relevant physics certainly *constrain* how possible bridges can be built, this is not sufficient for determining how exactly such a bridge will *look like*. Rather, bridging was a matter of conceptual reasoning, i.e., of reasoning how certain features of the model are not intended to represent something in the world (or are even 'fictional') and how they can, therefore, be adjusted so as to enable a seamless bridge from one modeling region to the other.

4 Bridging between Levels of Description

Winsberg and Bursten both stress that to successfully establish connections between different levels of description, the scientists of the silicon crack simulation could not simply logically infer descriptions at higher levels from those at lower levels. Rather, what was needed to bridge between different levels was physical, empirical or even conceptual insight.

While I am sympathetic to these conclusions, they are perhaps – in some regard – not really surprising. The reason is that empirical, pragmatic, or conceptual considerations come into play already *before* bridges between the different models are established, as I will discuss below. Winsberg and Bursten are certainly aware of this[4], and I do not aim to criticize their contributions regarding how bridges between different levels are established. Rather, I wish to build on these contributions and bring to light explicitly that rather than just underscoring the shortcomings of a (naïve) reductionist view, the silicon crack simulation can illustrate the mediatory role of models regarding both theories and the world.

Differently put, if already the choice and specification of the sub-models in a multiscale simulation context is, to some extent, a matter of empirical and pragmatic reasoning, it seems to be expected that this looks similar when it comes to the crafting of bridges between these models. At

[4]Winsberg, for example, actually explicitly endorses in another paper that computational models are, in the sense emphasized in this chapter, *semi-autonomous* and that working with them means making *transformations* and deviating from theory (Winsberg 2006b, 6). Thus, the contribution in this chapter is not to challenge Winsberg or Bursten's positive suggestions, but to underscore that they enable insight into something more interesting than the failure of naïve reductionism, namely the way in which models have a mediatory role in (multiscale) simulations.

the same time, it is interesting to explore as to what the concrete bridging strategies employed by the scientists of the silicon crack simulation may tell us about the role and character of models and their theories in multiscale simulation contexts.

4.1 Theoretical and Computational Models

When stating that the silicon crack simulation consists of different 'sub-models' or 'modeling regions', the use of the term 'model' requires further elaboration. When speaking of 'models' in the context of computer simulations, we can distinguish between 'theoretical' or 'conceptual' models on the one hand and 'computational' models or 'simulation' models[5] on the other hand.[6]

A theoretical model captures certain relevant properties of the respective object of investigation by drawing on resources from theory, often in the form of mathematical equations. In the case of the neighboring region of the emerging crack in silicon, for example, Broughton, Abraham et al. modeled the material as consisting of a lattice of atoms whose motion is approximatively described by Newton's equations of motion and whose interactions give rise to the dynamic behavior of the material in the respective region.

Theoretical models are certainly informed by theories of the dynamic behavior of matter, but they are not straightforwardly derived from them, because developing a theoretical model means making more concrete assumptions about the object of investigation and about the properties that are deemed relevant for the concretely-addressed problem or question, as I will outline in more detail later.

Once it is specified that the neighboring region of the crack should be described in terms of a lattice of atoms whose motion is approximated by Newton's equation of motion, the so-understood theoretical model still needs to be brought in a form that can be handled within a computer simulation framework. In the case of the silicon crack simulation, the Stillinger-Weber potential was chosen to reflect the interactions between the atoms, and together with an atomic starting configuration, algorithmically implementing the interatomic potential allowed it to numerically integrate Newton's equation of motion for the atoms, thereby approximatively tracing the dynamics of the material as the crack propagates.

When speaking of the three 'sub-models' in the silicon crack simulation, I will in the remainder of

[5] I will use these terms interchangeably here.

[6] In this terminology, the computational model coincides with what is often called a 'numerical model'. Sometimes, an even finer-grained conceptual distinction is made, and the term 'computational model' is used for the *algorithmically specified and implemented* numerical model. The idea of this finer-grained distinction is to highlight that besides needing discretization for numerical treatment, models in the context of computer simulations need to be algorithmically implemented – and this implementation can involve significant transformations (Boge 2019, 3). In this chapter, I will stick to the distinction between theoretical models and computational models; and by the latter I mean the discretized numerical models that are based on theoretical models and that can be implemented in a certain programming language by means of suitable algorithms (cf. also the introduction of this thesis).

the chapter be concerned with *computational models*, i.e., models which approximate and translate the content of theoretical models so as to bring it in a numerical form, allowing for a computational implementation and treatment.

Just as theoretical models are not straightforwardly derived from theories, computational models are not straightforwardly derived from theoretical models. Rather, careful decision-making and evaluation are inevitable at various points of our modeling activities[7]:

> "Not only does one have to worry about whether the theoretical model one starts with accurately represents the target system, but the discretization process required for making the theoretical model amenable to numerical solution can involve significant changes in the informational content of the original model. So, in addition to worries about whether the model has actually been solved correctly, there is the problem of determining the extent to which the solved simulation model can be taken as a more or less accurate representation of the target system." (Morrison 2015, 9)

The quote also suggests that the development of both theoretical and computational models requires careful thought and consideration regarding the features of the "target" we aim to address. This is a facet of them being mediators, as I will outline in the following.

4.2 Models as Mediators

Morgan and Morrison have famously suggested to understand models as 'mediators' between theories and the world (Morgan and Morrison 1999). Departing from the semantic tradition according to which theories are essentially *families of models* (Morrison 2007, 198), they suggest that models are based on and informed but not dictated by scientific theories, and that they have a mediatory and partly autonomous role between theories and the world:

> "[T]heory does not provide us with an algorithm from which the model is constructed and by which all modelling decisions are determined. As a matter of practice, modelling always involves certain simplifications and approximations which have to be decided independently of the theoretical requirements." (Morrison and Morgan 1999, 16)

The autonomous character of models concerns both their *construction* and *functioning* (Morrison and Morgan 1999, 12ff.). Neither can models straightforwardly be 'read off' from observational

[7] This point is also stressed by Lenhard (2014, 355) who deals with computational modeling in quantum chemistry and states that "[...] ab-initio methods (typically, though probably not always) rely on autonomous modeling steps that do not have a quantum theoretical rationale". Lenhard's position strongly resonates with the perspective developed in this paper. Arguing against someone who might claim that computational ab-initio methods have made true the dream of derivation from first principles, Lenhard states: "computational modeling, [...] is not a matter of theoretical derivation but rather a matter of mediation between theory, experiment, phenomena, and computational technology" (Lenhard 2014, 349).

data, nor can they simply be deduced from scientific theories (Eckert and Hillerbrand 2022, 139). Instead, the construction of models involves – as I will illustrate – making assumptions that go beyond what is incorporated in the respective theories, and it requires more than having observational data of a certain phenomenon. Also, since models are constructed partly independent from theory and data, they can autonomously fulfill various functions. For example, models can function as means for applying theories to particular situations by bringing abstract theoretical equations in a more concrete form, or they can be vehicles for experimentation, among other things (Morrison and Morgan 1999, 20).

To illustrate how the construction (and evaluation) of models requires more than what is comprised in the theories upon which they are based, consider the neighboring region of the crack in the silicon crack simulation. At the point at which the interatomic potential in the MD region is chosen, for example, pragmatic and empirical factors played a role. That is, the interatomic potential was not straightforwardly derived from theory of molecular dynamics. Interatomic potentials are – as indicated earlier – chosen specifically for certain classes of materials. Recall that in the silicon crack simulation, the overarching idea was to demonstrate the methodology and application of a concurrent multiscale simulation framework, using the case of silicon cracks as an example. As the scientists (1999, 2393) put it, silicon is a material for which many good empirical interatomic potentials exist, and they state: "we chose the Stillinger-Weber because of its computational simplicity" (Broughton et al. 1999, 2393). Interatomic potentials such as the Stillinger-Weber potential give an approximation of the potential energy of the system, and in the context of a simulation, parametrized interatomic potentials together with a suitable starting configuration can be means for studying a certain material beginning with atoms and their interactions.

Both the choice and parameterization of an interatomic potential depend not only on the material itself but also on the concrete question or problem at stake. For example, the Stillinger-Weber potential can reconstruct the elastic properties in the diamond structure of silicon, but it is considered to reach its limits when reproducing the brittleness of silicon under tension, among other things (Müser et al. 2023, 50).

The issue of choosing an interatomic potential and assessing as to whether it is a good fit for the respective problem addressed requires more than abstract theoretical knowledge, but it also requires more than knowing what material we are concretely dealing with. Rather, to develop a suitable computational model we act within pragmatic and computational constraints, and we need to make assumptions about what, in light of our problem, the relevant properties of the material are. This latter fact, that models are chosen and developed partly according to how they are assumed to reflect certain features of the object of investigation (while omitting others), is, I suggest, what

can be 'exploited' when establishing bridges between different models in a multiscale simulation context.

4.3 Bridging and the Mediatory Role of Models

This chapter has started with the observation that the success of multiscale simulations poses a puzzle: On the one hand, it seems that a naïve reductionist understanding according to which bridging between levels of description is a matter of logic or mereology is misguided. On the other hand, if the theoretical frameworks upon which the sub-models of multiscale simulations rely are, as has been pointed out by Winsberg and Bursten, *conflicting* – then how is this possible that the different levels of description can work together so well?

Examining the silicon crack simulation, Winsberg pointed out that rather than being a matter of logical or mereological reasoning, bridging between levels of description turned out to depend on physical and empirical reasoning. Going beyond this, Bursten identified two 'conceptual strategies' that figured in the silicon crack simulation and that made it possible to bridge between different levels of description, namely the manipulation of non-representational features and the tailored construction of certain fictional entities.

Taking one step back, I think it is interesting to ask more specifically what the employment of such strategies – in the context of a multiscale simulation – tells us about the role and character of the respectively involved models and theories. In the case of the MD and FE region for example, scientists could – according to Bursten – bridge between different levels of description because they manipulated non-representational features of their models. Besides examining how multiscale simulations offer nuanced insights into the concrete strategies by means of which bridges between different levels of description can be established, it is interesting to reflect more broadly as to what the employment of such strategies says about the role and character of the models and theories involved in multiscale simulations.

I suggest that the employment of certain strategies to bridge between different levels of description owes much to the mediatory role of models – models which are partly autonomous from 'their' theories, and whose communication is shaped not only by these theories, but by various modeling decisions that partly go beyond these theories, including considerations regarding what the relevant features or behaviors of the concrete system under investigation at a particular level of description are.

To elaborate on this idea, let me begin with the region of the emerging crack. Here, the tight-binding method was employed to provide a fine-grained account of the emerging tip of the crack.

Part of the motivation for using a model based on the TB method was that for the respective problem under investigation it was assumed that the breaking of molecular bonds at the emerging crack tip needs to be accounted for explicitly as it matters for the dynamics of crack propagation. As the scientists put it: "It is important to describe the energetics of this part of the system very accurately. Since it is a region where bonds are breaking it requires a quantum-mechanical description [...]" (Broughton et al. 1999, 2392). This quote also shows that choosing what the relevant features of the object of investigation for a given problem are is not necessarily something that is obtained *from* theory, but something that can *motivate* the use of a certain theory in the first place: We consider bond breaking events to be relevant in the area of the emerging crack, therefore we need an approach based on quantum mechanics.

In the neighboring region of the crack, a computational model relying on the Stillinger-Weber interatomic potential was employed. Again, the choice of this model was motivated by considerations regarding what the relevant features of silicon in the respectively investigated spatial region are. The scientists state that in the neighboring region of the crack, thermal fluctuations and pressure waves from rupturing bonds should be captured – which can be done by means of a model based on molecular dynamics (Broughton et al. 1999, 2392). Of course, the computational model based on molecular dynamics theory is also way more computationally efficient than the TB model. Still, this seems not to be the *only* reason for why the region was modeled as it was. Rather, the question of which features of the system matter in the respective region played a role: "The MD region may [...] contain those areas of the system that are defective but for which the primary dynamics are no longer important" (Broughton et al. 1999, 2392).

It is also such kinds of considerations, I suggest, that significantly shape how the handshaking strategies look like. The handshake from the crack tip region to the neighboring region was conducted with the help of 'silogens', which are – as Winsberg and Bursten put it – in some sense fictional. By introducing these hybrid atoms, Broughton, Abraham et al. could localize the energy at the TB/MD interface. The silogens allowed to preserve a bond energy equal to silicon-silicon bonds at the interface region and to obtain a longitudinal force constant which equals that of silicon (Broughton et al. 1999, 2398). At the same time, silogens differ from silicons in that they are univalent – they do not correspond to real atoms. I think an important assumption guiding the construction of this interface region was that the farer we move from the emerging crack tip, the less important is it to accurately precisely reflect all the properties of individual atoms. At the same time, since the idea of the computational model in the neighboring region of the crack is to trace pressure waves and thermal fluctuations, it is important to get the overall energetics at the interface region right.

The situation looks similar in the MD/FE region. As Abraham et al. suggest, part of the motivation for choosing a continuum model (which was then transformed into a *computational model*, discretized in the form of a FE mesh) for the region far from the emerging crack was that most of what individual atoms are doing in this region simply does not matter for the problem of studying crack propagation. Discussing the prospects of large-scale simulations of crack propagation, they state that "[in] the far-field regions, little of interest is happening: many computer cycles are spent describing trajectories of atoms that do very little apart from vibrating around lattice sites" (Broughton et al. 1999, 2392). The scientists point that the FE model was required because it allowed that the pressure pulses captured by the computational model in the neighboring region of the crack could "propagate harmlessly" through the material rather than reflect to the center at the edges of the MD simulation cell (Broughton et al. 1999, 2392). The region far from the crack is where some of the details captured by the model in the neighboring region of the crack do not matter anymore: "Here, the precise statistical mechanics is less important than allowing the free passage of (usually long wavelength) energy into or out of the system" (Broughton et al. 1999, 2392).

It seems that, again, the decision of modeling silicon as a continuous medium in the region far from the crack, and numerically approximating the behavior of this medium as a FE mesh was from the beginning on motivated by considerations concerning what, in light of the addressed problem, the relevant features of the material in the respective regions may be.

The interface regions are, perhaps, not so much regions at which two 'fundamentally mismatched' descriptions meet, but regions at which we need to be quite explicit about the factors that we deem relevant for the overall problem we are concerned with. When employing a computational model based on the FE method in the region far from the crack, this seems to be less an act of *questioning* that the material is made of atoms, and more an act of assuming that atomic details do not matter here. At the interface, as Bursten suggests, this is made explicit by adjusting the size of the mesh so as to align with the atomic lattice of the model in the neighboring region of the crack.

To sum up, all the sub-models had a mediatory role between theories and the world: their construction is informed by and partly 'constrained' by theories of matter at the respective levels of description; and at the same time they are shaped by considerations regarding what the relevant features of the material are in light of a given problem. Furthermore, the development of these models occurs within pragmatic constraints, e.g., regarding computational efficiency.

It is not, in a primary sense, theories that 'meet' at the interface regions, but models which are carefully crafted from and informed by these theories, and which are at the same time partly in-

dependent from them. That the region far from the crack tip is described as a continuous medium and computationally modeled as a discrete mesh, for example, is a somewhat contingent choice that is not straightforwardly obtained from theory, but also based on pragmatic (e.g. computational tractability and efficiency), empirical or conceptual reasoning; and, related to this, on considerations about what – in light of a given problem or question – the relevant features of the material in the respective region are. When bridging between the sub-models, assumptions play a role regarding what features should be 'preserved' across the models of the respective spatial regions of the overall investigated material sample. While these assumptions are informed and constrained by the respective theories (cf. Bursten 2018, 164), bridging between the different levels is not straightforwardly a connection of theoretical structures, but a connection of mediatory models which are developed and adjusted based on pragmatical, empirical – and, in Bursten's wording, conceptual considerations.

5 Conclusion

In this chapter, the intention was to emphasize that the *development of* and *interaction between* the models, theories and their physical targets in multiscale simulation contexts involve a multitude of modeling activities and considerations. Rather than having provided a 'recipe' for establishing connections between different levels of description, my aim was to render their apparent success in 'working together' less obscure, and to offer connecting points for further studies on how different levels of description can be combined in the context of a multiscale simulation.

As a starting point, I looked at the case of steel, as it has been discussed by Batterman, Green and Wilson. All three authors pointed out that no simple 'bottom-up' approach for describing and explaining the behavior of steel seems feasible, and that a naïve reductionist view of how different levels of description relate to each other should therefore be rejected. Instead, Batterman, Green and Wilson stressed that it is often models at intermediate levels between small and large that play a particularly significant role in capturing many scientifically-interesting phenomena.

As concerns the presented criticisms of a naïve reductionist view, the perspective offered in this chapter suggests a slightly different outlook. On the one hand, straightforward bottom-up approaches starting from simple descriptions at lower levels may typically seem misdirected. On the other hand, whether or not the behavior of certain systems can eventually be explained or understood from bottom up remains, I think, to some extent a different question.

It may, in a concrete situation, not be easy to say if we are unable to explain upper-scale behavior or structures from bottom up is because it is not possible to do so, or because we are sticking to

unwarranted simplifications about the properties of our system at small scales (Franklin, forth.). As Franklin (forth.) has pointed out, it seems not clear why a reductionist necessarily had to accept an idealized model of steel as a uniform lattice or as a homogeneous material at an intermediate level. According to the perspective suggested in this chapter, this is, to some extent, not something that is determined by theory alone, but a contingent modeling decision.

As concerns the second 'positive' suggestion by Batterman, Green and Wilson, however, the perspective offered in this chapter resonates with the idea that models at intermediate levels between micro and macro can be particularly relevant. The connection between different levels of description is a matter of working with models which take a mediatory role between theories and the world; and since models at intermediate levels often may code for particularly interesting and explanatory relevant properties of the objects of investigation (e.g. the stability of structures or organizational features), saying that 'lower-level' descriptions and models are superior to such intermediate models is not without problems.

In the second part of the chapter, looking at an example discussed by Winsberg and Bursten, I outlined that to understand how different levels of description can be connected to each other in multiscale simulation contexts, it is fruitful to consider how the models involved in a multiscale simulation act as partly autonomous mediators between theory and the world. The development of models such as those in the silicon crack simulation involves, as Winsberg and Bursten have outlined, pragmatic, empirical and conceptual considerations – and so does bridging between them. Since they are partly autonomous from their underlying theories, they can to some extent be manipulated and adjusted so as to make sure that certain features which are deemed relevant for the problem under investigation are 'preserved' across different modeling regions while others are omitted or artificially changed.

In this sense, rather than offering a miraculous reconciliation of otherwise incompatible theoretical frameworks, the bridging between different levels of description is a matter of careful adjustment and working with models which mediate between theories and the world.

Bibliography

Abraham, F. F., Broughton, J. Q., Bernstein, N., and Kaxiras, E. (1998). Spanning the continuum to quantum length scales in a dynamic simulation of brittle fracture. *EPL (Europhysics Letters)*, 44(6):783.

Batterman, R. W. (2021). *A Middle Way: A Non-Fundamental Approach to Many-Body Physics*.

Oxford University Press.

Batterman, R. W. and Green, S. (2021). Steel and bone: mesoscale modeling and middle-out strategies in physics and biology. *Synthese*, 199(1-2):1159–1184.

Boge, F. J. (2019). Why computer simulations are not inferences, and in what sense they are experiments. *European Journal for Philosophy of Science*, 9(1):13.

Broughton, J. Q., Abraham, F. F., Bernstein, N., and Kaxiras, E. (1999). Concurrent coupling of length scales: methodology and application. *Physical Review B*, 60(4):2391.

Bursten, J. R. (2018). Conceptual strategies and inter-theory relations: The case of nanoscale cracks. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 62:158–165.

Cole, I. S. and Marney, D. (2012). The science of pipe corrosion: A review of the literature on the corrosion of ferrous metals in soils. *Corrosion Science*, 56:5–16.

Eckert, C. and Hillerbrand, R. (2022). Models in engineering design as decision-making aids. *Engineering Studies*, 14(2):134–157.

Franklin, A. (fort). How the reductionist should respond to the multiscale argument, and what this tells us about levels. In Robertson, K. and Wilson, A., editors, *Levels of Explanation*. Oxford University Press.

Green, S. and Batterman, R. W. (2017). Biology meets physics: Reductionism and multi-scale modeling of morphogenesis. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 61:20–34.

Lenhard, J. (2014). Autonomy and automation: computational modeling, reduction, and explanation in quantum chemistry. *The Monist*, 97(3):339–358.

Morgan, M. S. and Morrison, M., editors (1999). *Models as Mediators*. Cambridge University Press.

Morrison, M. (2007). Where have all the theories gone? *Philosophy of Science*, 74(2):195–228.

Morrison, M. (2015). *Reconstructing Reality: Models, Mathematics, and Simulations*. Oxford Studies in Philosophy of Science.

Morrison, M. and Morgan, M. S. (1999). Models as mediating instruments. In Morrison, M. and Morgan, M. S., editors, *Models as Mediators*, pages 10–37. Cambridge University Press.

Müser, M. H., Sukhomlinov, S. V., and Pastewka, L. (2023). Interatomic potentials: Achievements and challenges. *Advances in Physics: X*, 8(1):2093129.

Raabe, D., Scheffler, M., Kremer, K., Thiel, W., Neugebauer, J., and Jansen, M. (2009). Multi-scale modeling in materials science and engineering.

Wilson, M. (2017). *Physics Avoidance: and Other Essays in Conceptual Strategy*. Oxford University Press.

Winsberg, E. (2006a). Handshaking your way to the top: Simulation at the nanoscale. *Philosophy of Science*, 73(5):582–594.

Winsberg, E. (2006b). Models of success versus the success of models: Reliability without truth. *Synthese*, 152:1–19.

Winsberg, E. (2010). *Science in the Age of Computer Simulation*. University of Chicago Press.

List of Figures

List of Tables