Paul Egan[1]
Department of Mechanical and
Process Engineering,
Swiss Federal Institute of Technology
(ETH Zurich),
CLA F 34.1,
Tannenstrasse 3,
Zurich 8092, Switzerland
e-mail: pegan@ethz.ch

Jonathan Cagan
Department of Mechanical Engineering,
Carnegie Mellon University,
5000 Forbes Avenue,
Pittsburgh, PA 15213
e-mail: cagan@cmu.edu

Christian Schunn
Department of Psychology,
University of Pittsburgh,
4200 Fifth Avenue,
Pittsburgh, PA 15260
e-mail: schunn@pitt.edu

Felix Chiu
Department of Mechanical Engineering,
Carnegie Mellon University,
5000 Forbes Avenue,
Pittsburgh, PA 15213
e-mail: Felixchiu92@gmail.com

Jeffrey Moore
Department of Biological Sciences,
University of Massachusetts Lowell,
One University Avenue,
Lowell, MA 01854
e-mail: Jeffrey_Moore@uml.edu

Philip LeDuc
Department of Mechanical Engineering,
Carnegie Mellon University,
5000 Forbes Avenue,
Pittsburgh, PA 15213
e-mail: prl@andrew.cmu.edu

The D_3 Methodology: Bridging Science and Design for Bio-Based Product Development

New opportunities in design surface with scientific advances: however, the rapid pace of scientific discoveries combined with the complexity of technical barriers often impedes new product development. Bio-based technologies, for instance, typically require decisions across complex multiscale system organizations that are difficult for humans to understand and formalize computationally. This paper addresses such challenges in science and design by weaving phases of empirical discovery, analytical description, and technological development in an integrative "D_3 Methodology." The phases are bridged with human-guided computational processes suitable for human-in-the-loop design approaches. Optimization of biolibraries, which are sets of standardized biological parts for adaptation into new products, is used as a characteristic design problem for demonstrating the methodology. Results from this test case suggest that biolibraries with synthetic biological components can promote the development of high-performance bio-based products. These new products motivate further scientific studies to characterize designed synthetic biological components, thus illustrating reciprocity among science and design. Successes in implementing each phase suggest the D_3 Methodology is a feasible route for bio-based research and development and for driving the scientific inquiries of today toward the novel technologies of tomorrow. [DOI: 10.1115/1.4033751]

1 Introduction

Scientific discoveries are often gateways for new technologies, and in recent years, engineers have begun turning to biology for design inspiration [1–6]. The usefulness of biological systems to engineers also extends past inspiration, and new advances in science [7] are revealing a promising future for redesigning and utilizing biological components directly [8]. The process of utilizing biological components in a new product, referred to as a bio-based design [9,10], has the potential to drive the development of new technologies for diverse applications [11–13].

In mechanical design domains, there is a great opportunity for leveraging biological components and design principles to develop technologies including nano-actuators, smart contractile materials, and shear-triggered medicines [14]. Recently, we have demonstrated the feasibility in utilizing computational approaches for designing myosin motor protein biosystems informed by scientific experiments [15,16] and have conducted empirical studies to understand and improve human bio-based design decision making [17,18]. In this work, findings and processes from these past studies are utilized to propose and implement a new general methodology that bridges science and design for bio-based product development.

In developing bio-based products, there is a need for new methodologies due to the inherent complexity of biological systems and the physical principles that differentiate biosystems from traditional mechanical systems. Large numbers of myosin motor proteins, for instance, use chemical energy to attach to protein filaments and exert force before detaching. Myosins are traditionally known for powering muscle contractions; here, the difference in scale from the myosin molecule to the muscle itself is very large. A muscle contraction occurs when myosins exert force on filaments that slide in relation to one another, which creates a cascading translational motion across the multiscale organization of muscle fibers. A single myosin operates analogously in many respects to a traditional mechanical motor [15] but differs greatly when considering each myosin has stochastic mechanical

[1]Corresponding author.

Contributed by the Design Theory and Methodology Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received October 12, 2015; final manuscript received May 31, 2016; published online June 16, 2016. Assoc. Editor: Andy Dong.

behaviors and must operate in large groups to smoothly power muscle contractions.

Although it is now possible to engineer new myosins with tailored properties [19,20], there are few formal approaches for guiding the design of bio-based systems [21], partly due to the challenges in both human and computational approaches in the domain. Human-only approaches are limited, since the vast number of parameter relationships and emergent behaviors of biological systems are difficult to understand [22]. Computational-only approaches are limited due to the challenges in creating models of emergent behavior [23,24]. Due to the limitations in both human and computational approaches, we propose a new approach that involves humans guiding computational processes.

Such human-in-the-loop design approaches are well suited for engineering complex systems, since they enable human designers to operate at a high-level and steer routine computational tasks [25,26]. Our approach for designing bio-based products, termed the D_3 Methodology, aims to bridge science and design perspectives through phases of empirical Discovery, analytical Description, and technological Development (Fig. 1).

At a high level, the Discover phase uncovers new information about a system, the Describe phase aims to model and predict the behavior of a system, and the Develop phase focuses on configuring systems to achieve high-performance designs. Specific goals of each D_3 phase (see Fig. 1) involve (1) performing experiments and analyzing data in the Discover phase, (2) proposing models and validating them in the Describe phase, and (3) conceptualizing and optimizing technologies in the Develop phase. This coordinated set of three phases is particularly suitable for bio-based design, since current scientific efforts in synthetic biology regularly consist of steps for understanding, design, and analysis [21], which map roughly to the Discover, Develop, and Describe phases, respectively.

The D_3 methodology differs from existing scientific approaches since it focuses on design outcomes, rather than primarily focusing on enhancing understanding and knowledge, and is best used with biological systems that are already known to exist and are potentially engineerable. Its steps are ordered such that scientific discovery and description may be driven by a specific end-design goal. Scientific discoveries facilitated by the methodology are aimed toward gaining a stronger understanding of the system for engineering purposes rather than for just generally promoting scientific discoveries [27].

The differing order of similar phases among science (understand, design, analyze) and our D_3 design approach (discover, describe, design) suggests that the starting point for implementing the methodology is flexible. That is, there is potential for designers to begin with any phase depending on their technical expertise. A designer is expected to initiate the methodology by inputting a base set of assumptions relevant to the starting phase of interest.

One potential path through the phases is following Discover, Describe, and Develop phases sequentially, which is illustrated in the Fig. 1 example for designing new myosin technologies, and follows a traditional progression of conducting new scientific experiments to inform engineering decisions. The Discover phase could consist of a defined experiment with variables to measure as an input, which is illustrated in Fig. 1 as myosins propelling filaments on a microscope slide. Empirical measurements from the Discover phase are used in the Describe phase for modeling experimental phenomenon and are illustrated in Fig. 1 with a parameterized myosin model. The myosin model is used during the Develop phase for evaluating designs that are optimized, as illustrated by the myosin nano-actuator in Fig. 1.

There are also input/output combinations that do not follow the sequential ordering of phases. If the run time of a model is too long to be useful for evaluation in the Develop phase, a constraint on evaluation time could be set in the Describe phase. Additionally, new designs found in the Develop phase may be experimentally tested as a means of model validation, since design exploration will interpolate and extrapolate myosins beyond available data. The overall methodology, therefore, promotes a reciprocal science and design relationship [28,29] manifested bidirectionally through (1) converting empirical discoveries into developed technologies and (2) extrapolating beyond known scientific findings to explore potentially useful designs that are then tested. Such reciprocity has great potential for streamlining the information flow across science and design disciplines, and thereby promoting the development of new products.

The aim of this paper is to examine the D_3 Methodology's feasibility, which is accomplished by implementing a sequential iteration of the D_3 Methodology using myosin technologies as an example case. A successful implementation of the D_3 Methodology should demonstrate how findings from each phase influence one another and demonstrate the synergies of concurrent scientific and design processes.

The paper is organized to guide a reader through one implementation of the D_3 Methodology, beginning with background relevant to implementing the D_3 Methodology for a specific example of developing biolibraries with standardized parts [30–32]. These biolibraries share similarities with traditional engineering product families [33]. Sections 4–6, respectively, demonstrate the implementation of Discover, Describe, and Develop phases for the specific myosin biolibrary design problem. Biolibrary configuration is a challenging engineering optimization problem that requires searching a large design space to find a set of myosin molecules that support high performance across a range of nanotechnologies. Section 7 discusses the feasibility and further applicability of the D_3 Methodology and relates the specific examples in designing

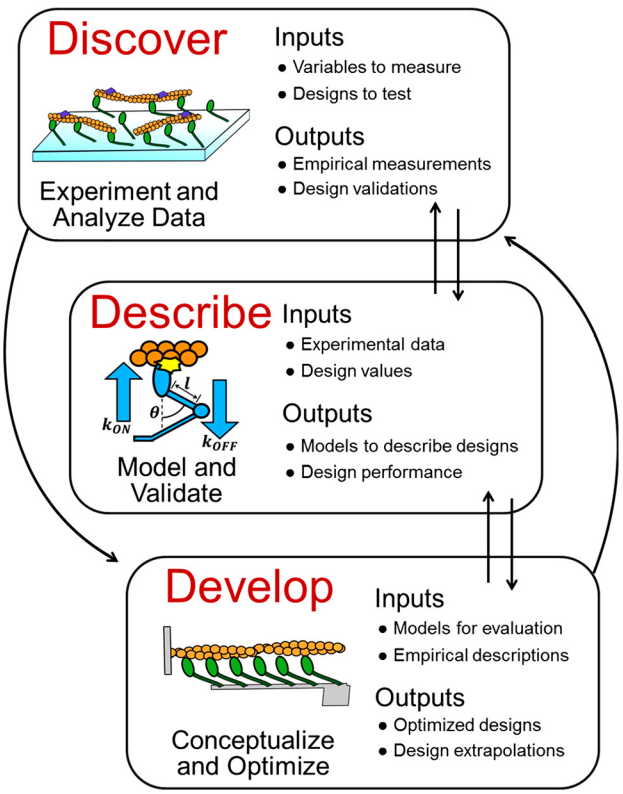

Fig. 1 The D_3 Methodology consisting of a Discover phase with generic steps for experiments and data analysis, a Describe phase with generic steps of modeling and validation, and a Develop phase with generic steps for conceptualization and optimization of new technologies; example inputs/outputs for each phase are presented in the context of myosin motor protein research and development

myosin nanotechnologies toward conjectures in how the D_3 Methodology may frame future design pursuits.

2 Background

Background relevant to implementing and assessing the D_3 Methodology is briefly reviewed in three parts covering (1) intersections between science and design, (2) the role of cognition for informing a human-in-the-loop approach, and (3) computational endeavors in bio-based design.

2.1 Science and Design Intersections. A core process in engineering is the application of scientific knowledge toward creating new designs, although there is much debate concerning which reasoning processes belong to science versus design [34–37]. Traditional arguments often characterize science as the study of things that exist naturally, while design is characterized as the creation of things that do not exist naturally. These notions tend to blur when considering biologically based design disciplines where natural systems are studied and utilized to engineer new systems.

Although not the focus of this work, bio-inspired approaches in engineering have had great success for promoting new technologies, which includes directly mimicking biological systems with biomimicry [8,38] and transference of functional principles to new designs with functional-based approaches [1–5]. Biomimicry approaches tend to map biological structures or behaviors directly to a new design, such as synthetic organisms that mimic jellyfish swimming [8] and materials that have repeating structures similar to organic materials [38]. Functional-based approaches tend to map design principles or functionality of a biological system to a new application, which could have an analogous function but very different structure and behavior [2]. Engineering design approaches have been developed for recognizing biological analogies and functionalities, such as retrieving functions from natural-language text [1]. There is also the development of tools for managing the large amount of biological knowledge available to engineers [5]. The growing use of and research on bio-inspired approaches suggest that engineers have great interest in learning about biology, which could lower the barrier for designers to utilize learned biological knowledge toward the development of technologies with biological components [39–42].

Synthetic biology, which is the focus of this paper, is different from bio-inspired design. Synthetic biology is an emerging field focused on the engineering of biological systems with a goal of either gaining greater scientific insights or developing new technologies that are biologically based [21]. There are approaches for both modifying existing biological systems and engineering new biological systems from existing molecular components. Although natural biological systems are often highly optimized for specific functions through evolution [43], there is much room for engineering improved systems for new applications that do not exist in nature. Design decisions in biological domains are inherently limited by the available building blocks that make up biological components, such as proteins. Despite these constraints on the building blocks, designing even a single protein is a complex engineering task [44], since stochastic properties of biological systems make many precise calculations and predictions challenging.

2.2 Cognitive Processes for Complex Systems Design. The inherent complexity of biological systems makes them difficult to formalize computationally, so there is a need to consider involving human designers when developing technologies. It is possible to leverage the strengths of both humans and computers in designing systems through a human-in-the-loop approach [25,26]. Human-in-the-loop approaches are well suited for complex systems design, since they enable human judgment to eliminate unrealistic designs rapidly or to provide creative input [25]. Experiments have demonstrated that utilizing human designers in the loop can result in finding higher quality designs with fewer function evaluations [26]. In order to effectively utilize human designers in the loop, it is important to consider not only their strengths but also their limitations, such as a human designer's limited capability to consider many parameters simultaneously.

A foundational psychology study long ago demonstrated that the amount of information humans may consider is severely limited to just a handful of independent elements at any one point in time [45]. Such limitations inhibit a human designer's ability to search a parameter space for optimal solutions [46]; human performance has been measured to geometrically decline as the number of considered parameters increases. The decline in human design performance was hypothesized to occur due to the limited amount of information a human may consider in their working memory. These considerations informed our recent experiment to measure designer search performance with four design inputs and up to two design outputs involving myosin technologies [17,47]. The study demonstrated that human designers' performance declined as the number of output constraints increased on a problem statement.

A human designer's search performance may be improved through appropriate system representation, such as using structure–behavior–function paradigms to facilitate understanding of system interactions [48,49]. Structure–behavior–function representations are particularly well suited for reasoning about complex systems. In the structure–behavior–function paradigm, structures refer to components of a system; behaviors refer to mechanisms and interactions of components; while functions refer to goals of designs or components. Our past work with myosins led to the development of a myosin agent-based simulation within a structure–behavior–function paradigm. In the simulation, each myosin's structure was designable and linked to behaviors that influence design performance [15]. Human designers interacting with the agent-based simulation gained understanding of the system that improved their ability to search a parametric design space [18]. These studies offer validation for using a parametric design representation for the human-in-the-loop approach. Determining the best representation of a system is difficult [50], especially since optimal representations will likely differ for novices and experts in a domain [51]. In particular, novices will likely perform better with representations that simplify systems to a few core variables, while experts may benefit from a more complicated representation that enables a higher degree of fine tuning design configurations.

2.3 Bio-Based Computational Design. The use of agent-based simulations is beneficial from a biological modeling perspective, since the development of analytical models is not always sufficient for concisely and accurately modeling systems [52,53]. For instance, a few bottom-up rules employed by agents that represent different proteins can replace large series of differential equations for describing molecular interactions. It is possible to validate agent-based models through using reverse engineering methods for fitting models to natural and complex systems [54–56]. Generally, these approaches propose a feasible model, simulate it, and then compare the simulated results to empirically measured results. If there is not a close match, a new model may be proposed and the process is repeated.

Once computational biological models are validated, they may be utilized to evaluate new system configurations that are representative of designs. In our past work, agent-based myosin models were validated with empirical data [16] and provided a means for evaluation in engineering optimization tasks representative of myosin technologies [18]. Since these tasks represent a variety of technologies, an interesting design problem emerges when considering the best set of myosins for use in all technologies, such as selecting myosins from a design catalog.

There is a rising interest in developing tools and standardized biological parts that could form the foundation of a myosin design catalog [30–32]. Although there is relatively little prior work in

the computational optimization of myosins for design, considerable effort has been put into studying, modeling, and engineering myosins to learn about their structures, behaviors, and functions [19]. These studies can form the foundation for developing a catalog, referred to in this paper as a biolibrary of myosins, for use in developing myosin-based technologies [39–42].

A biolibrary of standard useful myosins may enable a designer to focus on selecting myosins for use in new systems, rather than creating new myosins for every design application, which is similar to selecting standard nuts and bolts to configure products or reusing components to form a product family [33]. Biolibraries may consist of myosins already existing in nature [57] or novel synthetic myosin designs that may have superior behaviors and performance [20,58]. The optimization of a range of devices selected from biolibrary components could leverage existing design methods, such as software agents selecting design components from a catalog [59]. Agent-based optimization is well suited for searching complex design spaces and may be conducted in a domain-general manner through representing designs as binary strings [60] that could represent both individual myosins in a biolibrary and myosin-based technologies.

3 D_3 Methodology for Myosin Biolibrary Design

The D_3 Methodology aims to promote product design by bridging science and design processes across three iterative and flexibly ordered phases of empirical discovery, analytical description, and technological development. Since the overall methodology emphasizes engineering endeavors, the Discover phase seeks to uncover new understandings of bio-based components that are potentially well suited for use in designed products. The development of myosin biolibraries is used as a test case for sequentially implementing the three D_3 phases. Myosins are well suited for design endeavors since they are engineerable molecules that provide a basis for a number of nanotechnologies.

We define a biolibrary in the context of myosins as a set of varied myosin designs that are flexibly deployed across particular applications. Myosins often operate in groups to generate power, rather than in isolation; groups of myosins are referred to as biosystem blocks, which are small units that may be interfaced in various organizations to form myosin technologies such as synthetic muscle [39,40], contractile materials [41], and bio-based nano-actuators [42]. Therefore, biolibraries should be assessed based on their capability for promoting high performance across a set of myosin groups (Fig. 2).

All of the nanotechnologies shown in Fig. 2 operate through the fundamental function of myosins stochastically exerting force on filaments. For synthetic muscles (top right of Fig. 2) and nano-actuators (bottom right of Fig. 2), the net result of myosin force creates translational filament movement. Contractile materials (middle right of Fig. 2) generally contract to a smaller size as myosins generate force. The eight myosin isoforms illustrated on the left in Fig. 2 have varied geometrical configurations that are representative of different isoforms; up and down arrows illustrate differences in each myosin's chemical rates of attaching and detaching to a filament, respectively.

There are a number of existing myosin isoforms with different behavioral properties for use in nanotechnologies [57]; isoforms are families of molecules with slight structural differences that lead to different performances for a given function. New myosins are still being discovered [61,62], and there are potentially many undiscovered myosins since many animal species have unique isoforms not found in other species. Further, it is possible to design and fabricate synthetic myosins with altered behaviors that may provide superior performance over natural isoforms [20,58]. One of the most common approaches for manufacturing new myosins is the transference of parts from one myosin isoform to a second myosin isoform, thus creating a new myosin with hybrid properties. It is also possible to create protein parts for insertion in myosins, such as lever arms of varied lengths [19] that influence both how long a myosin remains attached to a filament and a myosin's output force.

The design of myosin biolibraries is relevant to cost and performance trade-offs for potential bio-industry applications, such as a company or laboratory's decision to utilize a biolibrary of natural isoforms or synthetic myosins custom designed for each application. Depending on available company resources, there may be different costs associated with using synthetic or natural myosins. Therefore, there is a need for early feedback on the likely performance of manufacturable myosins, since manufacturing of biological technologies is often costly and difficult [38]. The D_3 Methodology aims to aid in these decision making endeavors by presenting a framework for guiding a human scientist or designer toward isolating cost-effective solutions sooner.

Sections 4, 5, and 6 of this paper are focused on a specific implementation of the D_3 Methodology for the development of optimized myosin biolibraries, with each section concentrating on one phase of the methodology tied to understanding, modeling, and designing myosin systems. Specifically, natural myosin isoforms are characterized in Secs. 4 and 5 in the Discover and

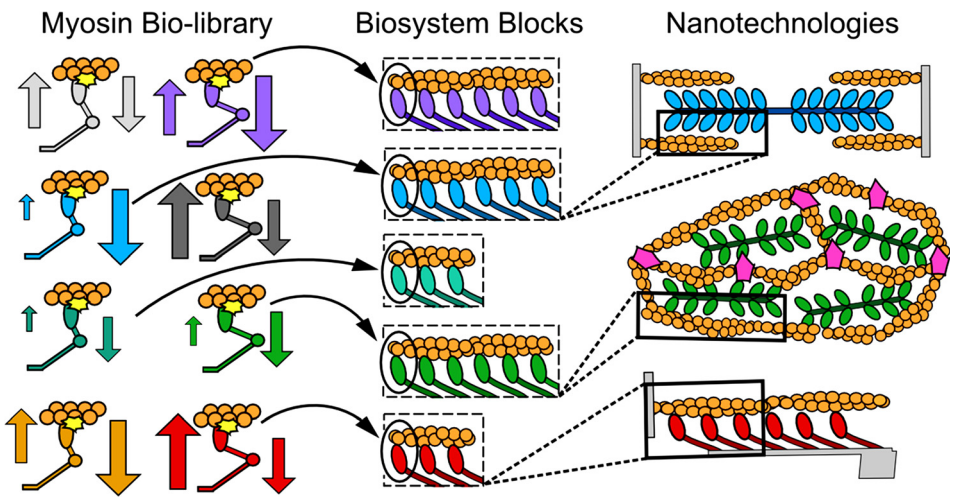

Fig. 2 Schematic of a biolibrary with varied myosin designs, biosystem blocks consisting of myosins from the biolibrary, and nanotechnologies constructed from biosystem blocks. Illustrated nanotechnologies from top to bottom include a synthetic muscle, contractile material, and nano-actuator.

Describe D_3 phases, respectively. Natural myosins are then contrasted with synthetic isoforms extrapolated from the model in Sec. 6, which focuses on the Develop D_3 phase.

4 Discover Phase: Experiment and Analyze

The Discover phase requires a human to carry out scientific experiments, which are then analyzed quickly through computational approaches. Experimental measurements are the inputs to the Describe phase, which validates models with experimental findings and enables design evaluation in the Develop phase. The human experimenter in this phase is important, because they can make decisions in how data are collected and what data are collected. These decisions influence which automated approaches may be used in later D_3 phases to model and optimize the system.

In the context of biolibrary development, empirically characterized natural myosins could form the foundation of a biolibrary of myosins with different performance characteristics for use in designed products. However, before they are usable in the Design phase, their behaviors must be measured and modeled. Myosin mechanical behaviors are often inferred by recording microscopy videos for measuring the velocity of filaments propelled by myosins anchored to a microscope slide [19]. Automated filament tracking methods may then be utilized to quickly and reliably measure filament velocities from a given experiment [63,64] and potentially provide a large body of data for validating myosin models.

In our work, chicken skeletal muscle myosins were utilized to validate an automated tracking and aggregation approach that tracked filaments based on their centroid location across images extracted from microscopy videos. Images were preprocessed to aid consistent finding of centroid locations across frames; preprocessing steps will differ based on video quality from different laboratory setups but are essential to reduce noise across frames and provide consistent filament shapes.

Automated tracking results were validated against manually tracked [65] measurements collected by two independent users that tracked filament locations across fames. Figure 3 presents manually and automatically tracked results of filaments propelled by chicken skeletal muscle myosins for 3.8 s.

Results show a close resemblance in track shapes obtained with both methods, which provides an initial qualitative validation. Data were then aggregated by averaging the velocities of each individual track within a video across multiple repeated experiments/videos. Data aggregation for the manually tracked filaments provides a calculated mean velocity and standard error of $5.16\ \mu m/s \pm 0.31\ \mu m/s$ that closely match the automatically tracked filaments' mean velocity and standard error of $4.93\ \mu m/s \pm 0.32\ \mu m/s$. These results demonstrate that the automated method is capable of analyzing empirical data quickly and accurately. The use of the automated tracker enables the gathering of large sets of data for validating models in the Describe phase that are used for evaluating designs in the Develop phase. Because the models are used for interpolation and extrapolation in the Develop phase, it is particularly important to validate the models in the Describe phase against a range of empirical data, which, in turn, depend upon efficient data analysis techniques in the Discover phase.

The data specifically provide the velocity of a filament traveling relative to a given myosin, which enables myosin behaviors to be inferred when different experimental conditions are introduced and by considering models that link structural features of myosins to their expected performance in propelling filaments. In this paper, we consider experimental conditions of using two different types of myosins and also an additional molecule type [66] that exerts a force on the filament that myosins must work against. These conditions enable the reverse engineering of myosin design parameters through comparing model predictions with measured empirical data.

5 Describe Phase: Model and Validate

Once data are obtained from the Discover phase, modeling is required to describe biophysical phenomena for evaluating configured designs. The Describe phase consists of proposing a model with design inputs, and then manipulating those design inputs until they provide a performance description consistent with measured empirical data. Validated models are crucial for design exploration in the Develop phase, since they enable analytic rather than trial-and-error exploration of the complex product space. The use of validated models enables the evaluation of myosin designs within a biolibrary to determine their usefulness for promoting desirable product performance.

Although other models for evaluating biological systems exist that could be used in the Describe phase [67], agent-based simulations are used in our implementation since they promote human understanding of complex systems [18,22] that is essential, since humans are expected to guide computational processes in the overall D_3 approach. Additionally, agent-based simulations are modular, which means that many mechanically based molecules are representable with few modeling modifications, such as using a single agent-based simulation to model both myosins and alpha-actinin molecules; alpha-actinins are proteins that interact with the same filaments as myosin. Alpha-actinins are commonly used

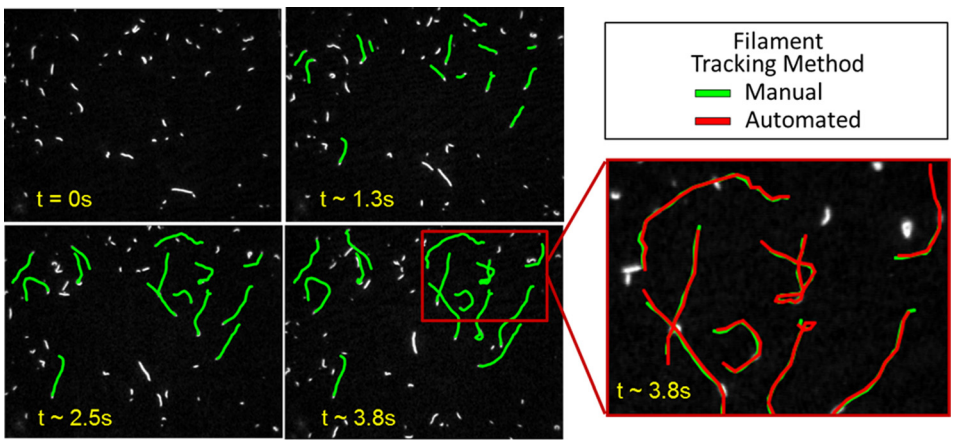

Fig. 3 Automated and manual tracking of filaments propelled by myosins over time. The movement of select filaments is indicated by lines that trace each filament's location across frames. In the $t = 0$ s frame, all displayed objects are imaged filaments, since there are no tracking lines.

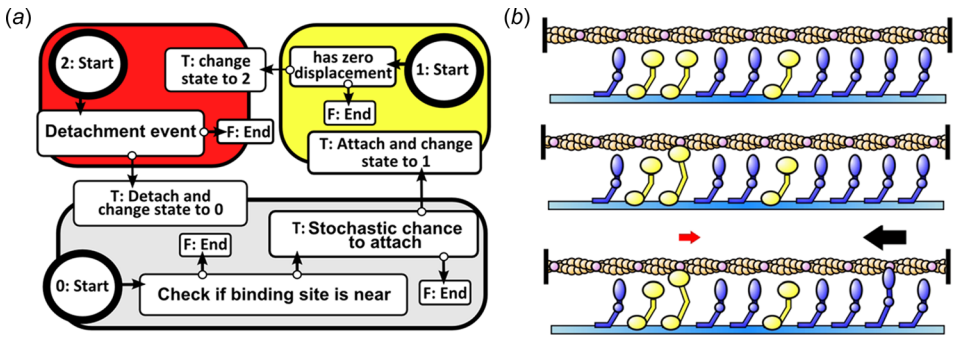

Fig. 4 Agent-based model with (*a*) rules for simulated molecules and (*b*) renderings of myosin (one globular end) and alpha-actinin (two globular ends) molecules. Left-facing arrows are positive forces that propel the filament; right-facing arrows are negative forces.

in myosin experiments for simulating a force load [66] and are generally assumed to attach to filaments and provide loading as passive springs that impede filament motion.

Our past agent-based implementation for simulating myosins is modified to a more generalized representation (Fig. 4(*a*)) that enables the simulation of alpha-actinin (Fig. 4(*b*)). All simulation details provided in this paper are identical to past implementations in Refs. [15,16], and [68], with only relevant details to the present discussion provided.

The simulation environment emulates experiments from the Discover phase, with a discrete number of myosins and alpha-actinins interacting with a filament moving at constant velocity. Each molecular agent operates autonomously according to a three-state cycle (Fig. 4(*a*)) of being detached (state 0), attached with positive displacement (state 1), and attached with negative displacement (state 2). Since alpha-actinins only generate force that impedes filament motility, they are assumed to attach with a displacement of zero and skip the positive force generation phase, which is reflected in Fig. 4(*b*) simulation rendering.

Parameters describing alpha-actinin configurations are determined by tracing alpha-actinin and myosin biophysical behavior through the logic of Fig. 4(*a*). Alpha-actinins have a stochastic chance to attach to the filament with attachment rate k_{on} and generate force f_{alp} based on their stiffness k and displacement d from their initial position, such that $f_{alp} = k \cdot d$. Alpha-actinins have a much slower detachment rate than myosins (possibly as low as $10\,s^{-1}$ [66]) and are assumed to detach at displacement d_r that causes bond rupturing [69], which is modeled deterministically.

The simulation is run with measured filament velocities from the Discover phase to empirically validate proposed parameter values for myosin and alpha-actinin. Known parameter values describing chicken skeletal myosin [19] assume myosins have a lever arm length of $l = 10\,nm$, an attachment rate of $k_{on} = 900\,s^{-1}$, and a detachment rate of $k_{off} = 1600\,s^{-1}$; these parameters are indicative of how long a myosin generates positive force, its chance of attaching to a filament, and its chance of detaching from a filament, respectively. Through algorithmic methods of proposing an alpha-actinin configuration and determining the steady-state velocity of the system, an alpha-actinin configuration of $k = 2.5pN/nm$, $k_{on} = 1500\,s^{-1}$, and $d_r = 30\,nm$ is found to fit strongly with empirical measurements (Fig. 5(*a*)).

Known alpha–alpha actinin parameter values are used to determine the parameter values of an unknown pig cardiac muscle myosin with experimental data (Fig. 5(*b*)). Pig cardiac myosins are assumed to have a $l = 10\,nm$ lever arm length that is typical of muscle myosins. A strong agreement of model to data in Fig. 5(*b*) data occurs when $k_{off} = 250\,s^{-1}$ and $k_{on} = 125\,s^{-1}$. The resulting parameter values for each myosin isoform are presented in Table 1.

Table 1 findings are consistent with pig cardiac myosin having a lower energy usage rate and velocity than chicken skeletal muscle myosin [70]. The data suggest that the myosins have contrasting attachment and detachment parameter values that may influence the performance of configured nanotechnologies from myosin biolibraries and importantly show that the reverse engineered myosin parameter values match myosin system behavior in Fig. 5. The results show the feasibility of reverse engineering myosin models through interpreting data collected in the Discover phase, which enables evaluation of myosin designs in the Develop phase.

6 Develop Phase: Conceptualize and Optimize

Models from the Develop phase form the basis for generating and evaluating designs that can be compared and considered for use in products. For the myosin example study, the empirically validated model from the Describe phase is utilized in the Develop phase to configure and evaluate myosins and biosystem blocks for the biolibrary design problem described in Sec. 3. Biolibraries are catalogs of varied myosin isoforms,

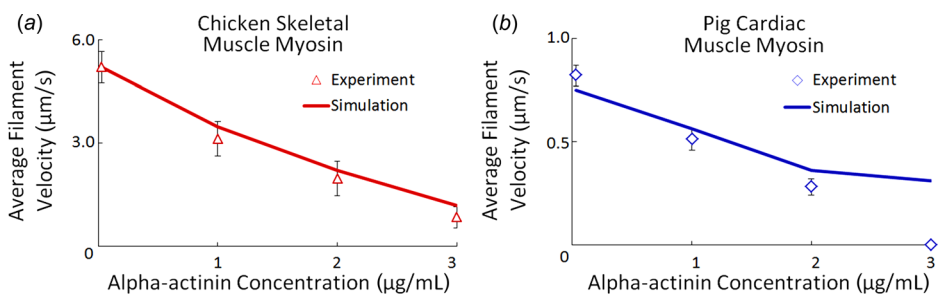

Fig. 5 Measured and simulated data of average filament velocity for (*a*) chicken skeletal muscle myosin and (*b*) pig cardiac muscle myosin experiments when varied concentrations of alpha-actinin molecules are introduced

Table 1 Reverse engineered myosin parameters

Natural myosin	Lever length (nm)	Attachment rate (s^{-1})	Detachment rate (s^{-1})
Chicken skeletal	10	900	1600
Pig cardiac	10	125	250

and a desirable biolibrary should contain a set of myosins that enables the configuration of a variety of high-performance nanotechnologies.

6.1 Biolibrary Design Representation. Myosin biolibraries potentially consist of many isoforms with multiple design variables. Concise and modular design representations are essential to promote human understanding of the multilevel design space. The lowest level of the design space refers to the configuration of a single myosin and the highest level refers to how myosins are organized to function as systems.

A biolibrary is represented as a discrete set of myosins, with each myosin design having its own set of design parameters. Myosin technologies are not modeled explicitly, but rather represented by biosystem blocks that are the smallest unit of multiple myosins with performance attributes characteristic of complete products. It is necessary to evaluate myosin performance when configured as biosystem blocks since myosins must operate in groups to function; a myosin operating by itself will not propel a filament because the filament is likely to travel out of the reach of a single myosin due to diffusion when the myosin is detached. Biosystem blocks are used for evaluation rather than specific nanotechnology applications to provide an application-independent approach for assessing a biolibrary's usefulness.

Myosin biolibraries are evaluated based on how well they fulfill varied biosystem block functionalities, rather than rating individual molecules within the library. Biosystem blocks are evaluated according to how well they fulfill performance requirements representative of potential nanotechnology functional requirements [17,18]; specifically, these evaluations rate how well a biosystem block achieves a particular performance goal without violating any design constraints. Each biosystem block has two design parameters: one for selecting a myosin from the biolibrary and one for determining the number of that particular myosin used to construct the biosystem block.

Binary string representations [60] are used that carry all information required to configure a biolibrary and place its myosins in a set of biosystem blocks for evaluation. The binary strings have two parts for (1) representing configured myosins in a biolibrary and (2) representing biosystem blocks that are configured from myosins in the biolibrary (Fig. 6).

Table 2 Genes, units, value ranges, and number of bits for myosins

Myosin gene	Symbol	Units	Minimum	Maximum	Bits
Myosin lever	l	nm	6	14	2
Myosin angle	θ	deg	30	45	1
Myosin attach rate	k_{on}	s^{-1}	100	1800	3
Myosin detach rate	k_{off}	s^{-1}	200	3200	3

Table 3 Genes, units, value ranges, and number of bits for biosystem blocks

Biosystem block gene	Symbol	Units	Minimum	Maximum	Bits
Selected isoform	myo	—	1	8	3
Number of myosins	N	#	30	150	3

In Fig. 6, each binary string represents a biolibrary with two myosin isoforms ($Myosin_1$ and $Myosin_2$) that each have two bits representative of a myosin's attachment (e.g., two bits for k_{on1} for $Myosin_1$ and two bits for k_{on2} for $Myosin_2$) and two bits representative of a myosin's detachment rate. There are three biosystem blocks that each have one bit for determining which isoform is placed in the biosystem (e.g., if the Myo_1 bit is turned off for $Block_1$, $Myosin_1$ is placed in $Block_1$; if the Myo_1 bit is turned on for $Block_1$, then $Myosin_2$ is placed in $Block_1$). Each biosystem block also has one bit for determining how many myosins are present. Each binary string contains 14 bits of information, resulting in 2^{14} ways to configure the biolibrary and place its myosins in biosystem blocks. In sum, each binary string contains information that describes the configuration of all myosin designs within a library, and how they are placed in each biosystem block for evaluation in relation to potential myosin technology performance metrics. Additionally, it is possible to alter the design space by increasing or reducing the number of genes that describe myosins or biosystem blocks, and the number of bits that determine the resolution of values for each gene.

In Tables 2 and 3, genes for each myosin and biosystem block are presented for the most complex optimization case considered in this paper. Myosin parameters include myosin lever arm length l, step angle θ, attachment rate k_{on}, and detachment rate k_{off}

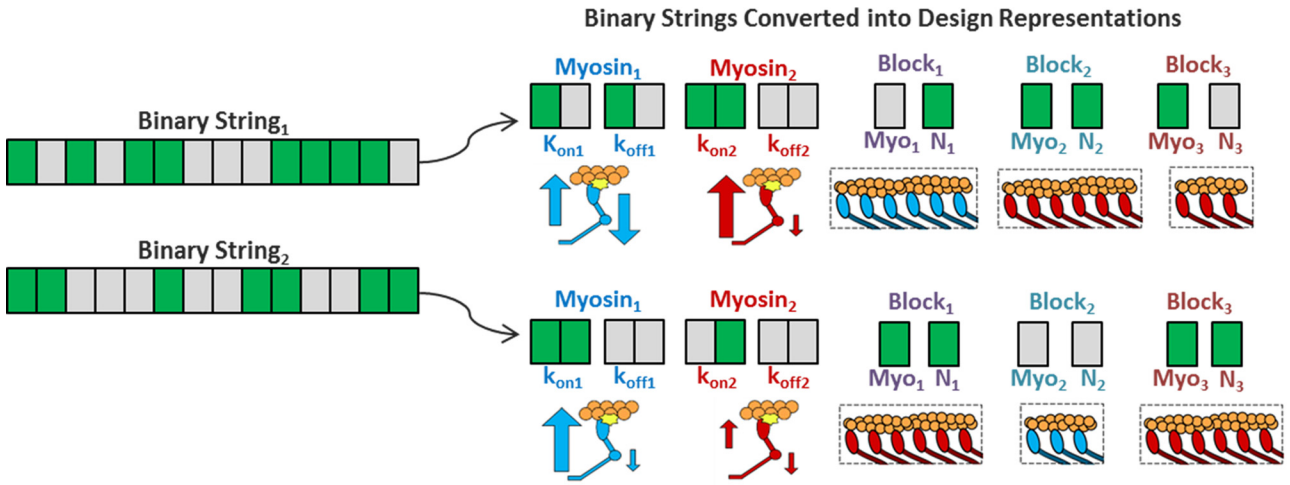

Fig. 6 Binary strings of biolibraries and biosystem blocks. Darkly shaded boxes represent turned on bits in genes mapped to design inputs.

(Table 2). Biosystem block parameters include selected myosin myo and the number of myosins present N (Table 3), which determine a biosystem block's size. Minimum and maximum parameter values were informed by findings in the Describe phase and known values from natural myosin and myosin engineering experiments [20,57].

Each Table 2 and 3 parameter has a minimum value when all bits are turned off and a maximum value when all bits are turned on. Intermediate values are calculated based on a gene's maximum number of bits, and which bits are turned on and off. It is possible to alter the number of bits for each Table 2 and 3 design parameter to broaden or restrict a design space search; changing the number of bits will also alter the binary string size. To elaborate, there are many more decision variables to consider when finding an optimal biolibrary of synthetic myosins, since each individual myosin has four design parameters in comparison to a biolibrary of natural myosins that has no alterable design parameters for individual myosins. Natural myosins represent already configured myosin isoforms that are extracted from existing biological systems. Therefore, evaluating the usefulness of a natural myosin biolibrary only considers decisions related to placing different available myosins in varied amounts for use in potential nanotechnology applications.

Functional requirements that represent myosin nanotechnology attributes are used to assess the performance of configured biosystem blocks. Eight functional requirements were generated according to effective rules from past studies [17,18] and represent a diverse set of metrics for evaluating how well a biolibrary enables the configuration of a variety of high-performing nanotechnologies, referred to as a biolibrary's robustness. In this sense, robustness [71,72] refers to the potential for the set of myosins that make up a biolibrary to enable the configuration of myosin-based technologies with insensitivity to the specific performance needs of each technology. A nonrobust biolibrary would be skewed toward enabling only high performance in a limited set of technologies, such as promoting actuating technologies that require myosins with force–velocity properties different than those well suited for medical diagnostics technologies.

Each functional requirement has a design goal representative of an objective function and up to two constraints. These requirements are shown in Table 4 and, as an example, the third set of requirements stipulates that an optimal biosystem block must have the lowest average number of attached myosins possible, while maintaining a filament velocity of at least 3.2 μm/s.

The functional requirements represent a wide range of conditions relevant for assessing myosin technologies, such as modulating energy effectively or achieving a fast filament translation. Force is considered as external stimuli, rather than a requirement, and a typical value of 10 pN is used for all evaluations. Having a set of eight requirements creates a large design space that makes it difficult for a human designer or a computational search to find the global optimum design. However, an initial computational search may find good designs that are then improved by humans, since each individual functional requirement is representative of a design configuration within the realm of human understanding.

6.2 Biolibrary Design Search. Biolibrary robustness is evaluated as an objective function for design searches by averaging the performance of a set of eight configured biosystem blocks, with a unique biosystem block evaluated for each Table 4 functional requirement. Robustness is a unitless value, since objective function values are nondimensionalized when evaluating Table 4 tasks to enable meaningful comparisons for tasks with different units (e.g., velocity is evaluated according to μm/s while energy is evaluated with units of ATP/ms). Specifically, the objective value for each functional requirement is assessed from 0 to 1 based on how well a biosystem block performs with respect to the global optimum performance possible for that particular functional requirement. All biosystem blocks that do not meet constraints are given a score of -1, which generally ensures that the most robust biolibraries found satisfy all requirements such that no biosystem block fails [73].

To determine an objective function score between 0 and 1, the difference in a biosystem block's goal performance and its global optimal goal performance is determined and then divided by the range of all possible goal performance values for that particular requirement. The global optimum is determined by placing an optimal synthetic myosin in each block for an optimal system size based on Tables 2 and 3 variable ranges. Therefore, the worst design that does not violate constraints always has a score of 0 and the best design has a score of 1 for each biosystem block. The evaluation of each biosystem block is straightforward using analytical models [15], which make it computationally nonintensive to assess the performance of a configured biosystem block but difficult to configure an optimal biolibrary. Although configuring and evaluating biosystem blocks is necessary for evaluating a biolibrary's robustness, biolibraries themselves only refer to a set of myosins. Having both myosin biolibrary and biosystem block configurations in a single binary string representation enables an efficient way for representing, manipulating, and evaluating a biolibrary's design and organization into biosystem blocks during design searches.

A stochastic optimization algorithm [60] was developed to search the biolibrary design space of eight synthetic myosins being configured for eight biosystem block functional requirements. The approach was modified to improve search results by separately altering the binary string portion that represents myosin biolibrary design and biosystem block configurations on alternate iterations, rather than manipulating the entire binary string on each iteration. The algorithm was run numerous times due to the stochastic nature of search results. Table 5 demonstrates the best biolibrary found after 20 runs of the algorithm for the longest possible binary string from Tables 2 and 3, which consists of 120 bits or 1.3×10^{36} possible configurations. These myosins are representative of synthetically designed isoforms, since they possess structural configurations not tied to any known myosins but are potentially manufacturable [19,58].

Table 5 shows that each configured biosystem block is unique. However, two of the myosin isoforms are used twice (for requirements two and three, and for requirements four and five), meaning the biolibrary design output by the optimization algorithm

Table 4 Biosystem block functional requirements

Req #	Design goal	Goal constraint	Secondary constraint
1	Highest avg# of attached myos	$\leq$5.7 (N_{att})	None
2	Highest filament velocity	None	System energy use $\leq$ 2.3 (ATP/ms)
3	Lowest avg# of attached myos	$\leq$5.4 (N_{att})	Filament velocity $\geq$ 3.2 (μm/s)
4	Highest energy density	None	None
5	Highest system energy use	$\leq$2.3 (ATP/ms)	None
6	Highest filament velocity	None	Avg# attached myos $\leq$ 5.8 (N_{att})
7	Lowest avg# of attached myos	$\leq$5.4 (N_{att})	System energy use $\leq$ 1.2 (ATP/ms)
8	Highest adjusted system energy	None	System energy use $\leq$ 4.0 (ATP/ms)

contains six unique myosins. Myosins from Table 5 include myosins with long (isoforms 1 and 2) and short (isoforms 3 and 4) lever arms that influence how long they remain attached generating positive force on a filament; myosins with high (isoforms 1 and 2) and low (isoform 3) attachment rates that influence how fast they consume energy, with higher energy use enabling them to cycle faster; and myosins with high (isoforms 2 and 3) and low (isoform 6) detachment rates that influence how fast they detach from a filament, which decreases the time they exert negative force that impedes filament movement.

There are no general trends for the best myosins or systems with respect to isoform/system configurations, and values for all parameters are expressed across their full range. The biolibrary also has high performance across all functional requirements, with only two of the functional requirements being less than 0.9, thus supporting the notion that this is a highly robust biolibrary. A human user could use this information to modify the biolibrary to improve objective function three and eight (the lowest performing cases) through making small changes to variables or considering configurations from alternate high performing biolibraries.

Due to the algorithm's capability for finding high-performance biolibraries in the most complex case, it is expected that the algorithm will also find near global optimum designs for less complex search cases. Based on these findings, and considering cost and performance trade-offs not explicitly considered in the computational search, it is possibly more beneficial to have a biolibrary with four isoforms of generally robust performance with a potentially lower cost than is required to manufacture and maintain six unique isoforms. Such considerations are difficult to program formally since they represent criteria unique to a particular design scenario, which further motivates the need for ensuring outputs from the optimization process are interpretable by human decision makers.

6.3 Natural and Synthetic Biolibrary Comparison. The myosins empirically measured and reverse engineered from the Discover and Describe phases resulted in design parameter values that describe two natural myosins that may be compared with synthetic isoforms. Results from such comparisons can inform scientists and engineers of differences among natural or synthetic myosins for design performance and motivate future experiments to better characterize myosins that are useful for configuring potentially high performing nanotechnologies.

To determine trade-offs in biolibraries consisting of different myosins, constraints are introduced that restrict the myosin biolibrary design space. For myosin biolibraries with only natural isoforms, a constraint is placed so myosins possess static values that reflect the measured values of existing isoforms found in nature. When these values are static, it results in the removal of genes from the binary string that represents biolibrary design decision variables, and therefore results in a smaller binary string. Once a binary string is created that reflects relevant design decision criteria, the optimization algorithm is used to optimize the configuration of the particular biolibrary for its use in Table 4 functional requirements when it is evaluated for robustness.

Since the algorithm was demonstrated to find near global optimum designs on the longest binary string possible to configure with Table 2 and 3 parameters, which are supported by Table 5 results, it is expected to perform as well or better in finding near global optimum designs on design representations with smaller binary string lengths. Therefore, differences in comparing biolibrary robustness likely reflect the highest possible robustness achievable with a given biolibrary, rather than limitations in the algorithm's capability for configuring bioblocks optimally for a given biolibrary.

In Fig. 7(*a*), biolibraries are optimized for robustness with one of the two natural isoforms placed in all eight biosystem blocks in varying amounts. The resulting binary string representation has no genes for describing myosin design parameters since all parameters have static values that reflect natural myosin configurations. All bits are retained for biosystem block genes that determine how many myosins are placed in a system, but biosystem blocks have no bits for describing which myosin isoform is chosen since there is only one natural isoform considered for placement in all systems. Design decision variables consist entirely of how many of each myosin to place in each biosystem block with functional requirements according to Table 4. If an output table for this design case was created, similar to Table 5, it would contain one myosin isoform with static values and eight biosystem blocks of potentially different sizes.

In Fig. 7(*b*), the design space is expanded by considering variable numbers of synthetic isoforms and how they are placed in different biosystem blocks. Therefore, the biolibrary portion of

Table 5 Most robust biolibrary for fulfilling biosystem block functional requirements

		Biosystem block		Myosin isoform				Illustration	
Req (#)	Objective function	Isoform (#)	N_{myo} (#)	Lever (nm)	Angle (deg)	k_{ON} (s^{-1})	k_{OFF} (s^{-1})	System	Myosin
1	0.96	1	81	14	45	1557	1914		
2	0.94	2	132	14	45	1314	3200		
3	0.79	2	47	14	45	1314	3200		
4	0.98	3	30	6	30	343	3200		
5	0.97	3	98	6	30	343	3200		
6	0.96	4	150	6	45	1071	2343		
7	0.99	5	81	8.67	30	1071	1486		
8	0.81	6	150	8.67	30	1800	200		

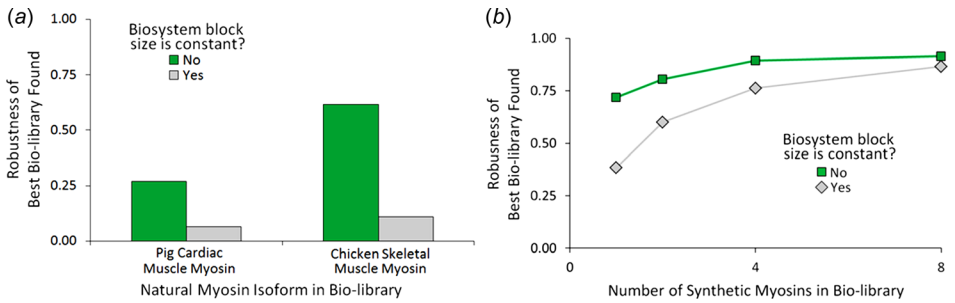

Fig. 7 Robustness of biolibraries that (*a*) consist of one of two natural isoforms and (*b*) up to eight synthetic isoforms

the binary string has a number of genes proportional to the number of designable synthetic myosins, with each myosin isoform having four genes reflecting design parameters in Table 4. The biosystem block portion of the binary string retains all bits for configuring biosystem block size, and 0, 1, 2, or 3 bits to determine which isoform is used in each block depending on whether the biolibrary has 1, 2, 4, or 8 synthetic myosins, respectively. If an output table for this design case was created similar to Table 5, it would contain a variety of myosin isoforms, up to the number of synthetic myosins allowed, and eight biosystem blocks of potentially different sizes.

An additional system level constraint is introduced for each design search to investigate the robustness a biolibrary may achieve when all biosystem block sizes are held constant. This constraint reduces the number of system level variables and therefore puts more emphasis on ensuring a diverse set of myosins are present in a biolibrary. The binary string for this configuration would change by having its genes for altering biosystem block size removed. This design space reduction is relevant to volume-limited applications, such as configuring a nano-actuator constrained to a particular length that would always require the same number of myosins. When biosystem block size is held constant, it is always representative of a system with 150 myosins, which is the number typically found operating together at the smallest scales in muscle [74]. If an output table for this design case was created, similar to Table 5, all eight biosystem blocks could potentially have different natural/synthetic myosin isoforms but would always have 150 myosins.

When considering only cases when biosystem block size is not constant, the results show that biolibraries with one synthetic isoform outperform both biolibraries with one natural isoform each. The findings suggest that the tuning of synthetic isoforms for particular applications is beneficial, and the fine-tuning of a single myosin can ensure it promotes performance across a range of potential applications for configuring nanotechnologies. Increasing the number of synthetic myosins in a biolibrary improves robustness, although there is not much gain in increasing biolibrary size from four synthetic myosins to eight synthetic myosins. Therefore, biolibraries consisting of four synthetic myosins are possibly the most cost effective when considering the cost to produce and maintain each myosin compared to the performance of the biolibrary for these potential applications. However, a final selection of the best biolibrary is dependent on contextual information currently outside of the formalized design framework that a human designer may possess, such as manufacturing and maintenance costs.

When the biosystem block size was set to a constant 150 myosins, it resulted in lower robustness for all cases, which is expected since there is a reduction in the number of design decision variables available. The drop in performance was lessened as more myosins were included in a biolibrary, such that a biolibrary design with eight synthetic myosins available had robustness that was insensitive to the constraint concerning biosystem block size, since there is always a myosin available to design appropriately for a given biosystem block.

The biolibraries and the evaluation of configured biosystem blocks at this point would require further empirical validation, as they represent extrapolations and interpolations from a model reverse engineered during the Describe phase based on limited data available from the Discover phase. In this sense, the designed biolibraries are conceptual in nature and early indicators of potentially high performing myosin isoforms and biosystem block configurations. Further steps continuing from the Develop phase would require a new iteration of the D_3 Methodology, with optimization findings informing which natural or synthetic myosins to investigate with empirical discovery and analytical description methods.

7 Discussion

The discussion section considers the feasibility of the D_3 Methodology, its use as a human-in-the-loop design approach, its findings for biolibrary development, and its potential for use in other application areas.

7.1 Feasibility of D_3 Methodology. The D_3 Methodology aims to weave concurrent scientific and design processes through phases of empirical discovery, analytical description, and technological development. Its feasibility was tested in this paper with a myosin biolibrary design problem. Each D_3 phase was implemented sequentially, with findings from earlier phases informing later phases. Empirical data output from the Design phase was considered as an input for the Describe phase, and the model validated in the Describe phase was used as an input for evaluating designs in the Develop phase. The successful transference of inputs and outputs across Discover, Describe, and Design phases resulted in the optimization of biolibraries from empirical data, which demonstrate the usefulness of the methodology.

This initial implementation provides a basis for proposing future steps that could explore the iterative and nonlinear nature of the proposed Fig. 1 D_3 Methodology, or applying the methodology for further applications, which is beyond the scope of this paper. For instance, optimization findings represent extrapolations and interpolations of myosin designs from empirical evidence, which requires further validation through new empirical experiments and modeling refinements. Additionally, new hypotheses may be tested in the Describe phase to investigate myosin behaviors not considered in the agent-based model, which is effectively a first-order approximation of myosin behavior. These considerations are important because optimization findings could deviate in an unknown manner from the reverse engineered model.

Care was taken in this implementation to extrapolate new isoforms with configurations near the empirically measured natural myosins and relevant to synthetic myosin experiments. It is likely that these extrapolations are generally accurate since they are conservative extrapolations relative to the more extreme myosin

configurations found in nature that still fundamentally operate on the same principles as myosins modeled in this study [61,75]. Such conservative extrapolations enable exploration of potentially useful designs at a much faster rate than possible through manufacturing and testing of new synthetic molecules through trial and error approaches. Empirical validation of extrapolated and interpolated molecules may be conducted through efficient design of experiment approaches that sample configuration spaces near potentially useful molecules isolated from the design searches.

7.2 Basis for Human-in-the-Loop Design. The D_3 implementation in this paper is primarily automated within each phase, but there is no automated passing of information across phases, which necessitates a human designer-in-the-loop. Including a human designer-in-the-loop has great potential for guiding nonlinear decision making across the methodological framework, since it may be difficult to automate nonlinear traversal of D_3 phases. For instance, if a modeling hypothesis in the Describe phase does not correspond to empirical data from the Discover phase, a human is likely required to generate new hypotheses or experiments by utilizing information outside of the formalized computational system. Human designers-in-the-loop should have reasoning abilities representative of common core processes from both science and design perspectives [29], which require deductive, inductive, and abductive decision making.

There are two primary aspects of the D_3 implementation informed by our previous cognitive studies. First, agent-based simulations were utilized because they can promote human reasoning of complex multilevel relationships that may directly improve design decision-making performance [18]. These agent-based simulations are representable in a structure–behavior–function framework that can promote human reasoning for design problems, especially for complex systems [15,22].

Second, the number of variables to describe and assess each biosystem block is informed by design decision-making experiments where humans were demonstrated to effectively make design decisions on similar representations [17,18]. The use of cognitive findings in the D_3 approach promotes the possibility of a designer interpreting findings from the computational search and improving them effectively. It also enables human designers to bring in information outside of the formalized computational system to make decisions. Some decisions may be obvious to a human designer but difficult to automate computationally, such as introducing a new unique myosin design to a biolibrary when a particular functional requirement is difficult to fulfill.

7.3 Development of Biolibraries. The development of biolibraries is an important step toward realizing a standardized repository of biological parts for bio-based products [30–32]. Biological repositories typically consider biochemically oriented parts; however, this study demonstrates the potential for building repositories of mechanically oriented motor proteins. Results from the Develop phase suggest that biolibraries with synthetic isoforms are more robust than those with natural isoforms, where robustness refers to a biolibrary's potential for enabling the configuration of a diverse set of myosin nanotechnologies. However, these results are limited due to only a couple of natural myosins being considered in comparison to the vast number of designable synthetic molecules considered. Further analysis of trade-offs in using natural or synthetic myosins should consider a larger set of natural myosins, a combination of both natural and synthetic myosins in a single biolibrary, and the cost of manufacturing and maintaining natural and/or synthetic isoforms.

The robustness of a biolibrary was also assessed when biosystem blocks were constrained to a particular volume, thus only differences in the types of myosins and not their number in the system are considered as design variables. The restriction of biosystem block size is representative of volume-constrained products such as myosin nano-actuators that may not exceed a particular length. It was found that more options in myosins in the biolibrary contribute to higher functioning across all biosystem blocks, which suggests the need for a diversity of myosin designs for volume-constrained product applications. Further exploration of potential products constructed from myosins could provide constraints and design scenarios that highlight other considerations for developing robust myosin biolibraries.

7.4 Broader Applicability of D_3 Methodology. The D_3 Methodology utilizes phases of empirical discovery, analytical description, and technological development poised to aid new product creation, especially in bio-based design applications where scientific experiments are required. The implementation of the methodology demonstrates its feasibility in the myosin domain and enables discussion concerning its use for further design applications. The D_3 framework is best utilized for designing new technologies where a phenomenon is already known to exist, and further empirical discoveries can promote understanding for engineering design and analysis. Although other bio-based design applications may require different specifics in computational tools utilized, the framework as a whole provides a basis for logically connecting essential processes in the research and development of new products.

In this paper, the methodology's implementation directly contributed to the basis of a human-in-the-loop approach for designing myosin biolibraries. Additionally, it provides a basis for assessing performance trade-offs in early myosin technology design, which is essential for selecting an appropriate number and type of myosins to manufacture and maintain. By identifying the myosin biolibrary design problem initially, Discover and Describe phases were driven toward results that aided design searches in the Develop phases. Without such a framework, new scientific experiments may explore potentially interesting spaces for generating knowledge of a system but provide little relevant information for design decision making.

The D_3 Methodology potentially has broad applicability to other applications, since science and design are inherently domain-independent processes in addition to the generic steps of empirical discovery, analytical description, and technological development. The framework provides a basis for embedding specific tools of a design domain in each phase, while suggesting critical inputs and outputs for communication across phases. In bio-based design, it is likely that specific computational tools will differ, such as using a variety of coarse graining approaches for model evaluations [67]. Differences in models could influence the optimal representation of a system for human-in-the-loop design approaches that are further investigated with human studies. There are a number of other motor proteins beyond myosin relevant for investigation with similar experiments and models [19]. Designing systems with these motor proteins could offer a practical first step for testing the methodology's feasibility in further bio-based applications for products with functions such as sensing, actuating, and transporting at a molecular scale [47].

8 Conclusion

The D_3 Methodology proposed and implemented in this paper integrates phases of empirical discovery, analytical description, and technological development to bridge science and design perspectives for configuring bio-based products. The methodology is set up for humans to guide computational approaches, which forms the basis of a human-in-the-loop design approach that is well suited for complex systems design. Agent-based simulations were used to model biosystems since they promote human understanding of complex systems and are adaptable to new rule sets for modeling biological phenomenon. The D_3 Methodology uses computational processes to search design spaces that are too complex for humans to traverse initially and finds individual designs with a number of variables suitable for high-level human decision making.

Optimization results from the final Develop phase suggest that using synthetic biological parts in designed products is advantageous, in comparison to only using naturally existing biological parts. These findings suggest the need for new scientific studies to characterize synthetic biological parts, thus highlighting science and design's reciprocal relationship. The successful implementation of each phase suggests that the D_3 Methodology is potentially useful for facilitating science and design endeavors and promoting the rate that scientifically investigated systems become designed technologies.

Acknowledgment

The authors would like to thank support from collaborations at Carnegie Mellon University, Harvard University, and Boston University. An early iteration of this work was accepted to the ASME International Design and Engineering Technical Conference [68].

Funding support was provided by the National Defense Science and Engineering Fellowship and the National Science Foundation under Grant No. CMMI-1160840.

References

[1] Cheong, H., and Shu, L., 2014, "Retrieving Causally Related Functions From Natural-Language Text for Biomimetic Design," ASME J. Mech. Des., **136**(8), p. 081008.
[2] Fu, K., Moreno, D., Yang, M., and Wood, K., 2014, "Bio-Inspired Design: An Overview Investigating Open Questions From the Broader Field of Design-By-Analogy," ASME J. Mech. Des., 136(11), p. 111102.
[3] Nagel, J. K., Nagel, R. L., Stone, R. B., and McAdams, D. A., 2010, "Function-Based, Biologically Inspired Concept Generation," Artif. Intell. Eng. Des., Anal. Manuf., **24**(04), pp. 521–535.
[4] Goel, A. K., Vattam, S., Wiltgen, B., and Helms, M., 2012, "Cognitive, Collaborative, Conceptual and Creative—Four Characteristics of the Next Generation of Knowledge-Based CAD Systems: A Study in Biologically Inspired Design," Comput.-Aided Des., **44**(10), pp. 879–900.
[5] Vandevenne, D., Verhaegen, P.-A., Dewulf, S., and Duflou, J. R., 2015, "A Scalable Approach for Ideation in Biologically Inspired Design," Artif. Intell. Eng. Des., Anal. Manuf., **29**(01), pp. 19–31.
[6] Cohen, Y. H., Reich, Y., and Greenberg, S., 2014, "Biomimetics: Structure–Function Patterns Approach," ASME J. Mech. Des., **136**(11), p. 111108.
[7] Cheng, A. A., and Lu, T. K., 2012, "Synthetic Biology: An Emerging Engineering Discipline," Annu. Rev. Biomed. Eng., **14**(1), pp. 155–178.
[8] Nawroth, J. C., Lee, H., Feinberg, A. W., Ripplinger, C. M., McCain, M. L., Grosberg, A., Dabiri, J. O., and Parker, K. K., 2012, "A Tissue-Engineered Jellyfish With Biomimetic Propulsion," Nat. Biotechnol., **30**(8), pp. 792–797.
[9] Pawelzik, P., Carus, M., Hotchkiss, J., Narayan, R., Selke, S., Wellisch, M., Weiss, M., Wicke, B., and Patel, M. K., 2013, "Critical Aspects in the Life Cycle Assessment (LCA) of Bio-Based Materials–Reviewing Methodologies and Deriving Recommendations," Resour., Conserv. Recycl., **73**, pp. 211–228.
[10] Becker, J., and Wittmann, C., 2012, "Bio-Based Production of Chemicals, Materials and Fuels–Corynebacterium Glutamicum as Versatile Cell Factory," Curr. Opin. Biotechnol., **23**(4), pp. 631–640.
[11] Weber, W., and Fussenegger, M., 2011, "Emerging Biomedical Applications of Synthetic Biology," Nat. Rev. Genet., **13**, pp. 21–35.
[12] Dababneh, A. B., and Ozbolat, I. T., 2014, "Bioprinting Technology: A Current State-of-the-Art Review," ASME J. Manuf. Sci. Eng., **136**(6), p. 061016.
[13] Slusarczyk, A. L., Lin, A., and Weiss, R., 2012, "Foundations for the Design and Implementation of Synthetic Genetic Circuits," Nat. Rev. Genet., **13**(6), pp. 406–420.
[14] Egan, P., Sinko, R., LeDuc, P., and Keten, S., 2015, "The Role of Mechanics in Biological and Synthetic Bioinspired Systems," Nat. Commun., **6**, p. 7418.
[15] Egan, P., Cagan, J., Schunn, C., and LeDuc, P., 2013, "Design of Complex Biologically Based Nanoscale Systems Using Multi-Agent Simulations and Structure-Behavior-Function Representations," ASME J. Mech. Des., **135**(6), p. 061005.
[16] Egan, P., Moore, J., Schunn, C., Cagan, J., and LeDuc, P., 2015, "Emergent Systems Energy Laws for Predicting Myosin Ensemble Processivity," PLoS Comput. Biol., **11**(4), p. e1004177.
[17] Egan, P., Cagan, J., Schunn, C., and LeDuc, P., 2015, "Synergistic Human-Agent Methods for Deriving Effective Search Strategies: The Case of Nanoscale Design," Res. Eng. Des., **26**(2), pp. 145–169.
[18] Egan, P., Schunn, C., Cagan, J., and LeDuc, P., 2015, "Improving Human Understanding and Design of Complex Multi-Level Systems With Animation and Parametric Relationship Supports," Des. Sci., **e3**.
[19] Howard, J., 2001, *Mechanics of Motor Proteins and the Cytoskeleton*, Sinauer Associates, Sunderland, MA.
[20] Hodges, A. R., Krementsova, E. B., and Trybus, K. M., 2007, "Engineering the Processive Run Length of Myosin V," J. Biol. Chem., **282**(37), pp. 27192–27197.
[21] Smolke, C. D., and Silver, P. A., 2011, "Informing Biological Design by Integration of Systems and Synthetic Biology," Cell, **144**(6), pp. 855–859.
[22] Vattam, S. S., Goel, A. K., Rugaber, S., Hmelo-Silver, C. E., Jordan, R., Gray, S., and Sinha, S., 2011, "Understanding Complex Natural Systems by Articulating Structure-Behavior-Function Models," Educ. Technol. Soc., Special Issue on Creative Design, **14**(1), pp. 66–81.
[23] Togashi, Y., Yanagida, T., and Mikhailov, A. S., 2010, "Nonlinearity of Mechanochemical Motions in Motor Proteins," PLoS Comput. Biol., **6**(6), p. e1000814.
[24] Marcucci, L., and Yanagida, T., 2012, "From Single Molecule Fluctuations to Muscle Contraction: A Brownian Model of AF Huxley's Hypotheses," PloS One, **7**(7), p. e40042.
[25] Simpson, T. W., and Martins, J. R., 2011, "Multidisciplinary Design Optimization for Complex Engineered Systems: Report From a National Science Foundation Workshop," ASME J. Mech. Des., **133**(10), p. 101002.
[26] Simpson, T. W., Carlsen, D., Malone, M., and Kollat, J., 2011, "Trade Space Exploration: Assessing the Benefits of Putting Designers 'Back-in-the-Loop' During Engineering Optimization," *Human-in-the-Loop Simulations*, Springer, Heidelberg, Germany, pp. 131–152.
[27] Kell, D. B., and Oliver, S. G., 2004, "Here is the Evidence, Now What is the Hypothesis? The Complementary Roles of Inductive and Hypothesis-Driven Science in the Post-Genomic Era," Bioessays, **26**(1), pp. 99–105.
[28] Galle, P., and Kroes, P., 2014, "Science and Design: Identical Twins?," Des. Stud., **35**(3), pp. 201–231.
[29] Dorst, K., 2011, "The Core of 'Design Thinking' and Its Application," Des. Stud., **32**(6), pp. 521–532.
[30] Smolke, C. D., 2009, "Building Outside of the Box: iGEM and the BioBricks Foundation," Nat. Biotechnol., **27**(12), pp. 1099–1102.
[31] Galdzicki, M., Rodriguez, C., Chandran, D., Sauro, H. M., and Gennari, J. H., 2011, "Standard Biological Parts Knowledgebase," PloS One, **6**(2), p. e17005.
[32] Cooling, M. T., Rouilly, V., Misirli, G., Lawson, J., Yu, T., Hallinan, J., and Wipat, A., 2010, "Standard Virtual Biological Parts: A Repository of Modular Modeling Components for Synthetic Biology," Bioinformatics, **26**(7), pp. 925–931.
[33] Luo, X., Li, W., Kwong, C., and Cao, Y., 2016, "Optimisation of Product Family Design With Consideration of Supply Risk and Discount," Res. Eng. Des., **27**(1), pp. 37–54.
[34] Farrell, R., and Hooker, C., 2012, "The Simon–Kroes Model of Technical Artifacts and the Distinction Between Science and Design," Des. Stud., **33**(5), pp. 480–495.
[35] Barnes, C. P., Silk, D., Sheng, X., and Stumpf, M. P., 2011, "Bayesian Design of Synthetic Biological Systems," Proc. Natl. Acad. Sci., **108**(37), pp. 15190–15195.
[36] Farrell, R., and Hooker, C., 2015, "Designing and Sciencing: Response to Galle and Kroes," Des. Stud., **37**, pp. 1–11.
[37] Cagan, J., Kotovsky, K., and Simon, H. A., 2005, "Scientific Discovery and Inventive Engineering Design," *Formal Engineering Design Synthesis*, Cambridge University Press, New York, pp. 442–465.
[38] Munch, E., Launey, M. E., Alsem, D. H., Saiz, E., Tomsia, A. P., and Ritchie, R. O., 2008, "Tough, Bio-Inspired Hybrid Materials," Science, **322**(5907), pp. 1516–1520.
[39] Spatz, J. P., 2005, "Bio-MEMS: Building up Micromuscles," Nat. Mater., **4**(2), pp. 115–116.
[40] Cvetkovic, C., Raman, R., Chan, V., Williams, B. J., Tolish, M., Bajaj, P., Sakar, M. S., Asada, H. H., Saif, M. T. A., and Bashir, R., 2014, "Three-Dimensionally Printed Biological Machines Powered by Skeletal Muscle," Proc. Natl. Acad. Sci., **111**(28), pp. 10125–10130.
[41] Koenderink, G. H., Dogic, Z., Nakamura, F., Bendix, P. M., MacKintosh, F. C., Hartwig, J. H., Stossel, T. P., and Weitz, D. A., 2009, "An Active Biopolymer Network Controlled by Molecular Motors," PNAS, **106**(36), pp. 15192–15197.
[42] Neiman, V. J., and Varghese, S., 2011, "Synthetic Bio-Actuators and Their Applications in Biomedicine," Smart Struct. Syst., **7**(3), pp. 185–198.
[43] Ayala, F. J., 2007, "Darwin's Greatest Discovery: Design Without Designer," Proc. Natl. Acad. Sci., **104**(1), pp. 8567–8573.
[44] Bölker, M., 2015, "Complexity in Synthetic Biology: Unnecessary or Essential?," *Synthetic Biology*, Springer, Heidelberg, Germany, pp. 59–69.
[45] Miller, G. A., 1956, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychol. Rev., **63**(2), p. 81.
[46] Hirschi, N., and Frey, D., 2002, "Cognition and Complexity: An Experiment on the Effect of Coupling in Parameter Design," Res. Eng. Des., **13**(3), pp. 123–131.
[47] Hess, H., 2011, "Engineering Applications of Biomolecular Motors," Annu. Rev. Biomed. Eng., **13**(1), pp. 429–450.
[48] Gero, J. S., and Kannengiesser, U., 2003, "The Situated Function-Behaviour-Structure Framework," Des. Stud., **25**(4), pp. 373–391.
[49] Goel, A. K., Rugaber, S., and Vattam, S., 2009, "Structure, Behavior, and Function of Complex Systems: The Structure, Behavior, and Function Modeling Language," Artif. Intell. Eng. Des., Anal. Manuf., **23**(01), pp. 23–35.
[50] Eastman, C. M., 1982, "Recent Developments in Representation in the Science of Design," Des. Stud., **3**(1), pp. 45–52.
[51] Björklund, T. A., 2013, "Initial Mental Representations of Design Problems: Differences Between Experts and Novices," Des. Stud., **34**(2), pp. 135–160.
[52] Hwang, M., Garbey, M., Berceli, S. A., and Tran-Son-Tay, R., 2009, "Rule-Based Simulation of Multi-Cellular Biological Systems—A Review of Modeling Techniques," Cell. Mol. Bioeng., **2**(3), pp. 285–294.

[53] Michael, S., Faeder, J., and Emonet, T., 2011, "Efficient Modeling, Simulation and Coarse-Graining of Biological Complexity With NFsim," Nat. Methods, **8**(2), pp. 177–183.
[54] Bongard, J., and Lipson, H., 2007, "Automated Reverse Engineering of Nonlinear Dynamical Systems," Proc. Natl. Acad. Sci., **104**(24), pp. 9943–9948.
[55] Schmidt, M., and Lipson, H., 2009, "Distilling Free-Form Natural Laws From Experimental Data," Science, **324**(5923), pp. 81–85.
[56] Gay, G., Menzies, T., Davies, M., and Gundy-Burlet, K., 2010, "Automatically Finding the Control Variables for Complex System Behavior," Autom. Software Eng., **17**(4), pp. 439–468.
[57] Resnicow, D. I., Deacon, J. C., Warrick, H. M., Spudich, J. A., and Leinwand, L. A., 2010, "Functional Diversity Among a Family of Human Skeletal Muscle Myosin Motors," PNAS, **107**(3), pp. 1053–1058.
[58] Tsiavaliaris, G., Fujita-Becker, S., and Manstein, D. J., 2004, "Molecular Engineering of a Backwards-Moving Myosin Motor," Nature, **427**(6974), pp. 558–561.
[59] Campbell, M. I., Cagan, J., and Kotovsky, K., 1999, "A-Design: An Agent-Based Approach to Conceptual Design in a Dynamic Environment," Res. Eng. Des., **11**(3), pp. 172–192.
[60] Landry, L., and Cagan, J., 2011, "Protocol-Based Multi-Agent Systems: Examining the Effect of Diversity, Dynamism, and Cooperation in Heuristic Optimization Approaches," ASME J. Mech. Des., **133**(2), p. 021001.
[61] Ito, K., Ikebe, M., Kashiyama, T., Mogami, T., Kon, T., and Yamamoto, K., 2007, "Kinetic Mechanism of the Fastest Motor Protein, Chara Myosin," J. Biol. Chem., **282**(27), pp. 19534–19545.
[62] Taniguchi, M., Okamoto, R., Ito, M., Goto, I., Fujita, S., Konishi, K., Mizutani, H., Dohi, K., Hartshorne, D. J., and Itoh, T., 2015, "New Isoform of Cardiac Myosin Light Chain Kinase and the Role of Cardiac Myosin Phosphorylation in α 1-Adrenoceptor Mediated Inotropic Response," PLoS One, **10**(10), p. e0141130.
[63] Work, S., and Warshaw, D., 1991, "Computer-Assisted Tracking of Actin Filament Motility," Anal. Biochem., **202**(2), pp. 275–285.
[64] Marston, S., Fraser, I., Bing, W., and Roper, G., 1996, "A Simple Method for Automatic Tracking of Actin Filaments in the Motility Assay," J. Muscle Res. Cell Motil., **17**(4), pp. 497–506.
[65] Meijering, E., Dzyubachyk, O., and Smal, I., 2012, "Methods for Cell and Particle Tracking," Methods Enzymol., **504**, pp. 183–200.
[66] Greenberg, M. J., and Moore, J., 2010, "The Molecular Basis of Frictional Loads in the In Vitro Motility Assay With Applications to the Study of the Loaded Mechanochemistry of Molecular Motors," Cytoskeleton, **67**(5), pp. 273–285.
[67] Saunders, M. G., and Voth, G. A., 2013, "Coarse-Graining Methods for Computational Biology," Annu. Rev. Biophys., **42**(1), pp. 73–93.
[68] Egan, P., Cagan, J., Schunn, C., LeDuc, P., Moore, J., and Chiu, F., 2015, "The d3 Science-to-Design Methodology: Automated and Cognitive-Based Processes for Discovering, Describing, and Designing Complex Nanomechanical Biosystems," ASME Paper No. DETC2015-47466.
[69] Ferrer, J. M., Lee, H., Chen, J., Pelz, B., Nakamura, F., Kamm, R. D., and Lang, M. J., 2008, "Measuring Molecular Rupture Forces Between Single Actin Filaments and Actin-Binding Proteins," Proc. Natl. Acad. Sci., **105**(27), pp. 9221–9226.
[70] Harris, D. E., Work, S. S., Wright, R. K., Alpert, N. R., and Warshaw, D. M., 1994, "Smooth, Cardiac and Skeletal Muscle Myosin Force and Motion Generation Assessed by Cross-Bridge Mechanical Interactions In Vitro," J. Muscle Res. Cell Motil., **15**(1), pp. 11–19.
[71] Mesmer, B. L., and Bloebaum, C. L., 2015, "An End-User Decision Model With Information Representation for Improved Performance and Robustness in Complex System Design," Res. Eng. Des., 26(3), pp. 233–251.
[72] Daniels, B., Chen, Y., Sethna, J., Gutenkunst, R., and Myers, C., 2008, "Sloppiness, Robustness, and Evolvability in Systems Biology," Curr. Opin. Biotechnol., **19**(4), pp. 389–395.
[73] Schwartz, B., Ben-Haim, Y., and Dacso, C., 2011, "What Makes a Good Decision? Robust Satisficing as a Normative Standard of Rational Decision Making," J. Theory Soc. Behav., **41**(2), pp. 209–227.
[74] Randall, D., Burggren, W., and French, K., 2001, *Eckert Animal Physiology: Mechanisms and Adaptations*, W. H. Freeman and Company, New York.
[75] Rome, L. C., 2005, "Design and Function of Superfast Muscles," Annu. Rev. Physiol., **68**, pp. 193–221.