Dynamics and Regulation of Cytoskeletal Proteins

by

Shashank K. Jariwala

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Bioinformatics)
in The University of Michigan
2019

Doctoral Committee:

       Professor David S. Sept, Chair
       Professor Daniel M. Burns Jr.
       Professor Heather A. Carlson
       Assistant Professor Peter Freddolino
       Professor Kristen J. Verhey

Shashank Jariwala

jari@umich.edu

ORCID ID: 0000-0001-5384-7393

To all the international students around the world pursuing their dreams away from home, and their families.

ACKNOWLEDGEMENTS

I have thoroughly enjoyed every bit of this journey. And that has been possible only due to the constant support and help by many souls along the way. First and foremost, I would like to express my deep gratitude to my advisor David Sept. I have greatly benefited from his passion for science and expertise, and his guidance in all my research projects. I also want to thank him for believing in me and accepting me, especially during a time when my faith in my journey as a graduate student was a bit shaken (but not stirred). I would also like to thank Kristen Verhey, who has been an important advisor to me during my graduate research. I had an enriching experience collaborating with her on a whole gamut of research projects in the cytoskeleton field. I also want to extend thanks to members of the Verhey lab, especially Breane Budaitis, Yang Yue, and my dear friend Jeremy Welsch.

I also want to express my gratitude to Barry Grant, who introduced me to the world of protein allosteric regulation, and inspiring me to choose the research life. Along with him, I would also like to thank Xinqiu Yao, Guido Scarabelli, Hongyang Li, and Lars Skjærven. I benefited tremendously from their expertise and constant guidance. I may have met them first as lab mates and collaborators, but they have become my friends.

I especially want to thank Daniel Burns, who has helped me at every turn during my graduate studies. I also thank the rest of my thesis committee members, Heather Carlson, and Peter Freddolino, for their advice and suggestions on my research projects. I would also like to thank Margit Burmeister and Maureen Sartor, for

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

LIST OF APPENDICES

Appendix

ABSTRACT

In this dissertation, I apply molecular dynamics (MD) simulations to improve our understanding of the dynamics, and hence, function and regulation of cytoskeletal proteins. Microtubules and kinesin motor proteins play a critical role in the cytoskeleton of the cell, providing structural support, facilitating cellular transport, beating of cilia and flagella, and separation of chromosomes during the cell cycle. The importance of tubulin dynamics as a vital therapeutic target is exemplified by the widely prescribed paclitaxel (Taxol), an anti-cancer drug that arrests tubulin dynamics, preventing cancer cells from undergoing cell division. Furthermore, the importance of understanding the structural, dynamical, and functional aspects of kinesin motor domains and their modifications is demonstrated by efforts in developing small-molecule inhibitors as antimitotic therapeutic agents in various cancers. However, despite strong conservation of the motor domain across the kinesin superfamily, how various kinesins have tailored their motility characteristics to best meet their functional needs in cells remains unclear.

Detailed comparison of structures from large heterogeneous protein families, such as kinesin motors, can inform on structural dynamic mechanisms critical for protein function including ligand binding, enzymatic catalysis, allosteric regulation and bimolecular recognition. However, existing tools for quantitative analysis of their sequence, structure and dynamics often require significant computational expertise and typically remain accessible only to expert users with relevant programming skills. In the first section of my dissertation, I describe the development of Bio3D-web, a free and open-source online application for interactive investigation of protein sequence-

structure-dynamic relationships. Bio3D-web requires no programming knowledge and thus decreases the entry barrier to performing advanced comparative structural bioinformatics analyses.

In second part of my dissertation, I discuss methods I helped develop in analyzing experimental structure data and dynamical data generated with MD simulations. Specifically, I discuss the ensemble distance difference matrix method (eDDM) for analyzing changes in residue-residue distances in protein structures and dynamical data to identify residues critical for protein regulation and function. I apply eDDM to three families of kinesin motor proteins: first, I elucidate the effect of a posttranslational modification in kinesin 5 mitotic motor Eg5. I show that acetylation of residue K146 in Eg5 alters its mechanochemical properties, wherein it acts as a "brake" during spindle separation in cells during mitosis. Second, I identify residues critical for force generation in kinesin 1 transport motor KIF5C. Mutating these residues in two important structural elements—A5G and S8G in the cover strand and N334A in the neck linker—severely cripple the ability of motors in ensemble to generate force during intracellular transport. Third, I characterize the allosteric effects of disease-associated variants in kinesin 3 neuronal transport motor KIF1A. KIF1A-associated neurological disorder (KAND) is associated with cognitive disability, spasticity, and cerebellar and optic nerve atrophy, typically with a progressive course.

In the third part of my dissertation, I highlight the divergent mechanism of tubulin polymerization in *C. elegans*. Through comparative analysis of MD simulations of *C. elegans* and *B. taurus* tubulin dimers, I found that sequence changes in the *C. elegans* tubulin lead to additional secondary structure formation in the lateral contact loops, and this changes the polymerization behavior as well as the structure of the microtubule. Finally, I also map the inter-conformer relationships of experimentally determined structures of tubulin and microtubules through principal component analysis (PCA). This PC-space serves as a reference map for comparing the intrinsic

dynamics of tubulin heterodimers, such as different isoforms, nucleotide states, and disease-associated mutations.

CHAPTER I

Thesis overview

My research interests broadly involve the study of protein dynamics as a regulatory element in protein function. Allostery, the coupling of structural dynamic changes at distal sites, is a key regulatory mechanism in numerous biomolecular processes, including ligand-binding and enzyme catalysis (*Boehr et al.*, 2009; *Henzler-Wildman and Kern*, 2007). Allosteric regulation is also an important factor to consider in rational drug design and protein engineering (*Wagner et al.*, 2016; *Raman*, 2018). Cytoskeletal proteins are also dynamic entities. Microtubules and kinesin motor proteins play a critical role in the cytoskeleton of the cell, providing structural support, facilitating cellular transport, beating of cilia and flagella, and separation of chromosomes during the cell cycle. The importance of tubulin as a vital therapeutic target is exemplified by the widely prescribed paclitaxel (Taxol), an anti-cancer drug that arrests tubulin dynamics, preventing cancer cells from undergoing cell division. Furthermore, the importance of understanding the structural, dynamical, and functional aspects of kinesin motor domains and their modifications is demonstrated by efforts in developing small-molecule inhibitors as antimitotic therapeutic agents in various cancers (*Perez-Melero*, 2014). In this dissertation, I use molecular dynamics (MD) simulations and methods to analyze MD trajectories to improve our understanding of the dynamics, and hence, function and regulation of cytoskeletal proteins. An

overview of studies covered in this dissertation are given below. I start with the development of online tools for rapid analysis of sequence, strucural, and dynamical features of large protein superfamilies, such as those of tubulin and kinesin motors. Next, I discuss my contributions to improving our knowledge about the dynamical effects of kinesin mutations and posttranslational modification on motor function. Finally, through sequence analysis and analyzing dynamics of free tubulin heterodimers in exclicit solvent MD simulations, I uncover how subtle sequence changes in tubulin sequence can give rise to different microtubule polymerization rates *in vitro*.

1.0.1 Accessible resources for structural bioinformatics

One of the first steps in understanding biomolecular functional regulation is an assessment of the sequence-structure-dynamics relationships within the given family or superfamily of proteins. To this end, I developed Bio3D-web and related tools for asssessment of protein sequence, structure and dynamics in an interactive and intuitive fashion. Bio3D-web requires no prior programming knowledge, thus crossing the barrier to entry for performing structural bioinformatics analyses. Bio3D-web is based on R-Shiny, a reactive programming and web application framework from RStudio (*RStudio*, 2013). The main purpose behind selecting Shiny for powering Bio3D-web was its reactive framework. Shiny's reactivity feature binds web controls to R functionality in a manner that caches expensive computational steps so that they are not re-computed unnecessarily during an interactive browser session. A brief overview of this feature is given in Figure 1.1. We can wrap computationally expensive operations in a "reactive" routine such that the output of operations are cached, and can be utilized by any number of dependent routines without re-executing such operations. Furthermore, any changes that may require re-executing such operations are automatically progated to the dependent routines, without having to explicitly check for updates.

A

```
# Calculate RMSD and perform clustering
reactCluster <- reactive({ rmsdCluster() )}

# Color 3D structures by cluster membership
renderWebGL({
  reactCluster()
  bio3d::view.pdb()
})

# RMSD cluster tree
renderPlot({
  reactCluster()
  plot()
})
```

B
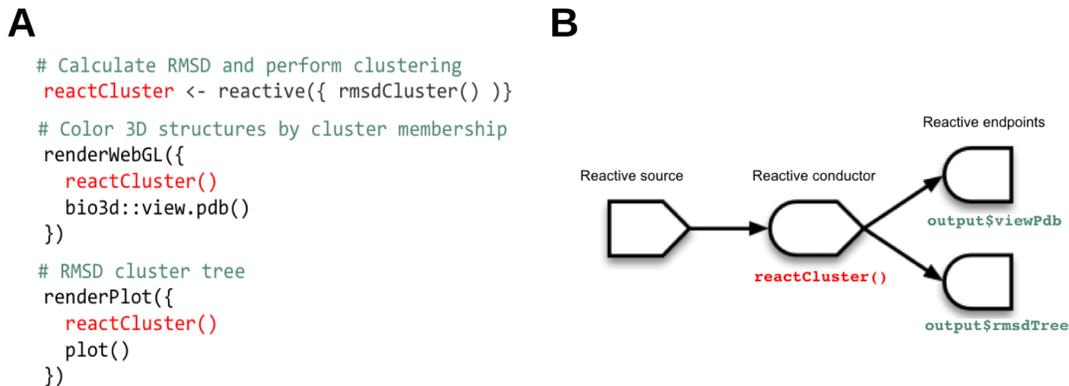
Figure 1.1: Encapsulating computationally expensive operations with Shiny. Bio3D-web makes use of Shiny's reactivity feature (**A**), where results of computationally expensive operations are cached, and any changes are automatically propagated to dependent user-facing renders (**B**).

For example we want to fit a large number of 3D structures from a particular protein family, calculate their all-to-all root mean-squared deviations (RMSD), perform RMSD-based clustering, and display the superimposed 3D structures colored by RMSD cluster membership. In this case, fitting the structures, calculating RMSD, and performing clustering can be considered as computationally expensive operations. These operations, part of a function called *rmsdCluster()* are wrapped in a reactive routine called *reactCluster()* (Figure 1.1A). A render of superimposed 3D structures, and a plot of the RMSD clustering results, colored by cluster membership, both access the cashed results of the *reactCluster()* routine. Any changes in the original *rmsd-Cluster()* function, such as addition of more protein structures, or changing clutering parameters, are automatically propagated to the 3D render and plotting functions, without any further effort on the user's part (and thankfully, the developer). We adopted this framework to develop Bio3D-web family of tools for performing structural bioinformatics analyses on the web (Figure 1.2). Details into Bio3D-web, and its applications to two protein families, are given in Chapter II (*Skjærven et al.*, 2016; *Jariwala et al.*, 2017). All major browsers are supported, just point them to

`http://bio3d.ucsd.edu/pca-app`. The full source code is available under a GPLv2 license from `https://bitbucket.org/Grantlab/bio3d`. The accessbile, interactive, and reproducible feature of my application has not only helped researchers, it is also used for classroom teaching and workshops—for example by Dr. Patrick Fleming at the Biophysics Department at Johns Hopkins University for teaching undergraduate courses.
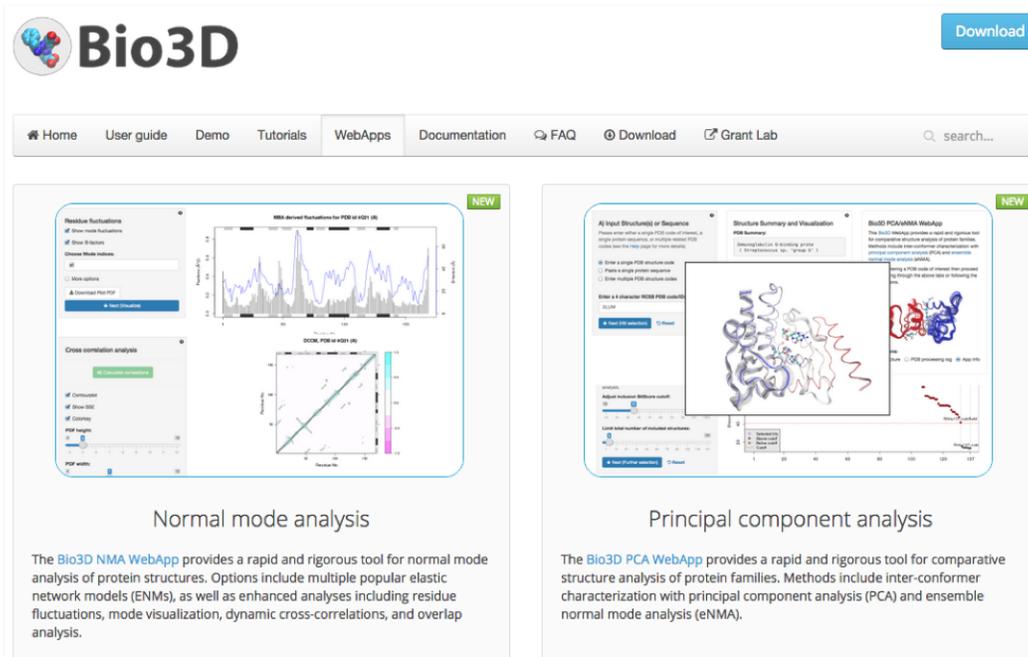
Figure 1.2: Bio3D-web tools for online structural bioinformatics. The tools are available for free at `http://bio3d.ucsd.edu/pca-app`, and `http://thegrantlab.org/bio3d/webapps`. Full source-code is open to all and is provided under a GPLv2 open-source license from: `https://bitbucket.org/Grantlab/bio3d`.

1.0.2 Dynamics of cytoskeletal proteins

The proteins that make up the cytoskeleton have both structural and functional roles in cells. Their critical role in cell viability arises from an interplay of both these roles, which include maintaining structural integrity of the cell, arrangement of organelles, beating of cilia and flagella, cell division, and facilitating cellular transport.

In general, the cytoskeleton refers to an interconnected network of filamentous polymers and associated proteins. The three primary types of cytoskeletal polymers are microtubules, actin filaments, and intermediate filaments (Figure 1.3). Microtubules, made up of tubulin subunits, and actin filaments, made up of actin subunits, undergo nucleotide-coupled polymerization and depolymerization events. These events generate directed forces, driving changes in cell shape and, together with molecular motors that move along these respecitve polymers, facilitate intra-cellular transport and guide the organization of cellular components (*Fletcher and Mullins*, 2010). Intermediate filaments, like microtubules and actin filaments, are also polymers, consisting of subunits such as keratins and lamins. In contrast, however, intermediate filaments lack polarity, are more stable and their constituent subunits do not bind nucleotides (such as GTP and ATP) (*Goldman et al.*, 2012). Their primary role is to provide structural support against mechanical stress, forming complex networks both within and with microtubule and actin filament networks.

1.0.2.1 Dynamics of kinesin motors and tubulin

The focus of this dissertation, apart from developing online structural bioinformatics tools, is on two classes of proteins belonging to the cytoskeleton: motor transport protein kinesin and microtubule (MT) subunit tubulin. Members of the kinesin superfamily of molecular motors fulfill multifaceted roles in cell physiology. Some serve as intracellular transporters (Figure 1.4), others serve as regulators of MT dynamics, including as MT depolymerizing agents (*Walczak et al.*, 2013), and some as mitotic motors (*Hirokawa and Takemura*, 2004; *Vicente and Wordeman*, 2015; *Bachmann and Straube*, 2015).

While diverse in their sequence, localization, and function, all kinesin motors contain both an ATP coordinating/hydrolyzing site and MT-binding elements, and in general, bind MTs and hydrolyze ATP to drive MT plus-end directed motor activ-
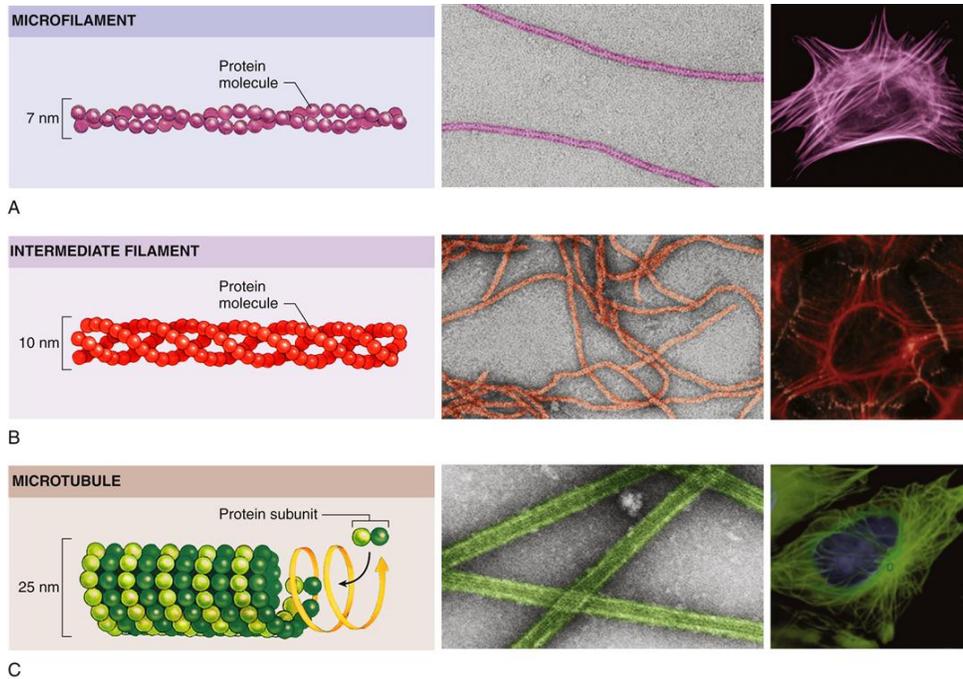
Figure 1.3: Polymers of the cell cytoskeleton. Schematics showing the physical dimensions and micrographs of the three primary types of cytoskeletal polymers. (**A**) Actin filaments (actin), (**B**) intermediate filaments, commonly made up of lamin subunits, and (**C**) microtutubles (tubulin heterodimer). Figure from (*Pollard et al.*, 2017).

ity (*Vale and Fletterick*, 1997). ATP binding brings about conformation changes to the active (ATPase) site, and this local conformation change causes allosteric effects transmitted through the motor domain that alter MT binding and neck linker orientation (*Atherton et al.*, 2014; *Goulet et al.*, 2014). Two key components of kinesin motor domain distal to the active site, a N-terminal sequence called the cover strand (CS), and a neck linker, interact to generate a "power stroke" for force production (*Hwang et al.*, 2008; *Khalil et al.*, 2008). The ATP-induced power stroke involves formation of a two-stranded β sheet called the cover-neck bundle (CNB) between the CS and N-terminal half of the NL. Following CNB formation, the second half of the NL "latches" to the core domain, fully docking with the kinesin core domain. A schematic of this CNB-assisted stepping mechanism is shown in Figure 1.5.

In Chapter III, Chapter IV, and Chapter V, I study the direct or indirect (al-
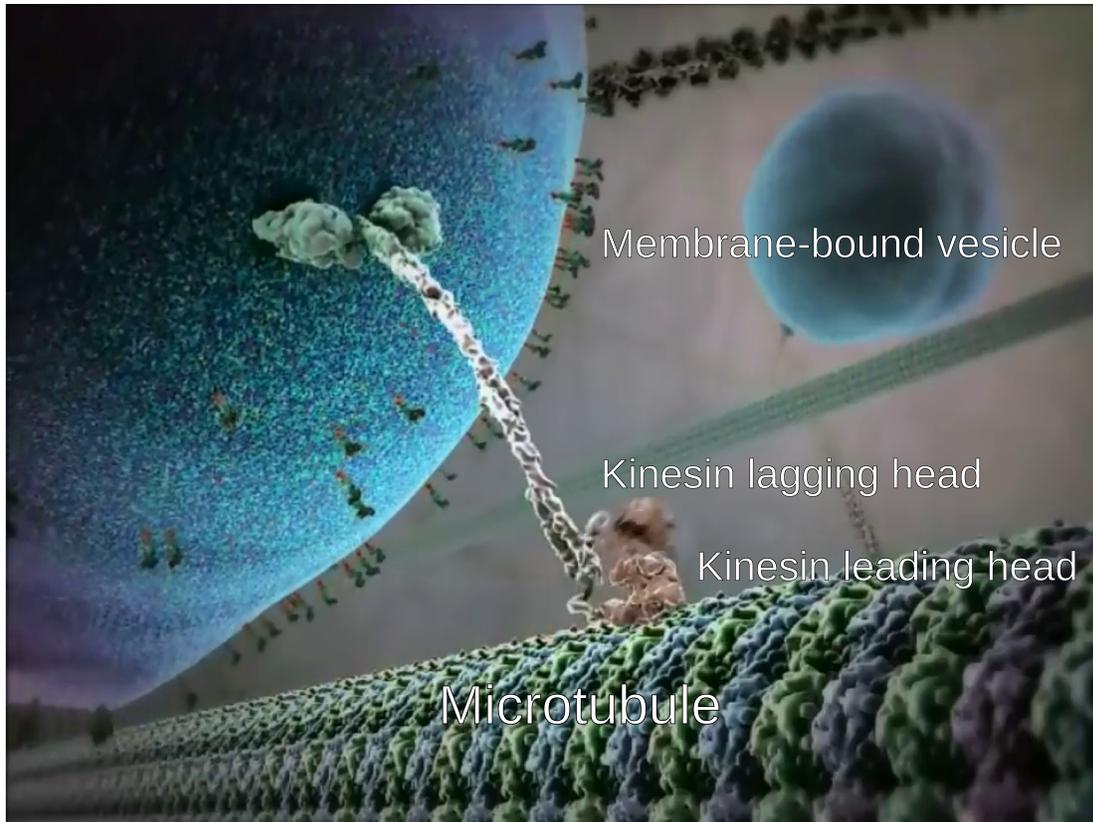
Figure 1.4: Kinesin walking on a microtubule. This rendering shows a dimeric kinesin motor walking on a microtubule filament in a "hand-over-hand" to transport cargo within a cell. Figure modified from The Inner Life series, BioVision at Harvard University (*Lue et al.*).

losteric) effects of posttranslational modifications, directed mutagenesis, and disease-associated variants on kinesin mechanochemistry and physiological function. Post-translational modifications (PTMs) have been identified in most kinesins, including many in the catalytic domain of these motor proteins (*Hornbeck et al.*, 2015; *Liu et al.*, 2014). Apart from mutations occurring near the active site or the MT-binding elements, the effects of these PTMs on motor function remain unknown. In Chapter III, I study the dynamical effects of an acetylation of lysine residue 146 (K146) of kinesin-5 mitotic motor Eg5 (*Choudhary et al.*, 2009; *Sol et al.*, 2012; *Nalawansha et al.*, 2017). Lysine 146, at the C-terminal end of helix $\alpha 2$ makes a salt bridge to aspartate 91 on helix $\alpha 1$, and its acetylation disrupts this ionic interaction. Using
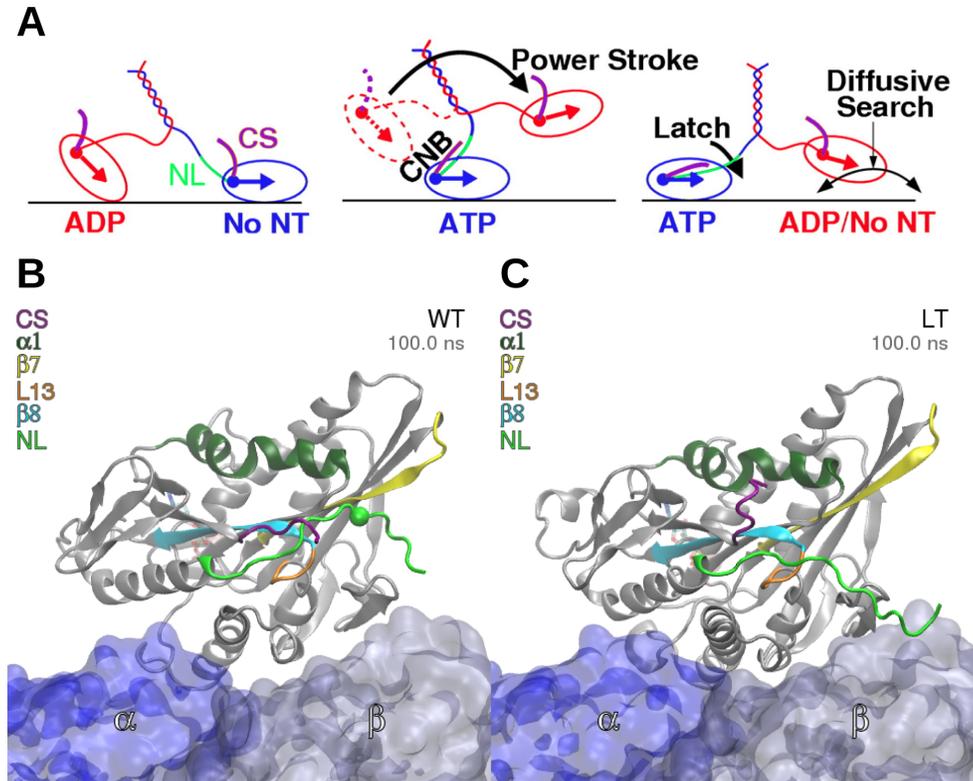
Figure 1.5: CNB-assisted stepping mechanism in kinesin motor. (**A**) ATP binding-induced formation of CNB provides the "power stroke", shunting the rear kinesin head forward, with NL docking completing the directionality of the step towards the MT plus-end (direction indicated by arrow heads). Panel adapted from (*Hwang et al.*, 2008). (**B**) Snapshot from molecular dynamics simulation of ATP-bound kinesin-1 motor in complex with tubulin heterodimer showing NL remains latched to core domain. (**C**) Disruption of the same latch in simulation of a mutant kinesin-1 motor (see Chapter IV for details).

molecular dynamics simulations and subsequent analyses, I uncover that disrupting this salt bridge enhances NL docking and increasing the allosteric coupling between the NL and the active site, improving the motility properties of this mitotic motor. Subsequent experiments using acetylation mimetic mutation K146Q support this prediction, showing enhanced coordination of NL docking and catalytic closure of the switch 1 in the active site. The interplay of enhanced NL docking and increase of NL and active site coordination was further tested under physiological conditions. Cell-based assays show that modified Eg5 in teams of motors act as a "brake", slowing

centrosome separation during mitosis (*Muretta et al.*, 2018).

In Chapter IV, I highlight key residues involved in CNB-formation and NL docking by studying the dynamics of kinesin-1 transport motor KIF5C in ATP-bound and nucleotide-free states. Using inter-residue distance difference analysis, I hypothesize that mutants disrupting CNB-formation and/or NL docking have an allosteric effect of enhancing active-site interactions. Subsequent motility assays indeed show that mutant kinesin-1 motors exhibit enhanced motility properties under single-molecule conditions. Under the hypothesis that CS and NL are mechanical elements of force generation in kinesin-1 motors, disrupting the CNB-formation and NL docking was predicted to impair force production. Indeed, the enhanced motility properties of mutant motors were not sufficient to overcome reduction in force generated to transport cargo under physiological conditions. Kinesin-1 motors with mutations that disrupted CNB-formation and/or NL docking severely reduced force generation, and were crippled in their ability to transport high-load cargo in cells (*Budaitis et al.*, 2019, *in final revision at eLife*).

Finally, in Chapter V, I study the dynamical effects of disease-associated mutations in kinesin-3 neuronal transport motor KIF1A. KIF1A is part of an important class of neuronal plus-end directed transport motors, primarily transporting synaptic vesicle precursors in cells (*Hirokawa et al.*, 2009). While certain isoforms of kinesin-1 are also neuron specific transport motor (such as KIF5A and KIF5C), studies have shown that it is the kinesin-3 family of motor proteins that are involved in long-range transport, with strikingly different motility properties: they are 10-fold more processive than the canonical kinesin-1 motors, and are called "superprocessive" motors (*Soppina et al.*, 2014). A number of disease-associated genetic variants and *de novo* mutations have been identified from clinical studies in the recent years (*Lee et al.*, 2015; *Esmaeeli Nieh et al.*, 2015; *Hotchkiss et al.*, 2016; *Iqbal et al.*, 2017). These mutations have been linked to neurodevelopmental disorders, typically with a

progressive course. The disease, collectively termed KIF1A Associated Neurological Disorder (KAND), is associated with cognitive disability, spasticity, and cerebellar and optic nerve atrophy (*Okamoto et al.*, 2014; *Ohba et al.*, 2015). How these variants and mutations affect KIF1A motor function is still not understood. Mapping currently known variants onto the 3D structure of KIF1A motor domain reveals a clustering around the ATP binding/hydrolyzing active and at the MT-interfacing regions. A third group of variants, away from these two regions critical for motor function, present a challenge: how do these variants and *de novo* mutations affect motor function, and how are they linked to disease? In this dissertation, I study a subset of these mutations and present results for two such mutations. A valine to methionine autosomal dominant mutation V8M (situated in $\beta1$, adjacent to the CS), allosterically impairs the ATPase activity of KIF1A motor domain. The hypothesis is further supported by single-molecule motility assays, which show that indeed V8M mutant motors are slower than the wildtype (WT). Another mutation, a *de novo* mutation of Y89D (situated in $\alpha1/\beta3$), causes neck linker undocking. In view of previous studies in this dissertion (pertaining kinesin-5 motor Eg5 in Chapter III, and kinesin-1 motor KIF5C in Chapter IV), I hypothesize this mutation to affect force generation in KIF1A motor. Experimental assays characterizing motility and force generation in KIF1A Y89D mutant motors are currently underway.

In the last section of this dissertaion, I present my contributions to a better undestanding of MT function and regulation through studying intrinsic dynamics of tubulin heterodimer. MTs are composed of dimer subunits made up of structurally homologous α- and β-tubulin subunits. MTs undergo nucleotide-coupled phases of stochastic growth and catastrophe events, terms "dynamic instability" (Figure 1.6). Dynamic instability is central to MT function and regulation in cells (*Kueh and Mitchison*, 2009). All euakryotic microtubules studied to date exhibit dynamic instability under *in vitro* conditions. However, we lack a clear understanding of how

subtle sequence changes in tubulin lead to differences in the dynamical properties of MTs. In Chapter VI, I compare the sequence, structure, and dynamics of the highly dynamic *C. elegans* tubulin to the canonical *B. taurus* tubulin. I find that the lateral contact loops important for polymerization, owing to sequence divergence, are more ordered in the *C. elegans* tubulin in solution. The pre-ordering of lateral contact loops reduces the entropic cost for incorporating the heterodimer in the MT lattice. Subsequent temperature-dependent tubulin reconstitution assays and CryoEM imaging show that, indeed, the lateral loops in *C. elegans* tubulin are ordered, and have a higher free energy in solution, compared to the bovine counterpart (*Chaaban et al.*, 2018).
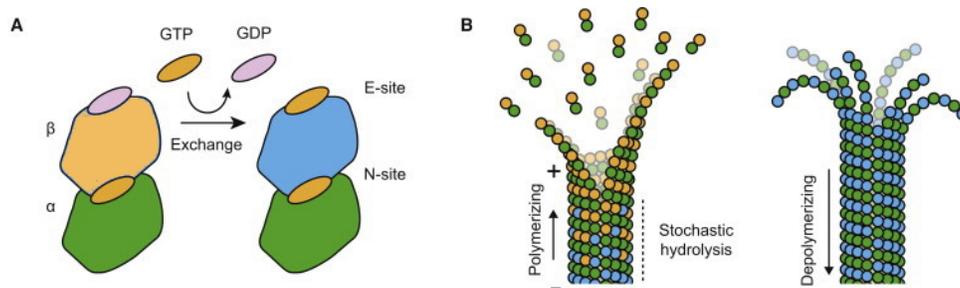
Figure 1.6: Dynamic instability of microtubules. (**A**) Tubulin dimers exchange bound guanosine diphosphate (GDP) for guanosine triphosphate (GTP) at the (exchangeable) E-site, rendering them competent for polymerization. The non-exchangeable N-site buried at the dimer interface remains bound to GTP. (**B**) Microtubules undergoing stochastic growth (polymerizing) and catastrophe (depolymerizing) events. Figure adapted from (*Alushin et al.*, 2014).

1.0.3 Molecular dynamics simulations as a "computational nanoscope"

With the exception of Chapter II (online interactive analysis of protein sequence, structure, and dynamics), I have mentioned the use of molecular dynamics simulations in studying the effects of mutations, sequence variations, and posttranslational modifications on protein function. But what is molecular dynamics (MD)? Broadly, it is a useful computational technique for studying the dynamic evolution of a sys-

tem (Figure 1.7). A molecular mechanics force field is used to calculate the forces between interacting atoms. Then, integration of Newton's laws of motions generates the new configurations of all atoms in this evolving system, providing trajectories that specify their positions and velocities over time. From these MD trajectories, a variety of properties can be calculated, including conformational dynamics, free energy, ligand-binding or other kinetic measures, and is routinely used in protein modeling, materials science, and drug discovery (*De Vivo et al.*, 2016; **?**). In this dissertation, I have made extensive use of MD simulations and application of methods for analyzing MD trajectories in improving our understanding of dynamics, and hence, function and regulation of cytoskeletal proteins.

A Divide time in discrete time steps (Δt)
For integrating equations of motion

B At each time step calculate pair-wise atomic forces (*F(t)*)
By evaluating **force-field** gradient
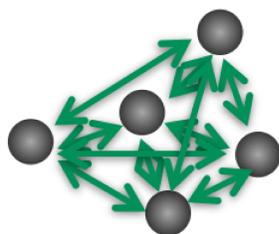
Empirical force field
$$E(\vec{R}) = \sum_{\text{bonded}} E_i(\vec{R}) + \sum_{\text{non-bonded}} E_i(\vec{R})$$

Atomic motion described classically
$$m_i \frac{d^2}{dt^2}\vec{R}_i = -\vec{\nabla}_i E(\vec{R})$$

C Use the forces to calculate **velocities** and move atoms to new **positions**
By integrating numerically
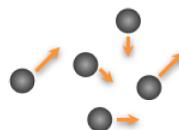
Figure 1.7: Basic anatomy of molecular dynamics simulations. (**A**,**B**) Inter-atomic forces are calculated at each discrete time-step, for example every two femto (10^{-15}) seconds. (**C**) Forces from step (B) are used to propagate the trajectory of atoms. Steps (B) and (C) are repeated many, many times (for example, for about 10^8 times for a simulation of around 200 nano (10^{-9}) seconds).

CHAPTER II

Analysis of Protein Sequence, Structure, and Dynamics

Contents of this chapter were partially published in (*Jariwala et al.*, 2017). Citation: Jariwala, S., L. Skjærven, X.-Q. Yao, and B. J. Grant (2017), Investigating Protein Sequence-structure-dynamics Relationships with Bio3D-web, *Journal of Visualized Experiments*, doi: 10.3791/55640.

2.1 Introduction

With the recent dramatic increase in the amount of high-resolution strucural data, many structures are now available for the same protein familiy determined under different crystallization conditions and oligomeraization states. Detailed comparison of these structures can inform on dynamics and mechanisms critical for biological processes including ligand binding, enzyme catalysis, and allosteric regulation (*Boehr et al.*, 2009; *Henzler-Wildman and Kern*, 2007). A wide range of online servers and molecular visualization tools enable researchers to explore and analyze individual biomolecular structures. However, existing tools for analysis of the sequence, structure and dynamics of large heterogeneous protein families often require significant computational expertise and typically remain accessible only to users with relevant

programming skills. For example, the Bio3D package requires R (*Skjærven et al.*, 2014), ProDy requires python, and Maven requires Matlab knowledge (*Bakan et al.*, 2011; *Zimmermann et al.*, 2011). To this end, we have developed Bio3d-web, an online application that implements an interactive workflow for analysis of protein sequence, structure, and dynamics.

Bio3D-web requires no programming knowledge and thus increases the accessibility and decreases the entry barrier to performing advanced comparative sequence, structure and dynamics analyses. Furthermore, the preparation, curation, annotation and clean-up of molecular structures that is frequently necessary for efficient analysis is included with the Bio3D-web application. Additionally, the restriction to performing such analysis on capable computational resources is alleviated by our server instance that enables large-scale analysis of many structures that can be initiated and controlled from any modern web browser.

Bio3D-web provides a variety of functions to support the five major steps of data analysis shown in Figure 2.1 and discussed in detail in the **Application Overview** section. These steps constitute a workflow that spans from query sequence or structure input, through multiple levels of sequence-structure-dynamic analysis, to summary report generation. Results are available immediately through extensive in-browser visualization and plotting devices, as well as through downloading result files in commonly used formats. In addition to a convenient easy to use dynamic interface for exploring the effects of parameter and method choices, Bio3D-web also records the complete user input and subsequent graphical results of a user's session as a sharable reproducible report in PDF, Word and HTML formats. User sessions may be saved and reloaded at future times and complete results downloaded and further interpreted by the Bio3D R package on a user's local machine.

2.2   Application Overview

A typical Bio3D-web session proceeds through five consecutive and dependent steps (see Figure 2.1 for a schematic representation). Each step is implemented as a consecutive navigation tab of the web application namely **SEARCH**, **ALIGN**, **FIT**, **PCA**, and **eNMA**.

1. Structure search and selection (SEARCH)

This tab enables the identification and selection of structures in the Protein Data Bank (PDB) format (a standard format for representing molecular data from structure-determining experients) related to a user input PDB code or protein sequence. Once a user enters their four character long PDB code, or pastes a protein sequence, the server instantaneously searches the RCSB PDB database (*Berman et al.*, 2002) for sequence related structures. The user may also upload a structure in PDB format, for example a recently crystallized protein previously unlisted in the PDB database, to begin analysis. Identified structures are presented in rank order of decreasing sequence similarity to the query with a subset of the most related structures selected for further analysis by default. Further refinement of this subset (via user initiated inclusion/exclusion of structures) is facilitated through extensive annotations presented in table format together with a schematic plot and linked interactive sliders that alter the inclusion thresholds of sequence similarity. Selected structures from this table are clearly highlighted in both the table and schematic plot. It is important to note that only these structures will be subject to ensemble analysis of their sequence, structure and conformational relationships upon visiting (i.e. clicking on) subsequent tabs.
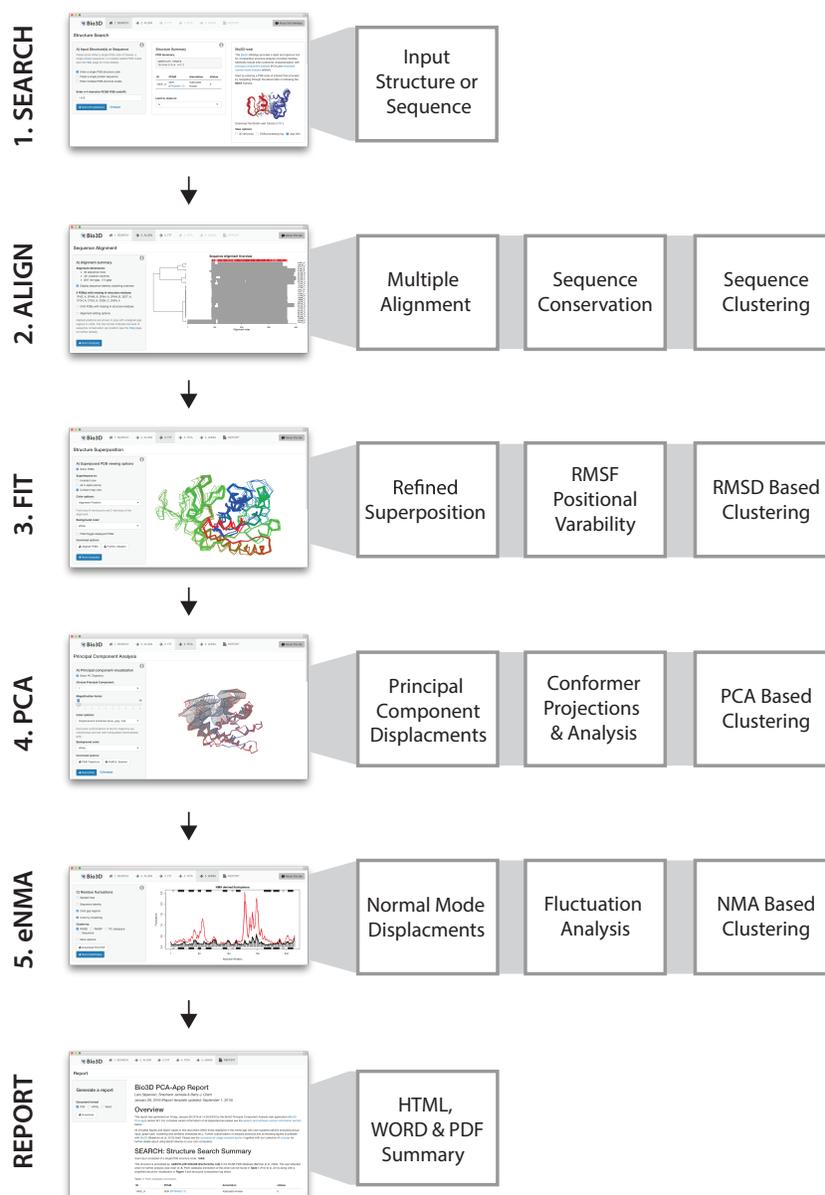
Figure 2.1: Bio3D-web overview. Bio3D-web takes a user provided protein structure or sequence as input in the SEARCH tab (**1**). The server provides a list of related structures, which can be selected for further analysis. (**2**) The ALIGN tab provides sequence alignment and analysis of the structures selected in the SEARCH tab. (**3**) In the FIT tab all structures are superimposed and visualized in 3D together with the results of conventional pair-wise structure analysis. (**4**) Principal component analysis of the structure set is performed in the PCA tab to characterize inter-conformer relationships. (**5**) Normal mode analysis on each structure can be carried out in the eNMA tab to explore dynamic trends for the available structural states.

16

2. Multiple sequence alignment analysis (ALIGN)

In this tab, all previously selected structures are subject to multiple sequence alignment and initial sequence similarity and sequence conservation analysis. Upon entering this tab, all selected structures from the SEARCH tab are parsed and their sequences aligned. The sequence alignment is available as a schematic representation as well as a more traditional in-browser alignment visualization. A FASTA format alignment file can also be downloaded for optional alignment visualization and editing in an external sequence viewer. Edited alignments can be optionally uploaded if required.

This tab further facilitates clustering of the selected PDBs based on the pair-wise sequence alignment which can be visualized either as a dendrogram or a heatmap. Positional sequence conservation within the structure set can further be plotted using various scoring functions including entropy, similarity and identity (*Valdar*, 2002). Conservation in the context of the associated PFAM seed alignment can also be optionally displayed (*Finn et al.*, 2014).

3. Structure fitting and analysis (FIT)

In this tab the aligned structures are superimposed on their invariant structural core and visualized in-browser via an OpenGL-Bio3D based molecular viewer with multiple coloring options. These include coloring by alignment position, structural variability per position, RMSD cluster groups, sequence cluster groups, aligned regions and secondary structure. The superposed structures can be downloaded as either conventional PDB files or as a single PyMOL session file. Analysis of pair-wise structural deviations (RMSD), fluctuations (RMSF) and multiple structure visualization is also provided along with RMSD clustering heatmaps, dendrograms and histograms.

4. Principal component analysis (PCA)

In this tab PCA is performed on the coordinates of all superimposed structures to characterize inter-conformer relationships (*Grant et al.*, 2006). This analysis effectively captures and summarizes the main conformational features and structural displacements in the structural ensemble. Visualization of the PCs is provided in-browser via an OpenGL-Bio3D based molecular viewer as well as via a downloadable PDB format trajectory file and a PyMOL session (vector field). Projections of the superimposed structures onto user selected PCs provides a low-dimensional representation of the relationship between all structures in terms of the conformational variability captured by the chosen PCs. This so-called 'conformer plot' can be colored by sequence, RMSD, and PC based clustering results and represents a key result from Bio3D-web. This tab also provides a range of PC based clustering options.

5. Ensemble normal mode analysis (eNMA)

In this tab NMA across the structure set is performed to predict large-scale motions (*Skjærven et al.*, 2014). Here NMA is performed on selected structures in a way that facilities the interoperation of structural similarity and dissimilarity trends. This allows a user to explore dynamic trends of all crystalized states in relation to each other without the conventional caveat of potentially over-interpreting the differences between extreme cases or a single artifactual structure. This tab also provides clustering of the structures based on the calculated normal modes and fluctuation profiles. In our experience, applying NMA across multiple structures, and then carefully contrasting the results, can provide new information on state specific global and local dynamics of potential functional relevance. In the examples below we demonstrate how this eNMA approach can predict ligand dependent modes of motion, distinct local flexibilities in functional regions, and yield a root mean square inner product (RMSIP) based clustering that clearly distinguishes functionally distinct structural

states.

2.3 Methods

Bio3D-web is powered by the Bio3D R package for analysis of biomolecular structure, sequence and molecular simulation data (*Grant et al.*, 2006; *Skjærven et al.*, 2014). In particular, Bio3D algorithms for rigid-core identification, superposition, principal component analysis (PCA) , and ensemble normal mode analysis (eNMA) form the basis of the application. We also utilize Bio3D protocols that depend on pHMMER (*Eddy*, 2011) for the identification of related protein structures, and MUS-CLE (*Edgar*, 2004) for sequence alignment. Structure and sequence annotations are derived via Bio3D utilities from the RCSB PDB (*Berman et al.*, 2002) and PFAM databases (*Finn et al.*, 2014).

Bio3D-web employs Shiny's reactive programming and web application framework (*RStudio*, 2013). Shiny provides the JavaScript web widgets together with the CSS/HTML Bootstrap web framework. We make extensive use of Shiny's reactivity feature that binds web controls to R functionality in a manner that caches expensive computational steps so that they are not re-computed unnecessarily during an interactive browser session.

Bio3D-web can be run from our online server or installed locally on any computer running R. Our online server is accessible at `http://bio3d.ucsd.edu/pca-app`. Bio3D-web is open to all users and is provided free of charge under a GPL open-source license from: `http://thegrantlab.org/bio3d/webapps`. The source-code and information on pull requests can be viewed at our Git page at: `https://bitbucket.org/Grantlab/bio3d`.

2.4 Example Applications

Here we present example Bio3D-web analysis of the Adenylate kinase (Adk) and GroEL chaperonin families. Summary report detailing complete results for Adk as an exmple is provided in Appendix A.

2.4.1 Adenylate Kinase

Adenylate kinase (Adk) is a ubiquitous enzyme that functions to maintain the equilibrium between cytoplasmic nucleotides essential for many cellular processes. Adk operates by catalyzing the reversible transfer of a phosphoryl group from adenosine triphosphate (ATP) to adenosine monophosphate (AMP). This reaction is accompanied by a well-studied rate limiting conformational transition of regions that close over the two nucleotide-binding sites (*Henzler-Wildman et al.*, 2007; *Kerns et al.*, 2015). Here we analyze available Adk structures with Bio3D-web to reveal features of this transition.

We can begin our Bio3D-web analysis of Adk by entering the RCSB PDB code of any known Adk structure. For example, entering the PDB ID 1AKE in panel A of the SEARCH tab returns 167 sequence similar structures from which the top 26 are automatically selected for further analysis (see panel B). The annotation presented in panel C indicates that these selected structures are all from *E. Coli*; were solved by x-ray diffraction in a range of space groups; have a resolution range of 1.63 to 2.8Å, were co-crystalized with a range of different ligands (including no ligands, AMP, ADP, MG and the inhibitor AP5). Note that additional annotation details can be displayed by clicking on "Show/Hide Columns" option in panel C. Multiple sequence alignment is performed on the ALIGN tab, followed by structure superposition on the FIT tab. The superimposed structures, displayed interactively in panel A, indicate the presence of a relatively rigid core region (encompassing residues 1–29, 68–117, and 161–214; see the 'optional core and RMSD details' panel at the bottom of the FIT tab for

details). Two more variable nucleotide-binding regions (residues 30–67 and 118–167) are also clearly visible (Figure 2.2). RMSD-based clustering groups these structures into two distinct conformations.
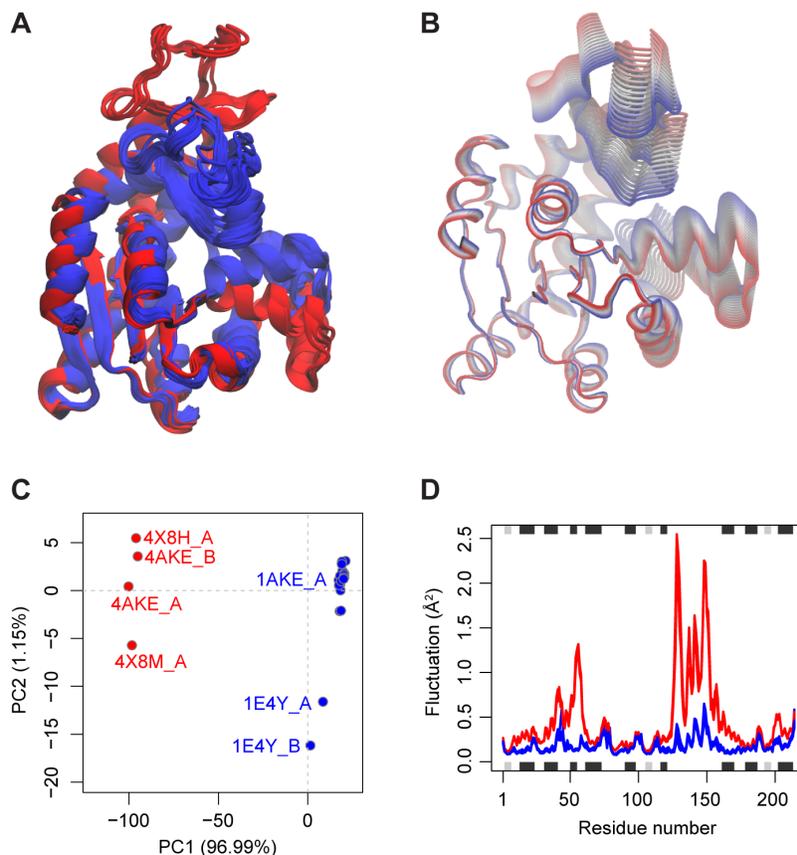
Figure 2.2: Results of Bio3D-web analysis of Adenylate kinase. (**A**) Available PDB structures of Adenylate kinase superimposed on the identified invariant core. Structures are colored according to RMSD-based clustering provided in the FIT tab. (**B**) Visualization of the principal components is available from the PCA tab to characterize the major conformational variations in the data set. Here, the trajectory corresponding to the first principal component is shown in tube representation showing the large-scale closing motion of the protein. (**C**) Structures are projected onto their two first principal components in a conformer plot showing a low-dimensional representation of the conformational variability. Each dot (or structure) is colored according to user specified criteria, in this case PCA-based clustering results. (**D**) Normal mode analysis in the eNMA tab suggests enhanced local and global dynamics for structures in the open state (red) in comparison to the closed form (blue) structures.

Clicking on the PCA tab more clearly shows the relationship between the struc-

tures in terms of the displacements of these regions that effectively close over the bound nucleotide species in related structures (Figure 2.2B and C). The majority of structures are in the 'closed' form (blue in Figure 2.2C) and are associated with a bound ligand or inhibitor. In contrast more 'open' conformations are nucleotide and inhibitor free. This is consistent with the extensive body of research on Adk structure and dynamics indicating that an open configuration of these regions is required for nucleotide binding and a closed conformation for efficient phosphoryl transfer and suppression of unproductive hydrolysis events (*Hanson et al.*, 2007). It is notable that a single PC captures 97% of the total mean square displacement in this Adk structure set and provides a clear and compelling description of the open to closed transition along with the individual residue contributions to this functional displacement (panel C of the web-app and Figure 2.2C).

Visiting the NMA tab and increasing the number of structures considered for calculation (via decreasing the cutoff for filtering similar structures) indicates that open state structures display enhanced local and global dynamics in comparison to the closed form structures (Figure 2.2D and panel C of web-app). Clustering based on RMSIP values again displays a consistent partitioning of open and closed state structures (panel D of web-app). Comparing PCA and NMA results for individual structures (panel E of web-app) indicates that the first mode of all open form structures displays a relatively high overlap to PC1 (with a mean value of 0.37 ± 0.04). In contrast closed form structures display lower values (with a mean of 0.30 ± 0.01). RMSIP values for open form structures (0.62 ± 0.003) are also higher than those of closed structures (0.56 ± 0.008).

Collectively these results indicate the existence of two major distinct conformational states for Adk crystal structures. These differ by a collective displacement of two nucleotide-binding site regions, and show distinct flexibilities upon nucleotide binding. The conformational transition between these states is well captured by a

single low frequency normal mode of open state structures.

2.4.2 GroEL

The chaperonin GroEL is a molecular chaperone that aids in the folding of a wide range of essential proteins in *Escherichia coli* (*Hartl et al.*, 2011). GroEL assembles into a barrel-shaped complex composed of two heptameric rings. Each of these rings provides a protective folding chamber for nascent polypeptides. Extensive structural and functional characterization has revealed that the GroEL-stimulated folding process is driven by large-scale conformational changes in GroEL and a set of ATP-dependent allosteric transitions (*Skjærven et al.*, 2015). Here we utilize Bio3D-web to characterize the conformational heterogeneity and flexibility of all available GroEL subunit structures.

Providing a GroEL RCSB PDB code (PDB ID 1AON) as input on the SEARCH tab results in the listing of 574 highly similar subunit structures (i.e. annotated hits above the default sequence similarity threshold). The provided annotation table in panel C of the SEARCH tab shows that these hits all correspond to *E. coli* GroEL structures. The more than 600 additional hits below the inclusion cutoff correspond to other bacterial, mammalian and archaeal chaperonin variants. The schematic overview of the sequence alignment along with sequence identity clustering in the next ALIGN tab shows that the selected structures are conserved over all 524 alignment positions.

All selected 574 subunit structures are superimposed in the FIT tab and interactively displayed in panel A colored by alignment position. The structures are superimposed on the identified invariant core located in the equatorial domain (Figure 2.3A). Toggling between 'core' and 'Calpha atom' fitting options demonstrates how core-fitting greatly enhances the visualization and subsequent interpretation of displacements over conventional superposition. In particular, relative domain displacements

are more clearly evident from core-fitting.
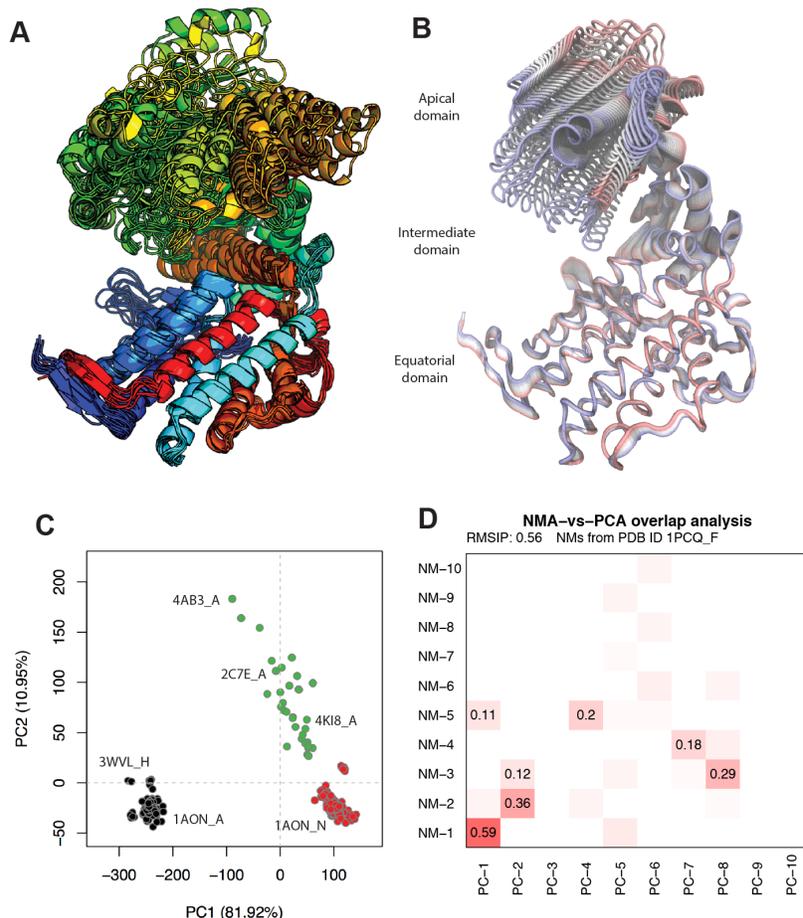
Figure 2.3: Results of Bio3D-web analysis of GroEL. (**A**) Representative PDB structures of the GroEL subunit superimposed on their identified invariant core. Structures are colored based on alignment position. (**B**) Visualization of the first principal component display a rotation and tilt of the apical domain in the PCA. (**C**) Low-dimensional conformer plot reveals the presence of three conformational forms with the closed (red) and open (black) being the most abundant. (**D**) Comparison of the normal modes of PDB ID 1PCQ with the principal components from the PCA-tab. The result is represented as a heatmap with coloring from white (low similarity) to red (high similarity).

RMSD-based clustering of the structures shows that they can be divided into two major groups corresponding to the main functional end-states of GroEL: (1) the substrate receptive state in which the subunits obtain a closed structure, and (2) the ADP- and GroES-bound, folding active state in which the subunit is in an open

24

conformation. ATP-bound states captured by electron microscopy define a third conformation with the apical domains rotated ~20° clockwise as compared to the closed state. In panel B a plot of the variability per position (in terms of RMSF of C-alpha atoms) provide information on the regions of structure that deviate the most. The plot reveals that the apical domains (residues 192-373) show large conformational variability. Further insight into these apparent displacements is provided on the next PCA tab.

PCA of the superimposed structure ensemble shows that the largest collective displacements correspond to a rotation and tilt of the apical domain coupled to a downward rotation of the intermediate domain (panel A of the app and Figure 2.3B). The low-dimensional conformer plot of PC1 vs PC2 projections in panel B clearly shows the relationship between structures in terms of these collective coordinates (panel B of the web-app and Figure 2.3C). Three major conformational groups are apparent. These correspond to apo closed forms (red), open ADP-bound forms (black), and ATP-bound states captured by electron microscopy (green). This demonstrates how we can effectively rationalize the heterogeneity in large structural sets such as this one in a manner that reveals functionally and mechanistically important inter-conformer relationships.

NMA of representative structures in the eNMA tab shows enhanced fluctuations at regions corresponding to the apical domain for the open conformation as compared to the more closed conformations (panel C of the web-app). The first normal mode of the subunits in the open conformation corresponds to a rotation and tilt of the apical domain (panel B of the web-app). Comparison of the normal modes with the principal components reveals that this first normal mode of the open subunit displays a high similarity with the conformational change described by the first principal component with an overlap value of 0.59 (see panel E of the web-app and Figure 2.3D). Thus, the most energetically favorable deformations of the GroEL subunit lie in the path

of the open-to-closed conformational states.

2.5 Conclusion

In this chapter, I presented Bio3D-web, an online application for analyzing biomolecular structure data. By reducing the required technical expertise, Bio3D-web facilitates structural bioinformatics analysis for a broader range of researchers. Bio3D-web is now being used actively by students, collaborators, and community of Bio3D users, who report that they are able to perform their analysis more quickly and efficiently. I encourage readers to take a look at the two publications related to this work for more details (*Jariwala et al.*, 2017; *Skjærven et al.*, 2016).

The design of Bio3D-web currently emphasizes simplicity over exhaustive inclusion of the many analysis methods available in the full Bio3D package. In many cases it is envisaged that researchers will use Bio3D-web to understand general trends in their protein family or superfamily of interest, which may then inform more specialized analyses. Bio3D-web is therefore designed to quickly explore biomolecular structure datasets and to act as a hypothesis-generating tool. I encourage users to further explore their data by making use of the example Bio3D code in the reproducible report that also stores all query details and analysis results. This shareable sharable summary reports also a link to revisit analysis sessions.

Bio3D-web is freely available and may be installed locally on any computer running R, or hosted on a remote server, to provide a customized multi-user instance with access to priority structural datasets such as those common in the pharmaceutical industry.

CHAPTER III

A Posttranslational Modification Of The Mitotic Kinesin Eg5

Contents of this chapter were partially published in (*Muretta et al.*, 2018). Citation: Muretta, J. M., et al. (2018), A posttranslational modification of the mitotic kinesin Eg5 that enhances its mechanochemical coupling and alters its mitotic function, *Proceedings of the National Academy of Sciences*, *115*(8), E1779–E1788, doi: 10.1073/pnas.1718290115. This has been a collaborative effort. The computational work was done by myself, with metadynamics simulations performed by G.S. at the University of Michigan. The experimental work was carried out by our collaborators at University of Minnesota (J.M.M. performed transient kinetics and conformational coupling experiments), University of California Irvine (B.J.N.R in S.P.G. lab performed force-velocity experiments), University of Vermont, Cleaveland Clinic Foundation, Ohio State University, and Mayo Clinic.

3.1 Abstract

Numerous posttranslational modifications have been described in kinesins, but their consequences on motor mechanics are largely unknown. We investigated one of these modification—acetylation of lysine 146 in kinesin 5 motor Eg5—by creating an

acetylation mimetic lysine to glutamine substitution (K146Q). Lysine 146 is located in the $\alpha 2$ helix of the Eg5 motor domain, where it makes an ionic bond with aspartate 91 on the neighboring $\alpha 1$ helix. Our molecular dynamics simulations predict that disrupting this bond enhances allosteric coupling between the catalytic site and the neck linker. Our collaborators tested this using time-resolved structural kinetics and single molecule force spectroscopy and found that the K146Q mutation increases motor performance under load and coupling of the neck linker to catalytic site. These changes convert Eg5 from a motor that dissociates from the microtubule at low load into one that is more tightly coupled and dissociation resistant—features shared by the transport motor kinesin 1. These features combined with the increased propensity to stall predict that the K146Q Eg5 acetylation mimetic should act in the cell as a "brake" that slows spindle pole separation. Our collaborators have confirmed this "braking" effect by expressing this modified motor in mitotically active cells. Thus, our results illustrate how a posttranslational modification of a kinesin can be used to fine tune motor behavior to meet specific physiological needs.

3.2 Introduction

Members of the kinesin superfamily of molecular motors fulfill specific roles in cell physiology. Some serve as intracellular transporters, others serve as regulators of microtubule (MT) dynamics, and some as mitotic motors (*Hirokawa and Takemura*, 2004; *Vicente and Wordeman*, 2015; *Bachmann and Straube*, 2015; *Walczak et al.*, 2013). However, not all molecular motors are limited in their roles. For example, while kinesin 1 transports organelles, it also slides MTs during axonal elongation (*Jolly et al.*, 2010). MCAK, an MT- depolymerizing mitotic kinesin, can also drive cell motility (*Braun et al.*, 2014). MTs also play multiple roles in cell physiology, and this adaptability is due in part to multiple posttranslational modifications (PTMs) (*Yu et al.*, 2015; *Janke*, 2014; *Sirajuddin et al.*, 2014). Thus, some of the multifaceted

functionality of kinesin motors might reflect PTMs that modify their mechanochemical behavior. PTMs have been identified in most kinesins, including many in the catalytic domain of these motors (*Hornbeck et al.*, 2015; *Liu et al.*, 2014). However, in nearly all cases, the consequences of these PTMs on motor function remain unknown. Multiple PTMs in the vicinity of the catalytic site have been reported in the motor domain of the mitotic kinesin Eg5, including in Loop 5 and helices $\alpha 2$ and $\alpha 3$ (*Bickel et al.*, 2017). As with other kinesins, Eg5 also plays multiple roles in cell physiology, including formation of the mitotic spindle in dividing cells, axonal branching, and cell motility (*Falnikar et al.*, 2011; *Venere et al.*, 2015). These multifunctional roles may be associated with distinct mechanochemical requirements, which might be achieved via specific PTMs. One such PTM of Eg5 is acetylation of a lysine residue (K146) at the C-terminal end of helix $\alpha 2$ (Figure 3.1A) (*Choudhary et al.*, 2009; *Sol et al.*, 2012; *Nalawansha et al.*, 2017). Lysine 146 makes a salt bridge to aspartate 91 on helix $\alpha 1$, and its acetylation disrupts this ionic interaction. This effect can be mimicked by a lysine to glutamine point mutation (*de Boor et al.*, 2015; *Gorsky et al.*, 2016; *Cohen et al.*, 2011). The importance of acetylation PTM in cell biology is underscored by the recent findings that it regulates not only chromatin activity through its effects on histones (*Marmorstein and Zhou*, 2014) but also transcription, metabolism, autophagy, and viral infection (*Menzies et al.*, 2016; *Kouzarides*, 2000; *Blee et al.*, 2015; *Xie et al.*, 2015).

Two key structural domains in kinesin 1 and Eg5, switch 1 (Sw1) and the neck linker (NL), alternate between two orientations during the ATPase cycle (*Muretta et al.*, 2015; *Rice et al.*, 1999a; *Gigant et al.*, 2013; *Cao et al.*, 2014; *Sindelar and Downing*, 2007). Sw1 senses the γ-phosphate of ATP, and it assumes "open" (capable of binding ATP) and "closed" (capable of hydrolyzing ATP) conformations. The NL moves in response to ATP binding between "undocked" (not force-generating) and "docked" (force-generating) orientations. The role of additional structure and

sequence features such as cover strand (which forms the cover-neck bundle along with NL) in force production in kinesin 1 (Kif5C) and kinesin 3 (KIF1A) has been discussed in Chapters IV and V, respectively. Here, we focus our work on the coupling between Sw1 and NL for force production.

Tight coupling is important for a highly processive motor that operates in isolation, such as kinesin 1, as it helps ensure that the motor does not enter a weak MT binding state before it has docked its NL and generated force. Conversely, this might be unnecessary for Eg5 which, by working in large ensembles, may not need to be highly processive. Our collaborators have previously found that while Sw1 and the NL are tightly coupled in kinesin 1 (e.g., closed Sw1 associated with docked NL), they are more loosely coupled in Eg5 (*Muretta et al.*, 2015). In this chapter, I describe our study of acetylation of K146 using molecular dynamics (MD) and metadynamics, and discuss our collaborators' results on transient time-resolved kinetics, single-molecule mechanics, and time-lapse microscopy of cells in mitosis. We find that a acetylation-mimetic mutant of Eg5 (K146Q) shows much tighter conformational coupling of Sw1 to the NL, and this is associated with improved motor performance under load: features characteristic of kinesin 1. Taken together, our results suggest that kinesin PTMs can act as chemomechanical regulators, broadening a specific kinesin's capabilities and enhancing the flexibility that a cell has to respond to a wide variety of physiological demands.

3.3   Results

3.3.1   Probing global and local effects of acetylation with MD simulations

ATP binding to Eg5, akin to all kinesin motors studied till date, produces a comformational change in the catalytic site. The acetylation mimetic mutation K146Q has little effect on this process, as measured through steady-state kinetics by our

collaborators (*Muretta et al.*, 2018). This local conformation change brought about by ATP binding causes allosteric effects transmitted through the motor domain that alter MT binding and NL orientation (*Atherton et al.*, 2014; *Goulet et al.*, 2014). We, therefore, performed MD simulations to characterize the dynamical effects of K146Ac and K146Q mutations on the motor domain. We performed 400ns of MD simulations each for WT, K146Ac, and K146Q Eg5 bound to tubulin heterodimer. Analysis across replicates was used to predict statistically significant differences in residue-pair interactions and energetics (see *Methods*). Althought the effects of K146Ac can be characterized computationally, the results that follow focus on the mimetic mutation K146Q to complement experiments where it is not yet possible to selectively acetylate lysine 146 residue. Significant changes in residue-residue interactions associated with the K146Q mutation are evident for two functionally important domains. The first involves structures that are responsible for generating the "power stroke", and they include the neck linker (NL), the cover-strand (CS), Loop 13, and $\beta 7$ (Figure 3.1A and B). The K146Q mutation shortens distances and enhances interactions between the CS and the NL, the CS and Loop 13, and the NL and $\beta 7$ (Figure 3.1B). Other charged residue interactions, including those between E14 (CS) and E92 ($\alpha 1$), the N-terminal portion of $\beta 1$ with the neighboring Loop 13/$\beta 8$, and E20 ($\beta 1$) to R329 (Loop13) are also enhanced (See Table B.1 in Appendix B). In aggregate, these explain why we find that the NL spends a significantly greater proportion of time in the docked orientation in the mutant compared to wild type (72% of simulation frames versus 40%, Figure 3.2A)) and they are a consequence of the small displacements of $\alpha 1$ that result from removing the salt bridge between D91 ($\alpha 1$) and K146 ($\alpha 2b$). Thus, while the average distance between K146 to D91 is $3.6 \pm 1.5 \, \text{Å}$ in WT Eg5, the corresponding Q146 to D91 distance in the K146Q mutant is $4.9 \pm 1.2 \, \text{Å}$, with a corresponding 5.3 kcal/mol decrease in interaction energy (Table 3.1 below and Table B.1 in Appendix B).
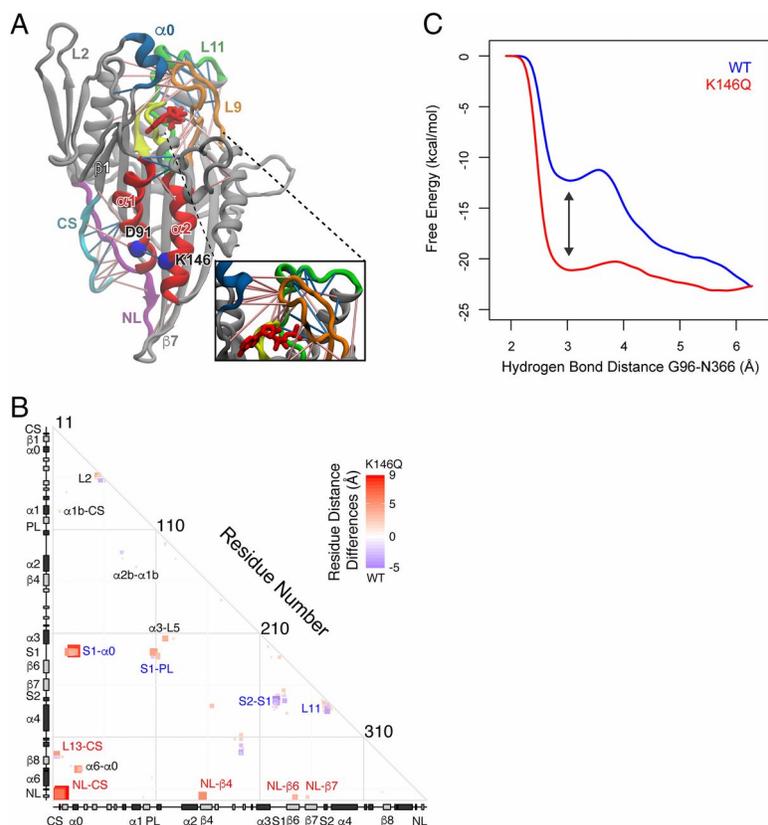
Figure 3.1: Dynamical effects of K146Q acetylation mimetic mutation. (**A**) Structure of Eg5 motor domain highlighting major structural elements, including Loop 9/Sw1 (orange), Loop 11/Sw2 (green), P loop (yellow), helix $\alpha 0$ (blue), NL (purple), CS (cyan), and helices $\alpha 1$ and $\alpha 2$ (red). Also shown is the WT salt bridge residue pair D91 and K146 (blue spheres). Significant ($P < 10^{-5}$) residue–residue distance changes between WT and K146Q mutant simulations are displayed as red and blue lines, with color intensity representing the magnitude of change. (*Inset*) Enlarged view of the catalytic site. (**B**) Pairwise residue difference distance analysis of WT and K146Q simulations. Significant ($P < 10^{-5}$) residue–residue distance changes are displayed with size and color intensity scaled by magnitude (red for shorter in K146Q and blue for shorter in the WT). Major secondary structure elements are displayed in the margins (α-helices in black and β-strands in gray). Specific structural regions noted in the text with distinct interactions with the NL and Sw1 regions have been labeled in red and blue, respectively. (**C**) Free energy profile from well-tempered metadynamics simulations for NL docking/undocking of WT (blue) and K146Q mutant (red) as a function of the distances between residues N366 (NL) and G96 (helix $\alpha 1$). The arrow denotes the energy difference between the WT and K146Q at point of hydrogen bond formation. Reprinted from (*Muretta et al.*, 2018) under PNAS license.

The second domain includes structures that bind and coordinate nucleotide, including Loop 9/Sw1, the P loop, β1/helix α0, and Loop 14 (Figure 3.1A). The largest change is seen for Loop 9/Sw1. The closed conformation for this loop, which is necessary for ATP hydrolysis, is more frequently seen in the mutant (75% of simulation time versus 18%, respectively), and this leads to shorter average distances to the P loop illustrated by a decrease in the G108 (P loop) to N229 (Sw1) distance of 3 Å. Distances between Sw1 and Sw2 and between Sw1 and Loop 14/α6 also decrease in the mutant (Figure 3.1B, blue labels and Appendix B, Table B.1). The K146Q mutation also causes the helix α0 to move away from Loop 2a and toward Sw1, Loop 14, and α6 (Figure 3.2B). Other inter-residue distances, including W127 (Loop 5) to E215 (α3), D118 (Loop 5) to bound nucleotide, and D186 (Loop 8) to R312 (Loop 12), are also significantly shorter in the mutant (Appendix B, Table B.1). Additional simulations of acetyl lysine at position 146 revealed equivalent enhanced catalytic site closure, with the Sw1 region exhibiting a closed conformation for 70% of simulation time versus 18% for the WT, and less dramatic but still significant ($P < 10^{-5}$) α1, α2b, and NL distance differences leading to enhanced docking of the NL (71% of simulation time versus 18% in the WT).

An analysis of MD-derived correlated motions reveal that the K146Q mutation produces significantly stronger couplings involving the nucleotide binding site (switch 1/P loop and switch 1/α0) and NL regions (NL/α1 and NL/β7) (Figure 3.3). These structural dynamic changes are likely a direct result of the tighter interactions between these regions in the K146Q acetylation mimetic, as noted above. Overall these results indicate that the K146Q mutation results in dynamic perturbations both locally—reflected in an increase in the α1 to α2 distance—and at more distant functional regions, which appear to collectively enhance coordination of the structural states of the NL with switch 1 (Figure 3.1A). Overall, these results indicate that the K146Q mutation results in dynamic perturbations both locally—reflected in an in-

crease in the $\alpha1$ to $\alpha2$ distance—and at more distant functional regions, collectively leading to enhanced coordination of the structural states of the NL with Sw1 regions.

Table 3.1: Residue-residue interaction energies for WT and K146Q Eg5. Results of residue-wise energy decomposition using molecular mechanics with generalized Born and surface area solvation (MM-GBSA) calculations on wild-type (WT) and mutant (K146Q) simulations. All residue pairs listed, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference larger than 2 kcal/mol between WT and K146Q simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop.

Residue position	ΔG_{WT} (kcal/mol)	ΔG_{K146Q} (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
K11 (CS) – E253 ($\beta7$)	$-0.03\,(0.02)$	$-3.1\,(5.81)$	-3.07
E14 (CS) – R327 (L13)	$-0.07\,(0.04)$	$-5.39\,(5.54)$	-5.32
K15 (CS) – E92 ($\alpha1$)	$0.11\,(0.61)$	$-2.22\,(4.02)$	-2.33
K17 (CS) – E92 ($\alpha1$)	$-1.9\,(3.44)$	$-5.4\,(5.07)$	-3.5
Q20 ($\beta1$) – R329 ($\beta8$)	$-1.01\,(1.68)$	$-3.22\,(2.96)$	-2.21
R24 ($\beta1$) – Q78 ($\alpha1$)	$-3.41\,(2.95)$	$-7.49\,(1.15)$	-4.08
R24 ($\beta1$) – T114 $\alpha12a$	$-3.1\,(1.87)$	$-6.1\,(0.9)$	-3
R26 ($\beta1$) – G108 (P Loop)	$-2.34\,(1.25)$	$-4.48\,(0.69)$	-2.14
N29 ($\alpha0$) – M228 (L9)	$-0.15\,(0.29)$	$-2.3\,(1.41)$	-2.15
D91 ($\alpha1$) – K146 ($\alpha2b$)	$-6.28\,(3.23)$	$-0.94\,(0.91)$	5.34
Q106 (P Loop) – E345 ($\alpha6$)	$-5.56\,(2.35)$	$-3.08\,(0.86)$	2.48
K111 (P Loop) – N229 (L9)	$0.14\,(0.11)$	$2.42\,(0.47)$	2.28
R119 (L5) – E215 ($\alpha3$)	$-1.95\,(3.69)$	$-4.16\,(4.73)$	-2.21
E128 (L5) – Y211 ($\alpha3$)	$-4.05\,(1.99)$	$-1.45\,(0.93)$	2.6
E166 (L7) – N287 (L11)	$-2.2\,(1.73)$	$-4.76\,(3.21)$	-2.56
D186 ($\beta5b$/L8) – R312 (L12)	$-0.81\,(1.18)$	$-7.08\,(6.34)$	-6.27
R192 (L8b) – D322 ($\alpha5$)	$-8.85\,(8.09)$	$-1.19\,(1.74)$	7.66
R221 (L9) – D265 (L11)	$-16.13\,(2.19)$	$-19.34\,(3.28)$	-3.21
R234 (L9) – G268 (L11)	$-1.31\,(1.69)$	$-3.78\,(2.27)$	-2.47
R234 (L9) – E270 (L11)	$-10.58\,(7.6)$	$-1.35\,(0.5)$	9.23
R234 (L9) – E284 (L11)	$-5.09\,(3.91)$	$-10.92\,(6.29)$	-5.83
R274 (L11) – E345 ($\alpha6$)	$-13.26\,(5.82)$	$-7.93\,(6.54)$	5.33
R327 (L13) – E364 (NL)	$-6.16\,(3.17)$	$-9.34\,(4.43)$	-3.18
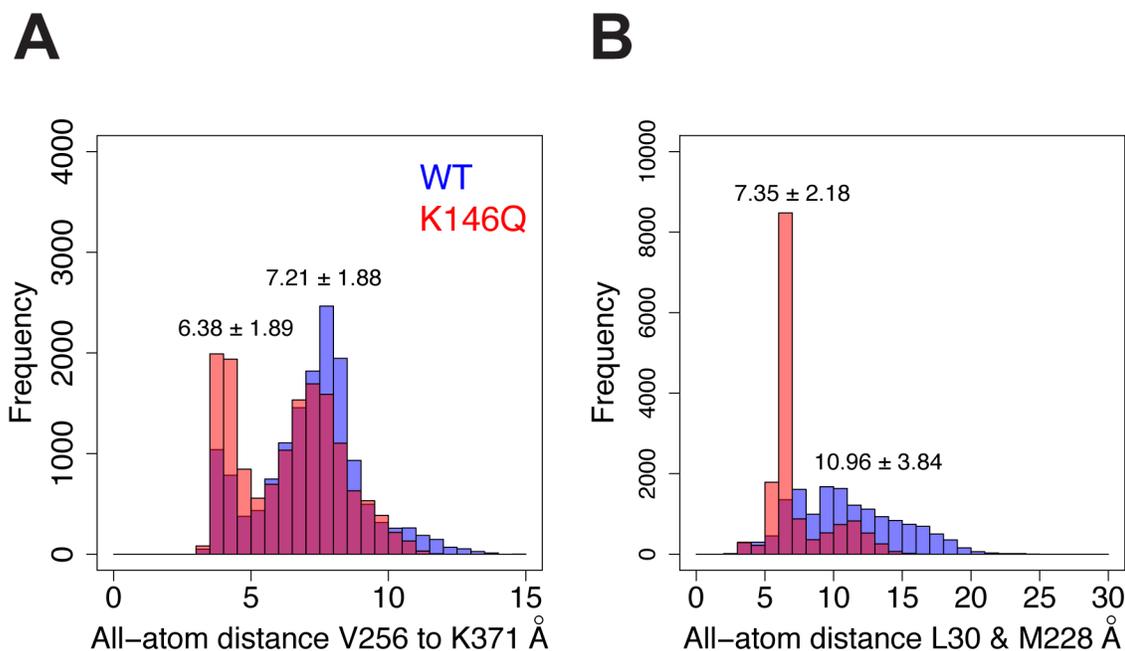
Figure 3.2: Distance distribution histograms. Histograms of distance distributions of (**A**) $\beta 7$ to the neck linker, and (**B**) $\alpha 0$ to SwI, during molecular dynamics simulations. Plotted values and their listed averages were obtained from the last half of all simulation replicates.

3.3.2 Change in free energy of neck linker docking

Analysis of MD simulations of WT and mutant Eg5 revealed enhanced interactions in the NL and Sw1 regions in the mutant motor domain. Next, we wanted to further probe the energetic effects of K146Q mutation on NL docking. Residue G96 at the C-terminal end of helix $\alpha 1$ forms a hydrogen bond with residue N366 in the NL, and this interaction is important for NL docking in kinesin 1 (*Hwang et al.*, 2008). We, therefore, chose the G96–N366 distance as a collective variable for characterizing the free energy of NL docking via 700ns metadynamic simulations (Figure 3.1C). The resulting free energy profiles indicate that the K146Q mutation favors formation of this hydrogen bond compared to the WT, with a relative total system free energy difference of 8.8 kcal/mol for docked versus undocked (Figure 3.1C, arrow).This predicts that NL docking should be energetically more favorable in the K146Q mutant
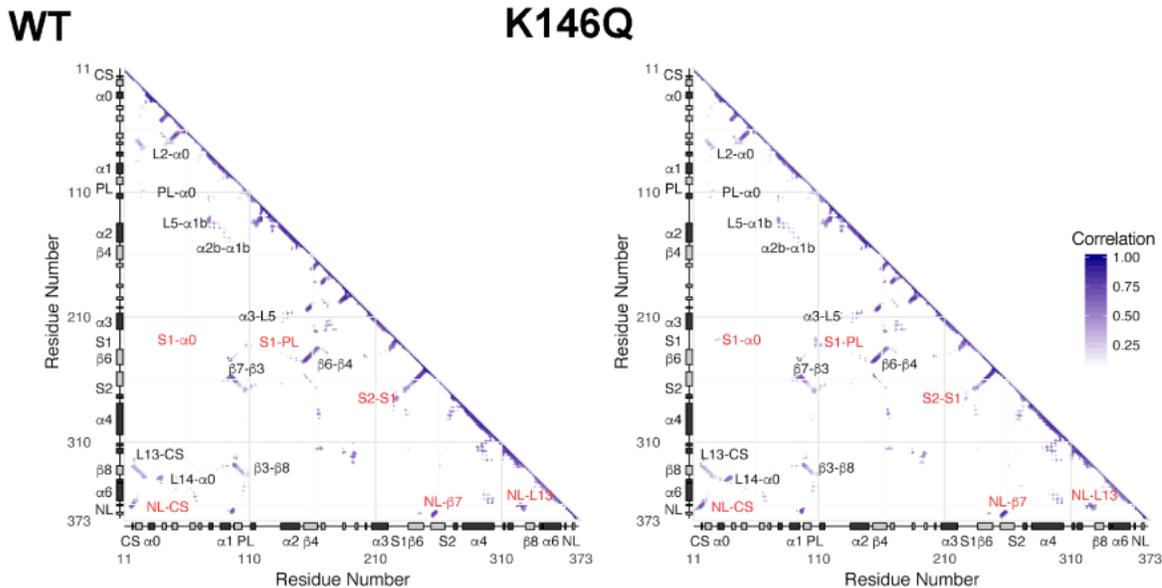
Figure 3.3: Residue-residue plot of correlated motions. The extent of correlation of atomic displacements for all residue pairs during WT (**left**) and K146Q mutant (**right**) simulations. Marginal black and grey rectangles indicate the location of major alpha helix and beta strand secondary structure elements. Key regions of couplings that differ between WT and K146Q are labeled in red.

and further highlights the structural importance of the salt bridge between $\alpha 1$ and $\alpha 2$b for modulating NL docking. Taken together, our simulations imply that a major consequence of K146 acetylation is enhanced conformational coupling between Sw1 and the NL.

3.3.3   Kinetics of neck linker docking and switch 1 closure

In a previous study (*Muretta et al.*, 2015) our collaborators showed how the conformational coupling between Sw1 and NL linker changes during duing the motor mechanochemical ATPase cycle through use of a combination of transient kinetics and time-resolved FRET (TR^{2}FRET). Their findings show that while NL docking and Sw1 closure are tightly coupled in kinesin 1, they are less so in Eg5, explaining one of the many functional differences between these two motors. Our MD simulations suggest that the K146Q mutation enhances the conformational coupling

of the NL with Sw1, similar to that observed in kinesin 1. Our collaborators applied the TR^{2}FRET technique to experimentally test these predictions. Our collaborators first generated two monomeric Eg5 constructs, each containing the K146Q mutation, that have pairs of reactive cysteines for labeling with an FRET donor [*N*-acetylaminoethyl-8-naphthylamine-1-sulfonate (AEDANS)] and acceptor [*N*-(4-dimethylamino-3,5-dinitrophenyl)maleimide (DDPM)]. The first, Eg5$_{NL}$K146Q, has been described above. The second, Eg5$_{Sw1}$K146Q, has reactive cysteines in the Sw1 loop (residue 228) and in a reference point (residue 30) along with the K146Q acetylation mimetic substitution. We mixed donor only (AEDANS)- or donor + acceptor (AEDANS/DDPM)-labeled Eg5$_{NL}$K146Q and Eg5$_{Sw1}$K146Q + MTs with ATP in a stopped flow instrument and then acquired TR^{2}FRET waveforms during subsequent ATP binding, hydrolysis, and MT dissociation. Figure 3.4 illustrates representative waveforms for MT complexes of Eg5$_{NL}$K146Q (Figure 3.4A) and Eg5$_{Sw1}$K146Q (Figure 3.4B) after mixing with 2 mM ATP. The waveforms were analyzed as perviously described (*Muretta et al.*, 2015) to determine the mole fractions of docked versus undocked NL in Eg5$_{NL}$K146Q and closed versus open Sw1 in Eg5$_{Sw1}$K146Q as a function of time after mixing with ATP (Figure 3.4C and D, respectively).

In Eg5$_{NL}$, both ATP binding and subsequent hydrolysis induce NL docking, and we find that the same is true for Eg5$_{NL}$K146Q. However, the K146Q mutation does alter NL movement in two ways. First, it accelerates NL docking during the ATP binding step threefold (Figure 3.4E). Second, it increases the fraction of motors that dock their NL during this step from 50 to 76% (SI Appendix, Table S4). In combination, these effects should make force production more rapid and complete with initial ATP binding. The coordination of Sw1 and NL can be examined by monitoring the coupling ratio defined as the mole fraction of docked NL divided by the mole fraction of closed Sw1 which, in turn, is calculated from the evolution of corresponding TR^{2}FRET waveforms over time after mixing with ATP. A ratio of 1.0 implies

37
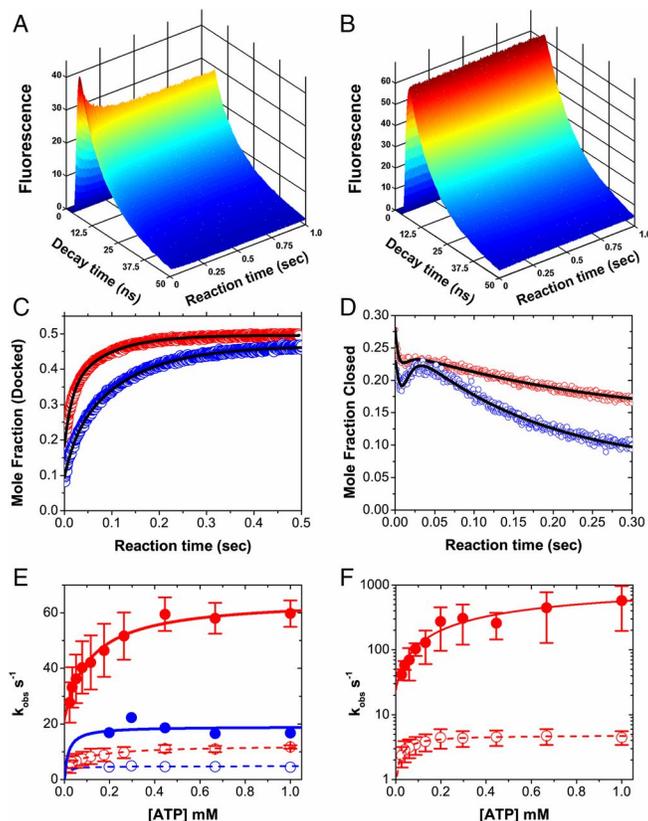
Figure 3.4: TR^{2}FRET of ATP binding to MT-bound Eg5. (*A* and *B*) Representative waveforms after mixing 2 mM ATP with 1μM AEDANS/DDPM-labeled Eg5$_{NL}$K146Q (**A**) or Eg5$_{Sw1}$K146Q (**B**) bound to 2.5μM MTs. (*C* and *D*) Mole fraction of docked NL (**C**) or closed Sw1 (**D**) for MT-bound K146 (blue) or K146Q (red) samples as in *A* and *B*. (**E**) Rate constant vs. [ATP] for the fast (closed) and slow (open) phases of biexponential fits to mole fraction transients as in *C* for Eg5$_{NL}$K146 (blue) or Eg5$_{NL}$K146Q (red). (**F**) Rate constant vs. [ATP] for the fast phase (closed) of the mole fraction transients as in *D* for Eg5$_{Sw1}$K146Q fit by a single-exponential function over the first 50 ms or a single-exponential function over a range from 50 to 300 ms. Reprinted from (*Muretta et al.*, 2018); refer to same for details on experimental methods.

tight NL/Sw1 coupling. For kinesin 1, this ratio is 0.99 in the absence of nucleotide (*Muretta et al.*, 2015). Mixing with 2 mM ATP causes the ratio to rise (Figure 3.5A, light gray), and fitting to a single-exponential rate equation (Figure 3.5A, solid black line) yields a rate constant of $32.3 \pm 0.5\,s^{-1}$ and a final coupling ratio of 2.1. By contrast, the coupling ratio for Eg5 (Figure 3.5A, dark gray) is 0.29 in the absence of nucleotide and fitting the data subsequent to mixing with 2 mM ATP yields a

rate constant of $3.8 \pm 0.1\,s^{-1}$ and an extrapolated coupling ratio of 7.9 (Figure 3.5A, solid black curve). The rate constants for both kinesin 1 and Eg5 are consistent with the values of k_{cat} at the temperature of this experiment ($10°C$). The results show that, for kinesin 1, Sw1 and the NL are tightly coupled through ATP binding and hydrolysis—when NL docking occurs—and remain moderately coupled with entry into the steady state. By contrast, Sw1 and the NL for Eg5 are poorly coupled in the absence of nucleotide. Coupling improves with ATP binding and hydrolysis, but Sw1 and the NL become markedly uncoupled again with entry into the steady state. We performed a similar analysis on the K146Q mutant as well as on a second mutant (K146M), which provides an additional way of testing the importance of the α1-α2b salt bridge in Eg5 function; as in our collaborators' prior study (*Muretta et al.*, 2015), we plotted the evolution of the coupling ratio over time from the TR^{2}FRET waveforms that are illustrated in Figure 3.4. The plot of coupling ratio versus time after mixing with 2 mM ATP is depicted in Figure 3.5B for the K146Q (Figure 3.5B, red) and K146M (Figure 3.5B, magenta) constructs. Coupling ratios for these two constructs in the absence of nucleotide are 0.63 and 1.02 for K146Q and K146M, respectively. After mixing with ATP, the coupling ratios increase for both mutants, and fitting to the same rate equation (Figure 3.5B, solid black lines) reveals rate constants and final coupling ratios of $4.2 \pm 0.1\,s^{-1}$ and 3.4 for K146Q and $4.1 \pm 0.1\,s^{-1}$ and 2.1 for K146M, respectively. These results imply that the two Eg5 mutations that break the α1-α2b salt bridge enhance NL/Sw1 coupling to levels that resemble kinesin 1.

3.3.4   Motor function under load

In kinesin 1, the position of the NL regulates the kinetics of ATP hydrolysis—a process that depends, in turn, on the structure of Sw1 (*Clancy et al.*, 2011a). In such a highly coupled system, load could thereby regulate the ATPase cycle of this motor. Since the K146Q mutation increases the conformational coupling between Sw1 and
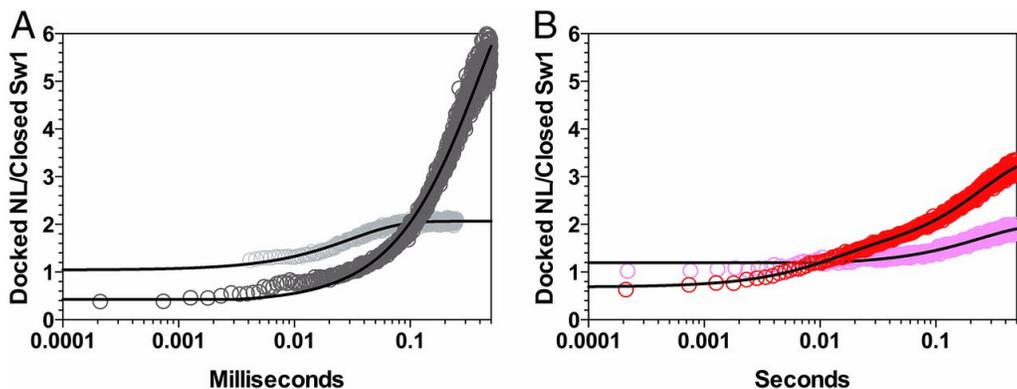
Figure 3.5: Transient changes in NL/Sw1 coupling. (**A**) Plot of the coupling ratio defined as the ratio of mole fraction of docked NL to mole fraction of closed Sw1 vs. time after mixing Eg5 (dark gray) and kinesin 1 (light gray) Sw1 and NL FRET-labeled constructs with 2 mM ATP. A ratio of 1.0 implies tight conformational coupling of the NL to Sw1, with deviation in either direction implying looser coupling. Fitting to a single-exponential rate equation (solid black lines) yields rate constants and extrapolated coupling ratios described in the text. (**B**) Corresponding experiments for Eg5K146Q (red) and Eg5K146M (magenta). For both constructs, the coupling ratio resembles kinesin 1 more than unmodified Eg5. Panel *A* reprinted from (*Muretta et al.*, 2015); panel *B* reprinted from (*Muretta et al.*, 2018).

the NL, we might expect that it would also alter the load dependence of the motor mechanochemical ATPase cycle. Therefore, our collaborators examined how load affects the mechanics of the K146Q acetylation mimetic at the single-molecule level using standard optical trap bead assays. The dimeric constructs used in these experiments consisted of the cys-light Eg5 motor domain (wildtype referred to as "D" and K146Q mutation as "DK146Q"), the NL, and neck coiled coil (human Eg5 residues 1-402), fused to the kinesin 1 hinge and coiled coil (human kinesin 1 residues 372-560). The maximum force that these motors experienced was about 0.07 pN at a position 150 nm from the trap center. Higher trap strengths were then used to characterize the response of the motors to load. Consistent with prior studies (*Muretta et al.*, 2015; *Valentine et al.*, 2006), our collaborators observed that D, like a WT Eg5 dimer, is quite insensitive to load, with single motors moving through a moderate-power optical trap at approximately constant velocity, although load increases as the motor

moves away from the trap center (*Shojania Feizabadi et al.*, 2015). Systematic set of measurements to determine the force–velocity curves for construct D revealed that it maintains an approximately constant velocity until $\sim$2.25 pN. It is able to advance under higher load, although velocity decreases as the load increases beyond 2.25 pN (Figure 3.6A). By contrast, the K146Q mutation increases the sensitivity to load, and a significant velocity decrease occurs by 1.75 pN and continues to decrease more substantially above this force. While moving quite slowly (Figure 3.6A), the K146Q mutation also allows the motor to access higher forces. Overall, the distribution of force production is shifted to higher forces for the K146Q motor (Figure 3.6B). We think several effects contribute to this. First, the decreased probability of detachment leads to achievement of higher forces and longer runs. Second, the slower velocity under load increases the duration of runs. Third, the increased stall probability and duration of stalls both contribute to longer periods of force generation. Figure 3.6B depicts the histogram from the "best" events, which will undercount the short duration events. In both cases short runs, especially those with lower force production, are likely to be undetected, and therefore, the actual single-molecule experimental data likely contain more short events than are detected.

It has previously been shown (*Mallik et al.*, 2005; *Kunwar et al.*, 2008) that sensitivity of velocity to load leads to optimal load sharing. In a team of such motors moving under load, the leading motors feels increased force and slow down, allowing the trailing motors to catch up. This improves load sharing and system performance, assuming that the forward motors do not detach prematurely. Consequently, our collaborators next examined the effect of load on the Eg5 detachment probability. They found that, for any given load, DK146Q has a lower probability of detachment than D. Next, the effect of the K146Q mutation on stall propensity and duration was examined. To identify stalls, a 100 ms minimum cutoff and ≤ 0.5 pN change in average force was used. Statistical distributions of observed stalls suggested that DK146Q
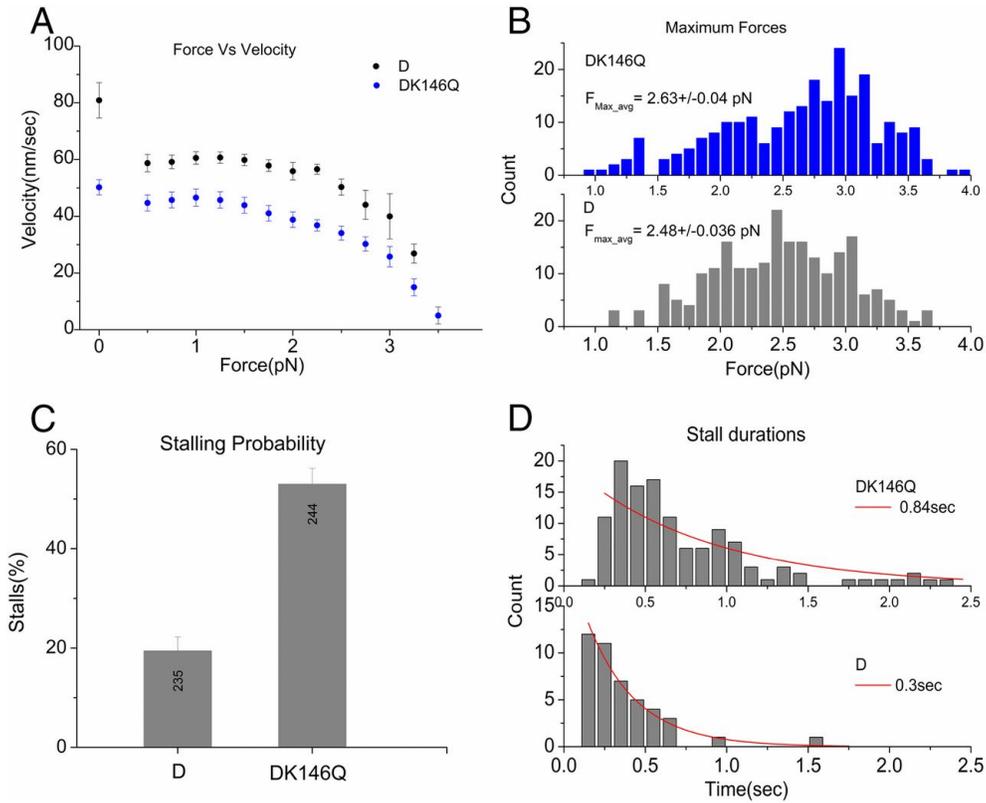
Figure 3.6: Single-molecule force–velocity curves for dimeric Eg5 constructs. (**A**) Force–velocity curves from experimental single-molecule data of D and DK146Q in bead assays. Compared with the D dimer, the K146Q-modified dimer shows increased sensitivity to load, with a slower velocity at all loads tested. (B)Maximum. (**B**) Maximum forces and durations for D and DK146Q single motor-containing beads (best15 traces from each single motor-containing bead). Overall, the distribution of force production is shifted to higher forces for the K146Q motor. **C** Stalling probabilities for D and DK146Q. **D** Stall durations for D and DK146Q single motor-containing beads. Reprinted from (*Muretta et al.*, 2018).

was much more likely to stall (Figure 3.6C) and that it stalled for a much longer duration (Figure 3.6D) than D. Furthermore, the smoothened velocities of individual traces as a function of load were obtained to determine the percentage of such traces with velocities below 20 nm/s at that load. At each load, more extremely low-velocity traces were observed for the K146Q mutant than WT. Thus, as well as decreasing the probability of detachment under load, the K146Q mutation increases stall probability and duration as well as the overall duration of force production. Since analyzing runs

that display stalls potentially introduces a selection bias, statistics on all recognizable runs using a 1 pN cutoff were also compiled to avoid noise. The presence of the K146Q mutation shifts the entire population of events to higher forces of longer mean duration. In summary, our data imply that Eg5 motors acetylated at K146 are less likely to fall off the MT while under load and more likely to maintain higher forces for longer durations—features that would improve the summation of forces generated from multiple motors in an ensemble.

3.3.5 Spindle pole separation in mitotic Cells

During metaphase, Eg5 works in large ensembles in a "tug of war" with cytoplasmic dynein. Our results suggest that, in the presence of opposing load, Eg5 acetylation will cause the motor to slow down, unlike WT Eg5. This effect along with its enhanced ability to stall mean that acetylated Eg5, even when present in low concentrations, will act as a "brake" by interfering with nonacetylated Eg5 motors. We, therefore, predict that expression of low molar fractions of the K146Q acetylation mimetic in cells will slow spindle separation during metaphase. To test this, our collaborators examined the effect of the K146Q mutation on spindle dynamics in cells. Constructs used in these experiments were mCherry (mCh)-tagged full-length WT or K146Q Eg5 proteins expressed in HeLa cells. Both WT and K146Q versions of the motor localized to the mitotic spindle in metaphase cells (Figure 3.7A). Measurements of mCh fluorescence along the pole to pole axis indicated that the spindle distribution of mCh-Eg5 K146Q did not differ from that for mCh-Eg5 WT (Figure 3.7B). To compare the functional activity of the WT and K146Q Eg5, HeLa cells expressing GFP-tubulin and equivalent levels of mCh-Eg5 WT or mCh-Eg5 K146Q were treated with the Eg5 inhibitor monastrol, resulting in mitotic arrest and the formation of monopolar spindles (Figure 3.7C and D). Bipolar spindle formation was then measured after monastrol washout (Figure 3.7C). Spindle lengths at
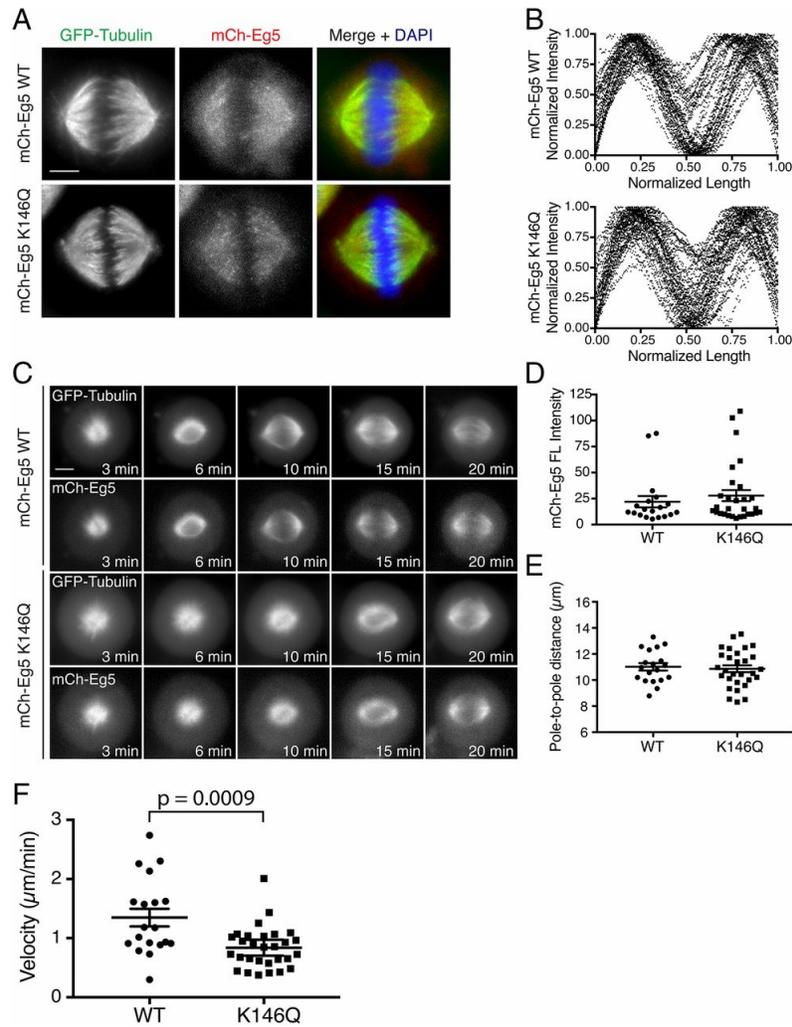
Figure 3.7: Expression of mCh-Eg5 K146Q slows pole separation during spindle formation. (**A**) Full-length mCh-Eg5 WT and mCh-Eg5 K146Q (red) localize to the mitotic spindle (green). (Scale bar: 5μm). (**B**) mCh-Eg5 WT (upper) and mCh-Eg5 K146Q (lower) are similarly distributed along the length of the spindle ($n = 43$ cells from three independent experiments). **C** Representative images of cells expressing GFP-tubulin and the indicated mCh-Eg5 construct. Time stamps indicate time after monastrol washout. (Scale bar: 5μm). **D** Stills from representative time-lapse movies of cells expressing GFP-tubulin and the indicated mCh-Eg5 construct. Time stamps indicate time after monastrol washout. (Scale bar: 5μm). **E** Plot of final spindle length at the conclusion of pole separation after monastrol washout (n = 19 mCh-Eg5 WT and n = 29 mCh-Eg5 K146Q cells from four independent experiments, $P = 0.71$ unpaired t test). **F** Plot of pole separation velocity in cells expressing mCh-Eg5 WT (1.35 ± 0.15, mean $\pm$ SEM, $n = 19$) and mCh-Eg5 K146Q (0.84 ± 0.07, mean $\pm$ SEM, $n = 29$, $P = 0.0009$ unpaired t test). Reprinted from (*Muretta et al.*, 2018).

the completion of pole separation were similar in cells expressing mCh-Eg5 WT or K146Q ($11.02 \pm 0.29\,\mu$m WT, $10.87 \pm 0.26\,\mu$m K146Q, mean $\pm$ SEM) (Figure 3.7E). However, pole separation occurred at a significantly slower velocity in cells expressing mCh-Eg5 K146Q ($0.84 \pm 0.07\,\mu$m/min) than in cells expressing mCh-Eg5 WT ($1.35 \pm 0.15\,\mu$m/min, $P = 0.0009$, unpaired t-test) (Figure 3.7F). Taken together, these data indicate that low levels of Eg5-K146Q in mitotic cells can act as a brake to slow spindle pole separation.

3.4 Methods

3.4.1 Molecular dynamics simulations

Simulated wild-type and mutant Eg5 systems in complex with ATP and tubulin were constructed based on the crystallographic structure of Eg5-AMPPNP (*Parke et al.*, 2010a) (PDB ID 3HQD) superimposed on helix α4 of kinesin-1-AMPPNP-tubulin (*Gigant et al.*, 2013) (PDB ID 4HNA). Simulations utilized the AMBER 12 package (*Case et al.*, 2012) and corresponding all-atom potential function ff99SB (*Hornak et al.*, 2006). The structures were inserted in a cubic box (margins located at 12 Å from the closest protein atom) filled with TIP3P water molecules and sodium counter ions to neutralize charge. For all systems, four runs of 4000 cycles each of energy minimization were performed alternating the steepest descent and conjugate gradient algorithms. Positional restrains of 500 kcal / (mol·Å^2) were applied on protein and ligand atoms in the first run, on backbone atoms in the second run, and on the water and ions in the third run. No atoms were restrained in the fourth run. Two consecutive molecular dynamics simulations runs were then used to heat the systems from 100K to 300K for 10ps and to equilibrate them at 300K for 300ps. Four production runs of 100ns (total 400ns) each were performed to evaluate the internal dynamics of the motor domain in the different systems. Full particle mesh-Ewald

electrostatics and a 12 Å cutoff value were used to treat non-bonded interactions. Periodic boundary conditions, 2fs time step, and constant pressure (1atm) were applied. Covalent bonds formed by hydrogen atoms were constrained with the SHAKE algorithm.

3.4.2   Residue-residue distance differences

Statistically significant residue-residue distance differences between WT and acetylated or mutant systems were identified with ensemble difference distance matrix (eDDM) analysis routine (*Muretta et al.*, 2018) which is described below.

Analysis of residue-residue distances have been used for studying functional dynamics of biomolecules, both in form of distances matrices (*Elber and Karplus*, 1987) and in terms of changes in residue contact maps generated from such distance matrices (*Doshi et al.*, 2016). Instead of averaging residue-residue differences or converting them into contact maps (adjacency matrices) by use of a cutoff, we modify long-range distances before comparison between systems. The distances are processed by applying a smooth function to mask long distance as follows:

$$f(x) = \begin{cases} x, & \text{if } x \leq c1 \\ c1 + 2 \cdot \left(\dfrac{c2 - c1}{\pi}\right) \cdot \cos\left(\dfrac{\pi}{2} \cdot \dfrac{c2 - x}{c2 - c1}\right), & \text{if } c1 < x \leq c2 \\ c1 + 2 \cdot \left(\dfrac{c2 - c1}{\pi}\right), & \text{if } x > c2 \end{cases} \quad (3.1)$$

where x is residue-residue distance, $c1$ and $c2$ are parameters of the smooth function, set to 4 Å and 8 Å, respectively. The above routine reduces the difference between long distances while difference between short distances are kept intact. This approach highlights statistically significant residue-residue contact differences between systems without the need for pre-alignment or structural superposition to reference conformation.

A total of 400 conformations were obtained for each of the WT, acetylated and

mutant systems by extracting 100 equally time-spaced conformations from the last 20ns of each simulation replicate. Distance matrices for each state were constructed from residue-residue distances, defined as the minimum distance between all heavy atoms of every residue pair in a given conformation. The resulting distance matrices of dimension $363 \times 363 \times 400$ from each state were processed as described above to mask long distance differences. The significance of residue-residue distance differences was evaluated with the Wilcoxon test (see section 8.0.2 for suggested improvements to this method). Residue pairs showing a p-value $< 10^{-5}$ and an average masked distance difference > 1 Å were considered statistically significant residue-residue distance differences for further analysis (see Figure 3.1B and Table B.1 in Appendix B).

3.4.3 Residue-residue correlated motion

To characterize correlated atomic fluctuations in each simulation, the heavy-atom residue-wise Pearson correlation was determined as follows:

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\sqrt{\langle \Delta r_i^2 \rangle \cdot \langle \Delta r_j^2 \rangle}} \tag{3.2}$$

where Δr_i is the displacement from the mean position of the ith atom in a trajectory. This analysis was performed separately on each replica, resulting in four matrices per protein state. The absolute value correlation matrices were then filtered by a contact map matrix built using the distance between carbon alpha atoms in the following way:

$$CM_{ij} = \begin{cases} 1, & \text{if } D_{C\alpha-C\alpha} \leq 10\text{Å and present in at least 1\% of frames} \\ 0, & \text{if } D_{C\alpha-C\alpha} > 10\text{Å or not present in at least 1\% of frames} \end{cases} \tag{3.3}$$

A principal component analysis transformation was then applied on the accumulated correlation matrices to identify the structural regions showing the largest variation

between simulated wild-type and K146Q mutant trajectories (red labels in Figure 3.3).

3.4.4 Residue-residue energetics

Molecular mechanics with generalized Born and surface area solvation (MM-GBSA) energy calculations were performed on each simulation with the GB^{OBC} model (*Onufriev et al.*, 2004) in AMBER 12 (*Case et al.*, 2012). Average pair-wise energy values (resulting from setting idecomp=4) are reported in the Results section for residue pairs found to differ significantly between wild-type and K146Q mutant simulations (*p*-value $< 10^{-5}$ and with a difference > 2 kcal/mol).

3.4.5 Metadynamics simulations

To study the partial undocking process of the neck linker, well-tempered metadynamics simulations (*Barducci et al.*, 2008) were then employed on both the Eg5 WT and K146Q systems. The Gromacs 5.0.4 package (*Abraham et al.*, 2015), AMBER ff99SB force field and Plumed 2.1.2 routine (*Bonomi et al.*, 2009) were used. The energetic bias was added on a collective variable corresponding to the distance of the hydrogen bond formed by the backbone oxygen atom of residue G96 (helix $\alpha 1$) and the side chain nitrogen atom of residue N366 (neck linker). Each system was studied for 700ns, setting the initial deposition height for the biasing Gaussian to 1kj/mol and the width to 0.125Å. The bias factor was equal to 9, the deposition rate to 10ps, and the temperature to 300K. An upper potential wall with an energetic constant of 700 kj/mol was applied at 12Å. The free energy profile was obtained using the Plumed routine *sum_hills*, and the convergence of the simulations was monitored through the changes in the free energy surface expressed using G96 to N366 collective variable.

For rest of the methods pertaining to experimental results, please see our publications (*Muretta et al.*, 2015, 2018).

3.5   Discussion and Conclusion

Multiple components of the cytoskeleton, such as the actomyosin and microtubule (MT) cytoskeleton, are frequently modified posttranslationally (*Buss and Kendrick-Jones*, 2008; *Terman and Kashina*, 2013; *Skoumpla et al.*, 2007). PTMs play a central role in regulating MT dynamics and function, and has been referred to as the "tubulin code" (*Yu et al.*, 2015; *Janke*, 2014; *Sirajuddin et al.*, 2014). In comparison, less is known about the roles that PTMs have on kinesin function, especially their effect on motor function. In kinesin 1, serine 175 at the amino-terminal end of the $\alpha 3$ helix can be phosphorylated by the JNK3 kinase, which reduces stall force by 20% (*DeBerg et al.*, 2013). A recent report has also described src-mediated phosphorylation of tyrosine residues in the motor domain of Eg5, including those located in the $\alpha 3$ helix and in Loop 5; in close proximity to residue K146 under study in this work (*Bickel et al.*, 2017). However, these studies do not provide insight into how PTMs alter motor mechanochemistry, only that these modifications reduce motor function. In prior studies of Eg5 (*Muretta et al.*, 2015; *Behnke-Parks et al.*, 2011), our experimental collaborators showed that Loop 5, which splits the $\alpha 2$ helix into $\alpha 2$a (N terminus) and $\alpha 2$b (C terminus), functions in part by regulating the conformational coupling between the catalytic site and the NL and that a mutation at the junction of Loop 5 with $\alpha 2$b (P131A) had critically affected on this coupling. This motivated us to look for a documented PTM in $\alpha 2$b helix with local structural and allosteric effects that could be predicted, and this led us to examine the mechanochemical consequences of lysine 146 acetylation.

Acetylation has long been recognized as a PTM of histones (*Marmorstein and Zhou*, 2014; *Shilatifard*, 2006; *Verdone et al.*, 2005; *Forsberg and Bresnick*, 2001; *Jenuwein and Allis*, 2001; *Reiner et al.*, 1993). Over the recent years it has been appreciated that it also occurs in other components throughout the cell, including transcription factors, metabolic enzymes, and structural proteins (*Menzies et al.*, 2016;

49

Kouzarides, 2000; *Blee et al.*, 2015). While we did not determine the acetyl trans-ferase(s) responsible for K146 acetylations, we have shown that low levels of acetylated K146 Eg5 in cells are sufficient to affect ensemble motor function——consistent with our collaborators' finding that modification of only a minority of Eg5 is sufficient to significantly slow spindle pole separation (Figure 3.7).

Our MD simulations predicted large-scale effects from eliminating the α1-α2b salt bridge. This includes increased interactions between α1, β7, the CS, L13, and the NL as well as between α0, the P loop, and Sw1 (Figures 3.1, 3.2, 3.3, and Table 3.1). In particular, a cluster of charged residues, including E92 (α1), K17 (CS), K362 (NL), and R329 (Loop 13), displays increased interactions on salt bridge disruption that together maintain the CS and NL in a state that is more frequently coordinated and docked. Metadynamics simulations also revealed an enhanced NL docking potential for the mutant (Figure 3.1C). In concert with these changes, Sw1 more frequently adopts a closed conformation primarily due to increased coordination of Loop 9, α0, and P-loop. Collectively, these results suggest that acetylation of K146 produces a motor where nucleotide binding/hydrolysis and NL docking are tightly coupled. Loose coupling in unmodified Eg5 would prevent one motor from acting as a brake that could slow or stop other motors in the ensemble. However, in some circumstances where Eg5 opposes the effects of cytoplasmic dynein, there may be a need for Eg5 to act more like kinesin 1—stalling rather than dissociating. Utilizing TR^{2}FRET, our collaborators found that both the K146Q and K146M acetylation mimetics acceler-ate ATP-induced NL docking 2.2- to 3.4-fold (Figure 3.4C and E). This ensures that Sw1 and the NL remain conformationally coupled through ATP hydrolysis and P$_i$ re-lease, when the power stroke occurs, and indicates that K146 acetylation makes Eg5 (mechanochemically) resemble kinesin 1. In a prior study where our collaborators compared kinesin-1 and Eg5 using TR^{2}FRET (*Muretta et al.*, 2015), they observed that when Sw1 in Eg5 is closed, it stabilizes Sw2 into a strong MT binding confor-

mation. The increased mole fraction of closed Sw1 that we observe with the K146Q mutant after ATP hydrolysis implies that acetylation should enhance the fraction of the time that the motor is strongly bound to the MT. Taken together, our MD and TR^{2}FRET studies predict that the consequences of K146 acetylation should be particularly apparent under load, in addition to single-molecule unloaded conditions.

Our MD and TR^{2}FRET studies predict that, unlike other kinesin PTMs, acetylation of K146 in Eg5 produces a "gain in function", which enhances Eg5's ability to work in teams to overcome dynein-produced opposing load. The single-molecule studies carried out by our collaborators reveal that acetylation of K146 leads to better ensemble motor function under load (*Muretta et al.*, 2018), and the capabilities that they provide Eg5 are remarkably similar to those provided for dynein by its cofactors NudE and Lis1. NudE and Lis1 alter dynein detachment kinetics under load, with individual dynein motors holding on to MTs longer, allowing better ensemble function as motors share load and enabling the complex to transport a cargo against significant opposition (*McKenney et al.*, 2010; *Reddy et al.*, 2016). Remarkably, acetylation of K146 does essentially the same thing for Eg5. On average, dynein with Lis1 and NudE holds on to MTs 60% longer before detaching compared with dynein alone (*McKenney et al.*, 2010), an effect of comparable magnitude to what we observed in the K146Q mutation. This seems to be a more prelevant form of modulating motor function, than previously appreciated. It implies that the ability to modulate a motor's performance under load is an important variable for multiple molecular motors, regardless of their evolutionary source.

By increasing conformational coupling and the lifetime of strong MT binding states, the K146Q acetylation mimetic mutation alters the force–velocity landscape of Eg5 to enhance motor efficiency in ensembles, reduce velocity in the presence of opposing load, and increase the probability of stalling. This would enable modified Eg5 to act as a "brake", slowing centrosome separation by generating drag force

against unmodified Eg5 motors. Spindle function involves a complex orchestration of multiple motors that work in opposition to each other, and regulation of motor mechanical output is necessary to fine tune this balance of forces. Our collaborators, using time-lapse microscopy, observed that only a fraction of Eg5 with acetylation mimetic mutation is cabable of slowing spindle pole separation (Figure 3.7). At least one other member of the kinesin 5 family of mitotic motors, BMK-1 from *C. elegans*, has also been reported to act as a spindle brake (*Saunders et al.*, 2007). Unlike Eg5, BMK-1 is not essential for mitosis or normal development. Its deletion accelerates spindle pole separation over two-fold, implying that it provides a force that opposes spindle elongation. Sequence analysis of the Eg5 and BMK-1 motor domains reveal that the $\alpha1$-$\alpha2b$ salt bridge seen in Eg5 is absent in BMK-1. In particular, Eg5 residues D91 ($\alpha1$) and K146 ($\alpha2b$) are replaced in BMK-1 by residues K83 and Q136, respectively. Our results in conjunction with this prior work imply that there is an evolutionarily conserved need for some kinesin 5 motors to act for at least part of the cell cycle as a brake and that this need in some organisms might be served by synthesis of a non-essential kinesin 5, while in others, it is served through a reversible PTM that provides cells with functional flexibility.

The increased conformational coupling that we observe with MD and TR^{2}FRET for K146Q Eg5 is largely a result of the threefold acceleration of NL docking, which allows this process to track with Sw1 closure. Our work, therefore, leads us to predict that NL docking should be relatively slow for kinesins with an $\alpha1$-$\alpha2b$ salt bridge and appreciably faster in kinesins that lack it. This prediction is consistent with the limited set of data on the kinetics of NL docking, which shows that it is slow in both Eg5 $(60 - 80\,s^{-1})$ (*Rice et al.*, 1999a) and kinesin 7 motor CENP-E $(\sim 29\,s^{-1})$ (*Rosenfeld et al.*, 2009). Both of these motors contain an $\alpha1$-$\alpha2$ salt bridge, with D91 and K146 in Eg5 corresponding to D72 and K116 in CENP-E, respectively (*Garcia-Saez et al.*, 2004). By contrast, kinesin 1 is devoid of this ionic interaction (*Jon Kull et al.*, 1996).

We would, therefore, predict that NL docking in kinesin 1 should be faster than in WT Eg5 or CENP-E, and in fact, it is ($> 800\,s^{-1}$) (*Rosenfeld et al.*, 2003). Rigorously characterizing the relationship between the $\alpha1$–$\alpha2$b interaction and NL docking kinetics will require additional comparisons with other kinesins. Two examples could serve as the basis for future studies. First, we note that kinesin 12 motor Kif15, another mitotic kinesin, is also devoid of an $\alpha1$-$\alpha2$ salt bridge (*Klejnot et al.*, 2014), and our results would predict that NL docking should be rapid in this motor. Second, at least one other kinesin (kinesin 10 Kif22) has a documented PTM in $\alpha2$b helix (*Rigbolt et al.*, 2011; *Olsen et al.*, 2010). Residue T158, corresponding to position 147 in Eg5, has shown to have been phosphorylated in this motor. Phosphorylated T158 could make a salt bridge with arginine 113, corresponding to residue D91 in Eg5. The computational and experimental approaches we have utilized in our study of K146 acetylation in Eg5 should be readily applicable to elucidate the consequences of altering the interaction between $\alpha1$ and $\alpha2$ in other kinesins, and discern the changes in their mechanochemistry.

CHAPTER IV

Force Generation in Kinesin-1 Kif5C

Contents of this chapter were partially published in (*Budaitis et al.*, 2019). Citation: Budaitis, B. G., S. Jariwala, D. N. Reinemann, K. I. Schimert, G. Scarabelli, B. J. Grant, D. Sept, M. J. Lang, and K. J. Verhey (2019), Neck Linker Docking is Critical for Kinesin-1 Force Generation in Cells but at a Cost to Motor Speed and Processivity, *eLife*, doi: 10.7554/eLife. This has been a collaborative effort. The computational work was done by myself. The experimental work was carried out primarily by B.G.B in K.J.V lab at the University of Michigan (work related to Figures 4.1C-D, 4.6, 4.8, 4.9), with optical trapping assays carried out by D.N.R. in M.J.L. lab at Vanderbilt University (work related to Figure 4.5).

4.1   Abstract

Kinesin motor proteins utilize energy from ATP hydrolysis to direct key microtubule-based processes such as cell division, intracellular transport, and cell motility. How ATP hydrolysis is coupled to force generation in kinesin motor, its mechanochemistry, is a fundamental question in the cytoskeleton field. Kinesin force generation involves ATP-induced docking of the neck linker (NL), a 12-18 amino acid sequence, along the motor domain core. However, the roles of steps involved in this docking, namely cover-neck bundle formation (CNB) followed by NL docking, are not clear.

In this chapter, I will discuss how I applied molecular dynamics simulations and used subsequent analyses to identify residues critical for CNB formation and NL docking in kinesin 1 transport motor Kif5C. Their predicted requirement for force-generation were then tested through single-molecule, optical trap, and cell-based assays by our collaborators. Kinesin 1 mutant motors impaired in CNB formation and NL docking showcase reduced force output, but surprisingly exhibited increased speeds, run lengths, and landing rates under single-molecule conditions. The ability of mutant motors to transport high-load cargo in cells was also impaired. This study shows that NL docking is critical for force production in kinesin 1, but at a cost to speed and processivity.

4.2   Introduction

Kinesin superfamily of proteins (KIFs) consists of a diverse set of motor proteins—for example ~ 45 genes have been identified in mammalians—that are involved in critical cellular roles, including microtubule-mediated cargo transport in cells. In general kinesins comprise a kinesin motor domain, which adopts a conserved fold among the KIF proteins, and a coiled-coil domain invovled in kinesin dimerization and/or binding to cargo, adaptors or scaffold proteins (*Hirokawa et al.*, 2009). The kinesin motor domain contains sequence and structural elements involved in ATP binding and hydrolysis, as well as for coordinating with microtubules (MT). The kinesin motor domain generates force by hydrolyzing ATP, by coupling nucleotide-dependent conformational changes in the catalytic site to mechanical output that drives cargo transport.

A flexible 12-18 amino acid sequence at the C-terminus of the motor domain, called the neck linker (NL), links the motor domains in most dimeric kinesin motors (*Hariharan and Hancock*, 2009; *Kozielski et al.*, 1997). The NL has been suggested to serve as a structural element critical for both directed motility and force generation

of kinesin motors. For the transport motor kinesin-1, the prototypical kinesin, structural studies have shown that conformational changes in the NL are coupled to the nucleotide state of the motor domain, undergoing a transformation from being flexible in both the ADP-bound and nucleotide-free states to being docked along the core motor domain in the ATP-bound state (*Rice et al.*, 1999b; *Rosenfeld et al.*, 2001; *Sindelar et al.*, 2002; *Skiniotis*, 2003; *Asenjo et al.*, 2006; *Sindelar and Downing*, 2010; *Gigant et al.*, 2013; *Shang et al.*, 2014). NL docking of the leading motor domain positions the lagging motor domain forward along the microtubule track, thereby specifying the direction of motility in initiating a "hand-over-hand" motion. NL docking also coordinates the alternating (out of phase) ATPase cycles of the two motor domains to ensure processive stepping (*Case et al.*, 2000; *Tomishige and Vale*, 2000; *Hahlen et al.*, 2006; *Yildiz et al.*, 2008; *Clancy et al.*, 2011b; *Dogan et al.*, 2015; *Isojima et al.*, 2016; *Liu et al.*, 2017). The mechanism of nucleotide-dependent conformational changes in NL driving processive stepping has been shown in other members of the kinesin superfamily as well (*Nitta et al.*, 2008; *Muthukrishnan et al.*, 2009; *Shastry and Hancock*, 2010, 2011; *Atherton et al.*, 2014; *Cao et al.*, 2014; *Atherton et al.*, 2017; *Ren et al.*, 2018).

So how does the NL dock, and how does the "power-stroke" occur? The ATP-induced NL docking involves distinct interactions of the NL with the motor domain. The first half of the NL interacts with the coverstrand (CS/β0), a sequence stretch at the N-terminus of the motor domain, forming a two-stranded β sheet called the cover-neck bundle (CNB). This NL-CS/β0 interaction is believed to provide the "power-stroke" for force generation in kinesin-1 (*Hwang et al.*, 2008; *Khalil et al.*, 2008). Point mutations in the CS designed to disrupt β strand formation, or deletion of the entire CS in the fly kinesin-1 motor significantly reduced the motor's ability to withstand load in optical trap assays (*Khalil et al.*, 2008). CNB formation in response to ATP-binding has also been reported in other members of the kinesin superfamily, indicating

its role as a mechanical element in kinesin (*Atherton et al.*, 2017; *Hesse et al.*, 2013). Following CNB formation, the second half of the NL (β10) is predicted to dock along the surface of the core motor domain. An asparagine residue (N334 in kinesin-1) is predicted to serve as a "latch" (asparagine latch or the N-latch), interacting with β7 of the motor core to hold the NL in a docked position. This asparagine residue is conserved in most kinesins, indicating that the N-latch formation may also be a conserved feature of kinesin force generation.

For an isolated, single-headed kinesin motor, do CNB formation and NL docking purely serve a mechanical role in kinesin processivity? Previous studies, including one of my own, have shown that the NL and active site are indeed coupled (*Muretta et al.*, 2015, 2018). So the next question was, is CNB and/or N-latch formation critical for multiple kinesin motors to drive transport of membrane-bound cargo under physiological conditions in cells? To address this, we first began by identifying the residues critical for CNB and/or N-latch formation using molecular dynamics. Next, our collaborators combined *in vitro* single molecule assays, and cell-based transport assays to delineate how NL docking influences kinesin-1 motors cooperating in teams to transport membrane-bound cargoes in cells. We found that while single mutant motors were faster and more processive under unloaded conditions, groups of mutant motors were severely crippled in their ability to transport high-load cargo in cells. Strikingly, the mutant motors were fully able to transport low-load cargo in cells. Overall, this work highlights the critical role of CNB and N-latch formation in the transport of high-load cargo in cells.

4.3 Results

4.3.1 Charazterizing CNB and NL dynamics in apo and ATP-bound kinesin-1

To identify residues critical for the nucleotide-dependent formation of CNB and NL formation, we first compared the experimentally determined structures of apo and ATP-bound kinesin-1 motor in complex with tubulin heterodimer. High resolution structures of kinesin-1 motor KIF5B were selected—PDB 4LNU for motor in apo (nucleotide-free) state (*Cao et al.*, 2014), and PDB 4HNA for motor in ATP-bound state (*Gigant et al.*, 2013), respectively. As with most available experimentally determined structures of kinesins, the NL is flexible in apo state and hence, unresolved. In the ATP-bound state, on the other hand, the CNB is formed and the NL is docked along the core of the motor domain. To identify nucleotide-dependent changes in residue-residue distances between the CS, NL, and the rest of the motor domain, we first generated models of *Rn*KIF5C in both apo and ATP-bounds states, and performed molecular dynamics (MD) simulations of the models (motor in complex with tubulin heterodimer). Four replicate simulations each were carried out for motors in the apo state and the ATP-bound state (see *Methods* section below). The differences in residue-residue distances between the two states were compared as described in Chapter III and also in my previous analyses of kinesin-5 (*Muretta et al.*, 2018). Briefly, we compared residue-residue distances between the apo and ATP-bound states to identify statistically significant distance differences ($P < 10^{-5}$). In the apo state, the NL is flexible (Figure 4.1A top) and forms few interactions with the motor domain (Figure 4.1B) while the CS interacts with residues in $\alpha 4$ and in Loop13 (Figure 4.1B, red boxes marked CS-$\alpha 4$ and CS-L13). Specifically, the C-terminal residue (CTR) of the CS (I9) points down into a hydrophobic pocket called the docking pocket (*Sindelar*, 2011) where it contacts residues I266, L269, and A270 of $\alpha 4$ and the remaining

residues of the CS contact Loop13. Collectively, these interactions sterically block the NL from accessing the docking pocket in the apo state.

In the ATP-bound state, the NL is docked along the core motor domain and maintains this interaction throughout the trajectories from MD simulations (Figure 4.1A bottom, Figure 4.1C). The N-terminal first half of the NL, $\beta9$ forms contacts with the CS to form the cover-neck bundle (CNB) (Figure 4.1B, blue box marked NL-CS) as well as contacts with $\alpha4$ and Loop13 (Figure 4.1B, blue boxes marked NL-$\alpha4$ and NL-L13). These contacts are made possible by the ATP-dependent formation of an extra turn at the end of $\alpha6$, the NL initiation sequence (NIS (*Nitta et al.*, 2008)), that positions $\beta9$ for insertion between the CS and $\alpha4$ (*Sindelar*, 2011; *Lang and Hwang*, 2010). The first residue of $\beta9$ (I327) now occupies the docking pocket and forms contacts with residues I266, L269, and A270 of $\alpha4$. The remaining residues of $\beta9$ interact with the CS via a series of backbone interactions to complete CNB formation. For the C-terminal second half of the NL, $\beta10$ docks along the core motor domain through backbone hydrogen bond interactions with $\alpha1$ and $\beta7$ (Figure 4.1B, blue boxes marked NL-$\alpha1$ and NL-$\beta7$). The N-latch residue (N334) forms backbone interactions with E76 and conserved residue G77 in $\alpha1$, and residues S225 and L224 in $\beta7$. The remaining residues of NL ($\beta10$) provide further interactions with $\beta7$ to complete NL docking.

Our MD simulations build upon previous work (*Nitta et al.*, 2008; *Hwang et al.*, 2008; *Khalil et al.*, 2008; *Hwang et al.*, 2017), and identify residues critical for regulating NL docking. In the nucleotide-free (apo) state, residue I9 of the CS occupies the docking pocket bordered by $\alpha6$, $\alpha4$, and L13, and the NL remains flexible and undocked. In the ATP-bound state, residue I327 at the N-terminus of the NL occupies this pocket and initiates the NL docking along the core motor domain. Furthermore, residue N334 (N-latch) interacts with both $\alpha1$ and $\beta7$ to position the NL along the core motor domain. As noted in previous studies (*Hwang et al.*, 2008; *Khalil et al.*,
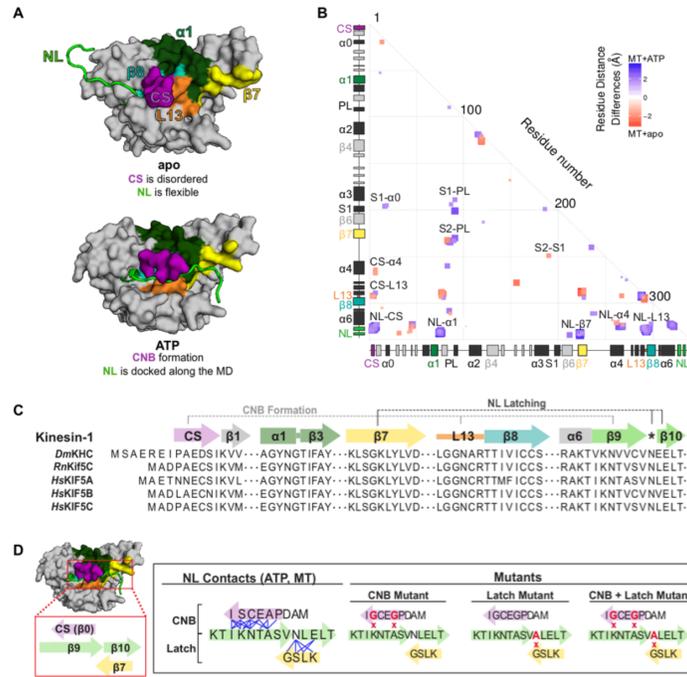
Figure 4.1: Key interactions between the kinesin-1 NL and motor domain. (**A**) Kinesin-1 (RnKif5C) motor domain in the nucleotide-free (apo) state (top) or ATP-bound, post-power stroke state (bottom), shown as surface representation. The neck linker (NL, light green) is represented as a cartoon and is flexible in the apo state and docked along the motor domain in the ATP-bound state. Coverstrand (CS, purple), $\alpha 1$ (dark green), $\beta 7$ (yellow), Loop13 (L13, orange), $\beta 8$ (teal), neck linker (NL: $\beta 9$ and $\beta 10$, light green). (**B**) Differences in residue-residue distances between kinesin-1 motors in the apo versus ATP-bound states as determined from MD simulations. Secondary structure elements along the axes with α helices colored in black, β strands in grey, or colored according to (*A*). Residue-residue interactions that are significantly closer ($P < 10^{-5}$) in the apo state (red) or ATP-bound state (blue) are indicated. The magnitude of the distance change is indicated by color intensity. Interactions between key structural elements are labeled. Complete list of inter-residue distances in Appendix C, Table C.1. (**C**) Sequence alignment of elements indicated in (*A*) from kinesin-1 motor domain across species (*Dm*, *D. melanogaster*; *Rn*, *R. norvegicus*; *Hs*, *H. sapiens*); asterisk indicates the asparagine-latch (N-Latch, N334 in *Rn* kinesin-1). (**D**) Key structural elements involved in CNB formation and NL latching in WT and mutant motors. The first-half of the NL ($\beta 9$, light green) interacts with the C-terminal end of the CS (purple) to form the cover-neck bundle (CNB). The second half of the NL (N-Latch and $\beta 10$) interacts with $\beta 7$ (yellow) of the core motor domain for NL docking. Residue-residue contacts for NL docking are depicted as blue lines. Point mutations generated in this study to disrupt CNB formation, N-latch formation, or both are shown in red text. Figure from (*Budaitis et al., 2019, in final revision at eLife*)

60

2008), an asparagine residue in the NL between $\beta 9$ and $\beta 10$ is a conserved feature of many kinesin motors with an N-terminal motor domain, potenially playing a critical role in kinesin motility and force generation in other kinesins as well (Figure 4.1C asterisk).

4.3.2 MD simulations of CNB and N-latch mutants of kinesin-1

To delineate the effects of disrupting CNB or N-latch formation, we carried out MD simulations of the Latch and CNB+Latch mutant motors in the microtubule- and ATP-bound states (post-power stroke). To disrupt formation of the N-latch, residue N334 of the NL was mutated to an alanine residue. We also wanted to test the effect of disrupting CNB formation in tandem. To hinder formation of CNB, CS residues A5 and S8 were mutated to glycine residues (Figure 4.1D, CNB mutant), which have a low propensity to form a β sheet (*Minor and Kim*, 1994). The A5G/S8G double mutant was previously reported to impair force generation for single *D. melanogaster* kinesin-1 motors in optical trap experiments (*Khalil et al.*, 2008). However, the effects of these mutations on force generation in mammalian motors has not been tested previously.

The simulation trajectories of Latch and CNB+Latch mutants were analyzed as described in the previous section (also see *Methods* section below). For the Latch mutant, the simulations predict that the N-latch and $\beta 10$ residues make fewer interactions with $\alpha 1$ and $\beta 7$ and that the distances of these interactions are increased (Figure 4.2). For the CNB+Latch mutant, the simulations predict that mutation of the CS (A5G,S8G) results in intra-CS interactions (Figure 4.3D,E) rather than interactions with $\beta 9$ of the NL (Figure 4.3A,B) and that mutation of the N-latch residue (N334A) results disrupts its interactions with $\alpha 1$ and $\beta 7$ (Figure 4.3A,C). Thus, mutations of CS and N-latch residues disrupt CNB formation and NL latching, respectively.

61
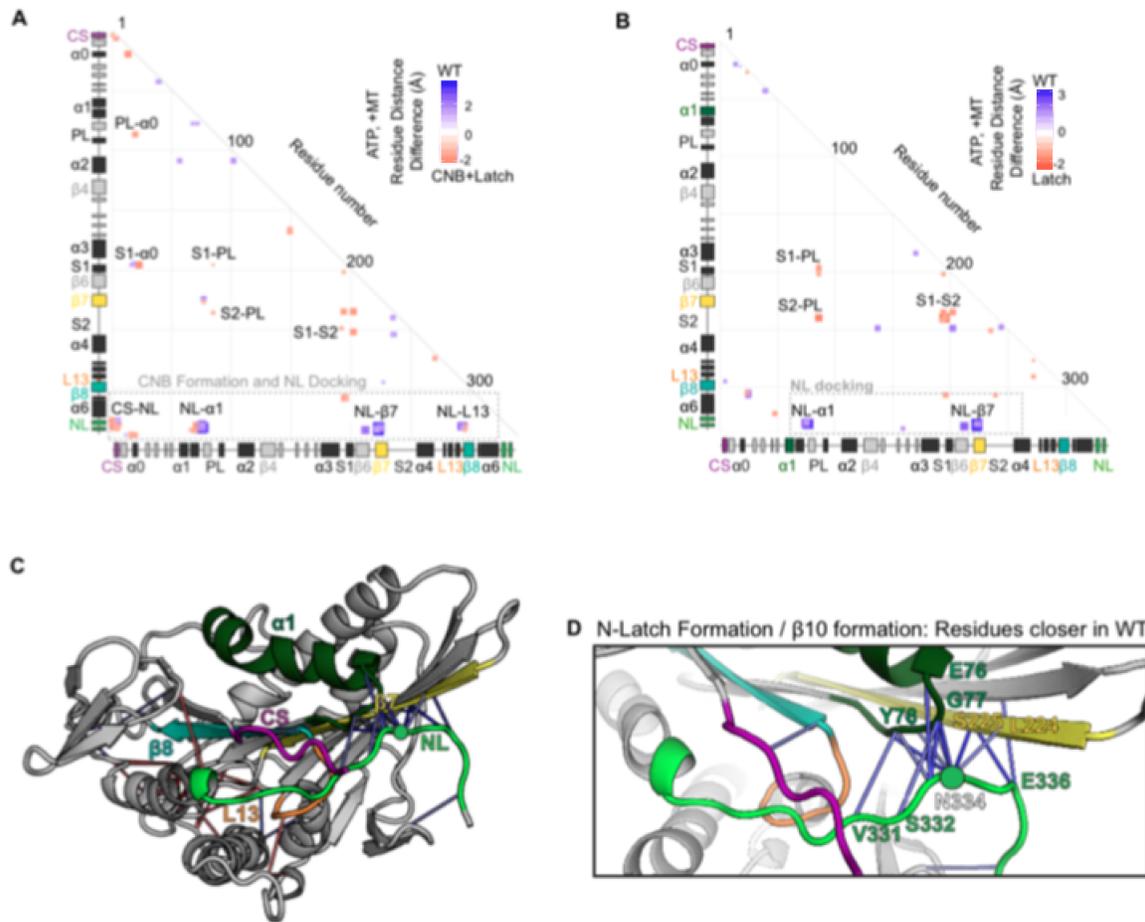
Figure 4.2: Differences in inter-residue distances for WT, Latch, and CNB+Latch motor domain. Differences in residue-residue distances between WT and (**A**) CNB+Latch mutant, or (**B**) Latch mutant based on MD simulations of microtubule-bound motors in the ATP-bound state. Colors according to Figure 4.1. Distances that are significantly ($P < 10^{-5}$) shorter in the mutant state (red) or WT state (blue) are displayed. The magnitude of the distance change is indicated by color intensity; interactions between elements are labeled. A complete list of residue distance differences are in Appendix C, Tables C.2 and C.3. Figure from (*Budaitis et al.,* 2019, *in final revision at eLife*)

The distance analyses also highlighted potential allosteric effects of Latch and CNB+Latch mutations on the nucleotide and microtubule binding regions of the motor domain. First, simulations of CNB+Latch mutant revealed enhanced interactions between elements important for coordinating and hydrolyzing nucleotide, as compared to the wildtype (WT) (Figure 4.4D,E,F). Specifically, the residue-residue distances
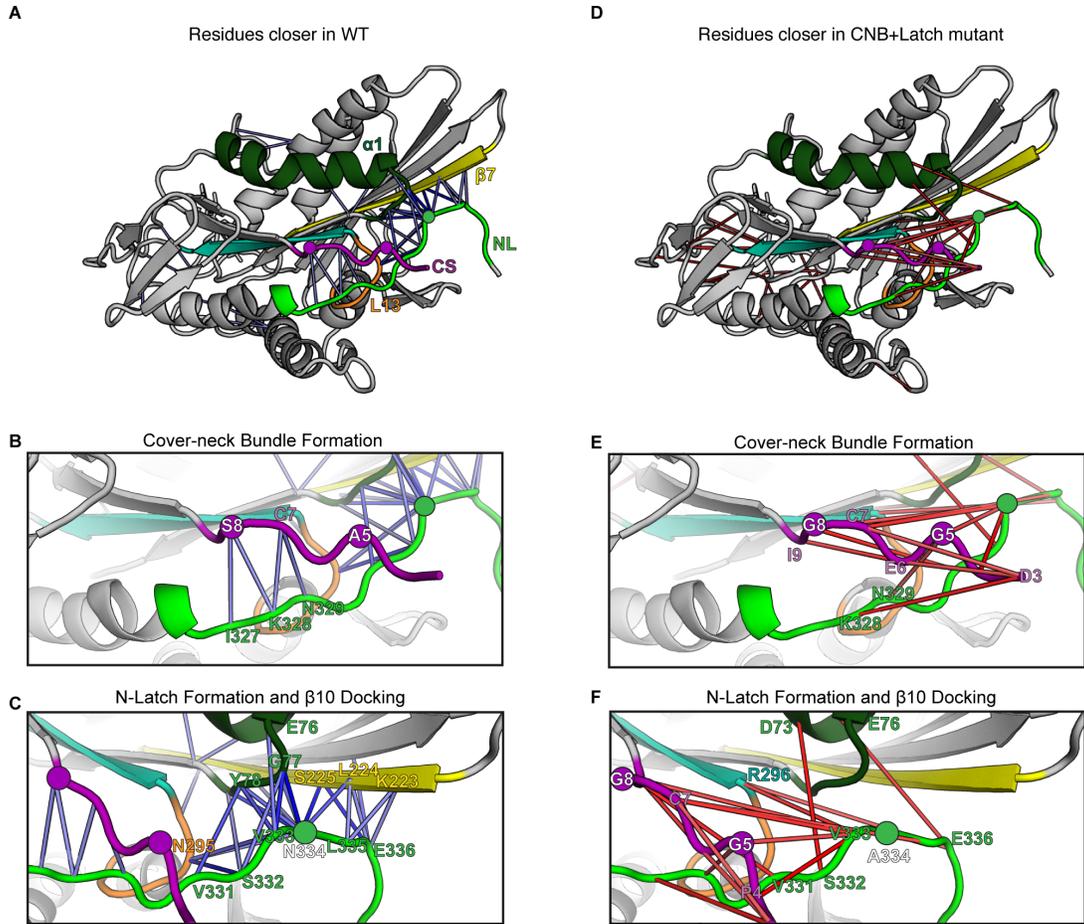
Figure 4.3: MD simulations predict that CNB+Latch mutations alter CNB formation and NL docking. (**A,D**) Cartoon representation of kinesin-1 motor domain in the ATP-bound, post-power stroke state. Residue-residue distances that are significantly ($P < 10^{-5}$) closer in (*A*) WT are shown as blue lines, and distances closer in (*D*) CNB+Latch mutant shown as red lines. (**B,E**) Contacts between the CS (residues S8, C7) and the NL (β9 residues I327, K328, N329) are shorter in the WT motor (*B*), suggesting that CNB formation is disrupted in the CNB+Latch mutant. The mutated CS (*E*) makes intra-CS contacts rather than interactions with the NL. (**C,F**) The WT motor shows shorter contacts for (i) the N-latch (N334) with β7 (L224, S225) and α1 (G77, Y78) residues, (ii) the N-terminal half of the NL (β9 residues V331, S332, V333) with the core motor domain (L13 residue N295 and α1 residues E76, G77, Y78), and (iii) the C-terminal half of the NL (β10 residue E336) with the core motor domain (β7 residues L224, S225). NL docking is disrupted in the CNB+Latch mutant (*F*) making interactions with the CS rather than β7. Figure from (*Budaitis et al.,* 2019, *in final revision at eLife*)

are shorter between the P-loop and $\alpha 0$ (Figure 4.4D, red square PL-$\alpha 0$; Figure 4.4E-F; Table 4.1: E22-S89 distance 4.40 ± 1.29 Å in CNB+Latch versus 7.49 ± 1.97 Å in WT). As the P-loop coordinates ATP in the nucleotide pocket and $\alpha 0$ gates ATP binding (*Hwang et al.*, 2017), this result suggests that modulating NL docking influences nucleotide binding and/or catalysis. Shorter residue-residue distances are also observed between switch 1 and $\alpha 0$ (Figure 4.4D, red square S1-$\alpha 0$; Figure 4.4E-F; Table 4.1: R25-M198 distance 3.92 ± 0.56 Å in CNB+Latch versus 7.38 ± 2.28 Å in WT) and between switch 1 and switch 2 (Figure 4.4D, red square S1-S2; Figure 4.4E-F; Table 4.1: T196-E237 distance 4.35 ± 0.64 Å in CNB+Latch versus 6.97 ± 1.38 Å in WT). Second, enhanced interactions between residues involved in coordinating and hydrolyzing nucleotide are also observed in the Latch mutant (Figure 4.4G-I; Table 4.1). As closure of the switch regions is necessary for ATP hydrolysis (*Clancy et al.*, 2011b; *Cao et al.*, 2014; *Turner et al.*, 2001; *Parke et al.*, 2010b), these results indicate that the Latch and CNB+Latch mutations result in enhanced catalytic site closure and ATP hydrolysis that could account for the increase in velocity of the mutant motors under single-molecule, unloaded conditions, as briefly discussed in the experimental results below. We note that the allosteric coupling between between the NL and active site has been hypothesized before, and we have previosuly shown—both computationally and experimentally—that perturbing NL docking has measured effect on active site closure and nucleotide hydrolysis, findings we have shown before in kinesin-1 and kinesin-5 (*Muretta et al.*, 2015, 2018).

Table 4.1: Differences in inter-residue distances from MD simulations of WT and CNB+Latch mutant. Only residue pairs mentioned in *Results* text are listed, see Table C.1 in Appendix C for all pairs. Residue pairs, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference between wild type (WT) and mutant (CNB+Latch) simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop.

Residue position	WT dist. (Å)	CNB+Latch (Å)	CNB+Latch − WT (Å)
E22 ($\alpha 0$) − S89 (PL)	7.49 (1.97)	4.40 (1.29)	−3.09

Residue position	WT dist. (Å)	CNB+Latch (Å)	CNB+Latch − WT (Å)
R25 (α0) – M198 (L9/S1)	7.38 (2.28)	3.92 (0.56)	−3.46
T196 (L9/S1) – E237 (L11/S2)	6.97 (1.38)	4.35 (0.64)	−2.26
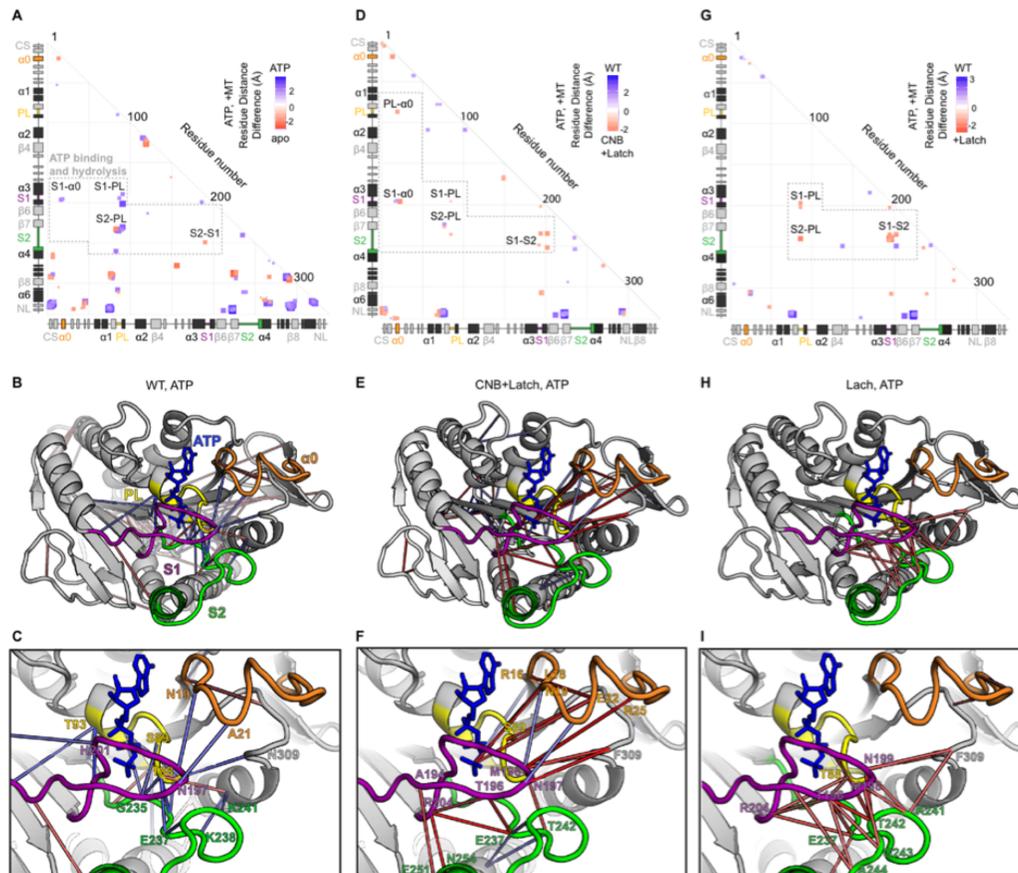
Figure 4.4: Interactions between nucleotide coordinating elements in WT, CNB+Latch, and Latch mutant motors. Significantly different ($P < 10^{-5}$) residue distances shorter in (**A-C** WT shown as blue squares or lines, shorter in (**D-F**) CNB+Latch mutant shown as red squares or lines, and shorter in (**G-I**) Latch mutant shown as red squares or lines, as compared to WT. (*C,F,I*) enlarged views of PL, S1, S2, and α0 interactions. Secondary structure elements colored as: P-Loop (PL, yellow), Switch 1 (S1, purple), and Switch 2 (S2, green), α0 (orange). (*E,F*) For the CNB+Latch mutant motor, enhanced interactions between S2-PL, S1-S2, α0-PL, and α0-S1 suggest enhanced nucleotide coordination and closure of the catalytic site. (*H,I*) For the Latch mutant motor, enhanced interactions between S1-PL, S2-PL, and S1-S2 suggest enhanced catalytic site closure. Figure from (*Budaitis et al., 2019, in final revision at eLife*)

4.3.3 Force output of CNB and NL mutants

Our collaborators at Vanderbilt University measured the force output of CNB and NL mutants using optical trap assays. The experimental setup is discussed in detail in our manuscript (*Budaitis et al.,* 2019, *in final revision at eLife*). The mean detachment forces of the motors under consideration (WT, CNB, Latch, and CNB+Latch), are listed in Table 4.2. Individual WT motors were motile in the absence of load, stalled on the microtubule when approaching the detachment force, and detached from the microtubule at an average force of $4.6 \pm 0.8\,$pN (Figure 4.5), consistent with previous studies (*Khalil et al.*, 2008; *Svoboda and Block*, 1994). In contrast, the CNB mutant detached from the microtubule before stalling (Figure 4.5B) and at much lower loads than WT motors (mean detachment force $0.91 \pm 0.6\,pN$, Figure 4.5A), overall similar to the behavior of the fly kinesin-1 with identical mutations in the CS (*Khalil et al.*, 2008). Motors with N334A mutation (Latch mutant) were also sensitive to small opposing forces exerted by the trap, with mean detachment force of $0.84 \pm 0.4\,$pN (Figure 4.5A). Thus, mutations that hinder either CNB or N-latch formation resulted in motors equally impaired in their ability to displace attached bead under load. The effects of the CNB and Latch mutations were not additive as individual CNB+Latch mutant motors displayed behaviors similar to the CNB and Latch motors—a tendency to detach rather than stall when subjected to load (Figure 4.5B) and detachment from the microtubule at low loads (mean detachment force $0.81 \pm 0.5\,$pN, Figure 4.5A). These results indicate that both CNB formation and N-latch formation are critical for single kinesin-1 motors to generate a strong power stroke and transport processively under load.

Table 4.2: Mean detachment forces of *Rn*KIF5C motors. Force output of wild type (WT), and CNB, LT, and CNB+Latch mutant motors as measured by optical trap apparatus. Standard deviation values in parantheses.

Motor system	Mean detachment force (pN)
WT	4.6 (0.8)
CNB (A5G/S8G)	0.91 (0.6)
Latch (N334A)	0.84 (0.4)
CNB+Latch (A5G/S8G/N334A)	0.81 (0.5)
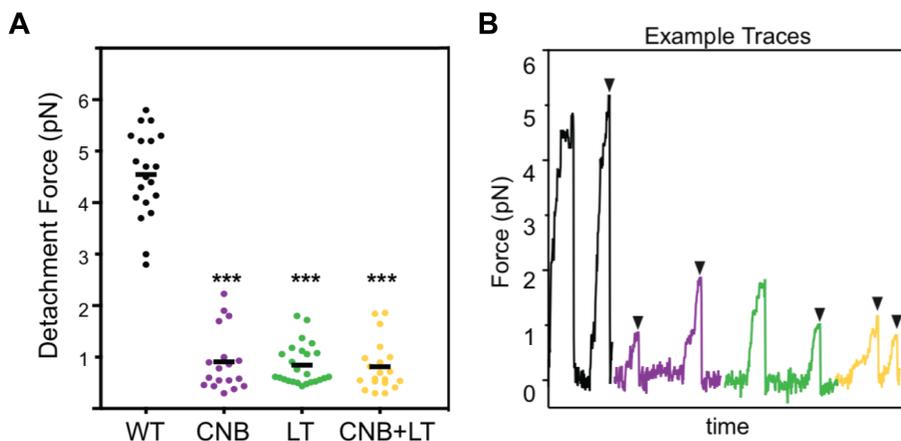
Figure 4.5: Force output of *Rn*KIF5C motors. (**A**) Maximum detachment force of individual events with the mean for each construct indicated by a black horizontal line. Maximum detachment forces include motility events where single motors reached a plateau stall before detachment and events where the motor abruptly detached from the microtubule. $N \geq 20$ events for each construct; *** indicates $P < 0.0001$, compared to the WT. (**B**) Representative traces from optical trap experiments. Arrowheads indicate abrupt detachment events. Figure adapted from (*Budaitis et al.*, 2019, *in final revision at eLife*)

4.3.4 Motility properties of CNB and Latch mutants in unloaded conditions

Our collaborators used single-molecule motility assays to examine the behavior of the CNB and Latch mutants under unloaded conditions. The experimental setup is discussed in our manuscript (*Budaitis et al.*, 2019, *in final revision at eLife*).

Briefly, the single-molecule motility properties of motors "walking" on polymerized microtubules (MTs) were characterized using total internal reflection fluorescence (TIRF) microscopy. In contrast to the crippled force output of mutant motors as discussed above (measured through optical trap assays), all mutant motors were faster and more processive than the WT motor under unloaded conditions. CNB, Latch, and CNB+Latch motors displayed faster velocities of $821 \pm 5 \, \text{nms}^{-1}$, $801 \pm 5 \, \text{nms}^{-1}$, and $778 \pm 3 \, \text{nms}^{-1}$, respectively, compared to $622 \pm 4 \, \text{nms}^{-1}$ for WT motors (Figure 4.6B). The mutant motors also displayed longer run lengths of $1.830 \pm 0.071 \, \mu\text{m}$, $2.380 \pm 0.071 \, \mu\text{m}$, and $3.279 \pm 0.103 \, \mu\text{m}$, respectively, as compared to $0.798 \pm 0.029 \, \mu\text{m}$ for WT motors (Figure 4.6C). Examination of the kymographs indicated an increase in the number of motility events for the mutant motors. Our collaborators, therefore, quantified how often motors landed on a microtubule to start a processive run (apparent landing rate) and measured landing rates of 0.525 ± 0.01, 1.463 ± 0.03, and 2.442 ± 0.6 events/$\mu\text{m}^{-1}\text{nM}^{-1}\text{s}^{-1}$, respectively, compared with WT motor rate of 0.172 ± 0.006 events/$\mu\text{m}^{-1}\text{nM}^{-1}\text{s}^{-1}$ (Figure 4.6D). The enhanced motility properties could be explained by the increased interactions (coupling) between elements important for coordinating and hydrolyzing nucleotide in the CNB and NL mutants as compared to the WT motor. These predictions are discussed in the sections above. However, it was still unclear how mutations that hinder CNB formation and/or NL docking can result in enhanced microtubule binding (landing rate) and processivity of the mutant motors. We discuss this further in the following section.

4.3.5 Principal component analysis of kinesin motor domain

We used principle component analysis (PCA) to create a map of the conformational differences of the microtubule-binding surface of the kinesin-1 motor domain in the microtubule-free (and ADP-bound) state as compared to the microtubule-bound (and ATP-bound) state, to gain an understanding of the observed enhanced micro-
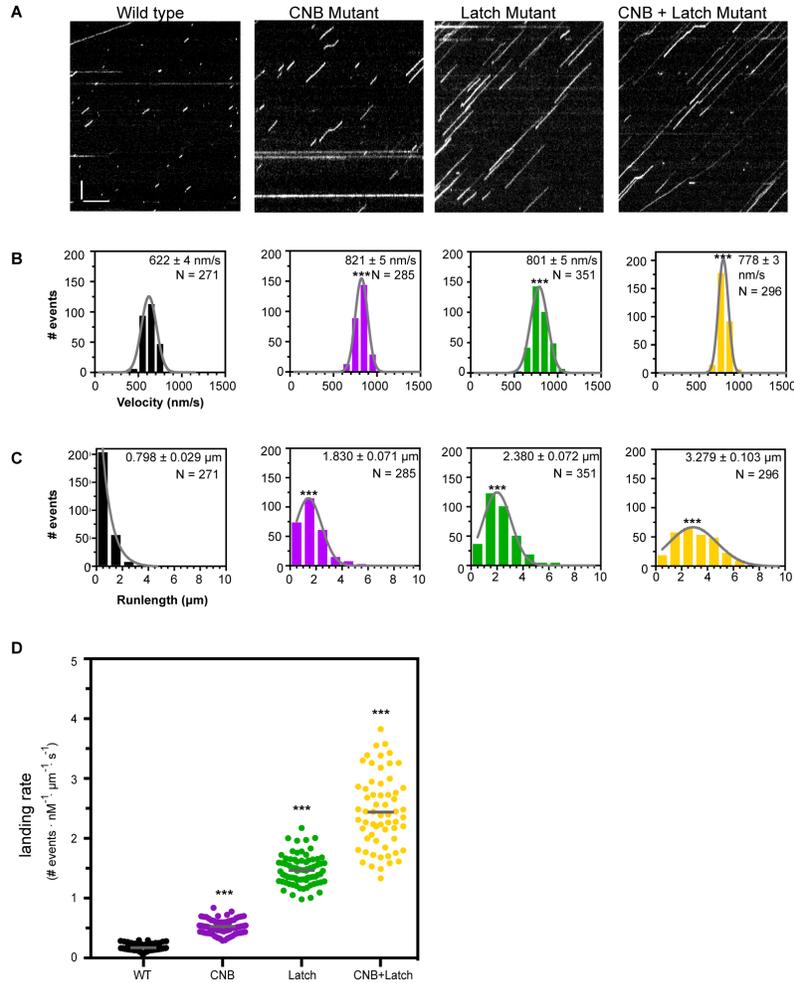
Figure 4.6: Motility properties of CNB and Latch mutants in unloaded conditions. (**A**) Representative kymographs from TIRF microscopy, time displayed on the x-axis (bar 2s) and distance displayed on the y-axis (bar, 2μm). Single-motor (**B**) velocities, (**C**) run lengths, and (**D**) landing rates as determined from the kymographs. Mean±SEM and number of motility events are indicated in upper right corner of each graph; $N \geq 250$ events across three independent experiments for each construct; *** indicates $P < 0.001$ as compared to the WT motor. Figure modified from (*Budaitis et al.,* 2019, *in final revision at eLife*)

tubule binding (landing rate) and processivity of the mutant motors.

The structures of seventeen motor domains from five different kinesins (Table 4.3) were subjected to interconformer analysis with PCA. The CS and NL regions were excluded from the analysis due to their absence from most ADP-bound structures. PCA analysis revealed that the first two dimensions account for over 80% of the vari-

ance in atomic positional displacements of the microtubule-binding surface between these states (PC1 79.66%, PC2 4.95%, Figure 4.7). Thus, PC1 and PC2 provide a suitable conformational space to describe the structural dynamics of kinesin motor domains transitioning from an ADP-bound, microtubule-free state to an ATP-bound, microtubule-bound state. The major conformational difference between these states can be described by PC1 which involves a displacement of $\alpha 4$, where $\alpha 4$ is in a "down" orientation in the ADP-like, microtubule-free structures and in an "up" orientation in the ATP-like, microtubule-bound structures (Figure 4.7), consistent with previous studies (*Scarabelli and Grant*, 2013).

Table 4.3: Kinesin experimental structures for PCA analysis. The PDB IDs of the kinesin structures analyzed with PCA are listed in the first column. The nucleotide state, presence of tubulin, neck linker conformation (A=absent, PD=partially docked, FD=fully docked), and family is reported for each structure.

PDB ID	Nucleotide	Neck linker	Tubulin heterodimer	Family (isoform)	Reference
1BG2 (A)	ADP	A	No	kin-01 (KIF5B)	Kull et al. Nature (1996)
4A14 (A)	ADP	A	No	kin-04 (KIF7)	Klejnot et al. Acta Crystallogr.,Sect.D (2012)
2XT3 (A)	ADP	A	No	kin-04 (KIF7)	Klejnot et al. Acta Crystallogr.,Sect.D (2012)
1CZ7 (A)	ADP	A	No	kin-14 (ncd)	Kozielski et al. Structure Fold.Des. (1999)
1CZ7 (B)	ADP	A	No	kin-14 (ncd)	Kozielski et al. Structure Fold.Des. (1999)
1CZ7 (C)	ADP	A	No	kin-14 (ncd)	Kozielski et al. Structure Fold.Des. (1999)
1CZ7 (D)	ADP	A	No	kin-14 (ncd)	Kozielski et al. Structure Fold.Des. (1999)
2NCD (A)	ADP	A	No	kin-14 (ncd)	Sablin et al. Nature (1998)
3L1C (A)	ADP	A	No	kin-14 (ncd)	Heuston et al. Bmc Struct.Biol. (2010)
4UXR (C)	ADP-Pi (ADP-AlF$_4^-$)	PD	Yes	kin-03 (KIF1A)	Atherton et al. Elife (2014)
4UXP (C)	ATP (ANP)	PD	Yes	kin-03 (KIF1A)	Atherton et al. Elife (2014)
4UXY (C)	ATP (ANP)	FD	Yes	kin-01 (KIF5A)	Atherton et al. Elife (2014)

PDB ID (isoform)	Nucleotide Reference	Neck linker	Tubulin heterodimer	Family	
4UY0 (C)	ADP-Pi (ADP-AlF$_4^-$)	FD	Yes	kin-01 (KIF5A)	Atherton et al. Elife (2014)
3J8Y (K)	ATP (ATP)	PD	Yes	kin-01 (KIF5B)	Shang et al. Elife (2014)
4HNA (K)	ADP-Pi (ADP-AlF$_4^-$)	FD	Yes	kin-01 (KIF5B)	Gigant et al. Nat.Struct.Mol.Biol. (2013)
5MIO (C)	ATP (ANP)	A	Yes	kin-13 (KIF2C)	Wang et al. Nat Commun (2017)
5ND4 (C)	ADP-Pi (ADP-AlF$_4^-$)	PD	Yes	kin-06 (Kif20a)	Atherton et al. Elife (2017)

MD simulations of WT and CNB+Latch mutant motors, carried out in replicate for a total of $1\,\mu$s each in the same manner as described in the *Methods* section, were then projected onto this PCA conformational space to compare how often the WT and CNB+Latch motor domains could adopt the ATP-bound, microtubule-bound state from the ADP-bound, microtubule-free state. Initial starting conformation was from the PDB 2KIN structure which partially adopts an ATP-like conformation (docked NL but weak microtubule binding (*Parke et al.*, 2010a)), the WT motor sampled a conformational space between the ADP-bound, microtubule-free and the ATP-bound, microtubule-bound states (Figure 4.7, blue contours). The CNB+Latch mutant, initiated from the same PDB 2KIN starting structure, sampled an additional conformational space closer to that defined by the ATP-bound, microtubule-bound kinesin structures (Figure 4.7, red contours). This suggests that the CNB+Latch mutant has a higher degree of structural flexibility in its microtubule-binding regions as compared to the WT motor domain, enabling it to more readily sample a conformation compatible with strong microtubule binding in response to ATP in the nucleotide pocket and could account for the enhanced microtubule-landing rate and processivity observed in single-molecule assays.
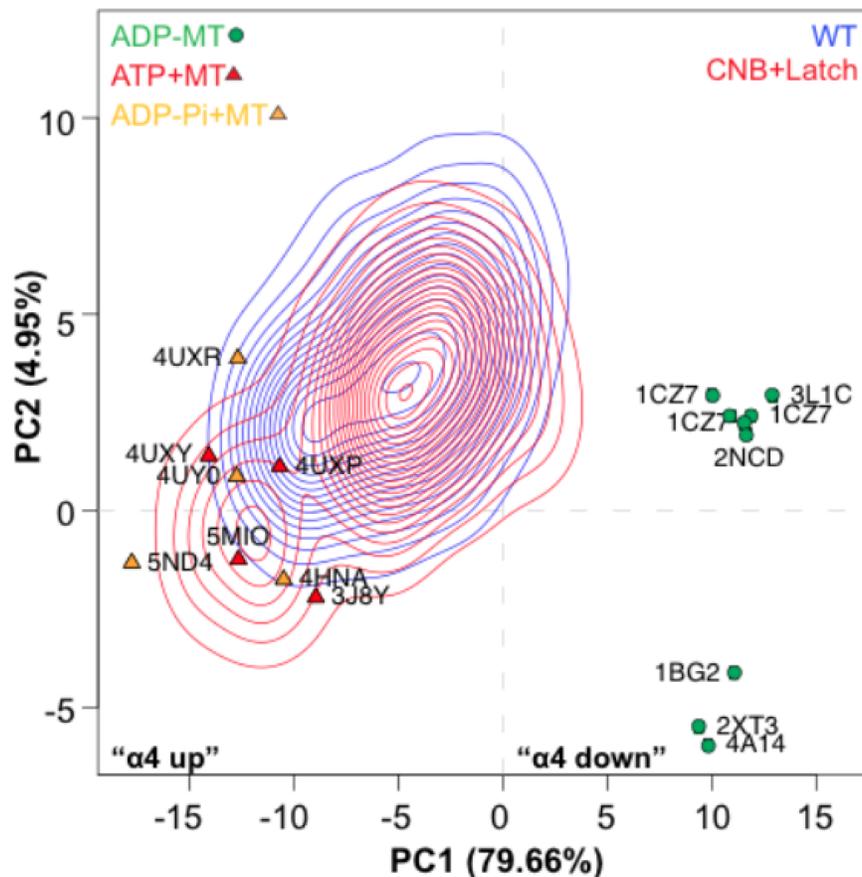
Figure 4.7: Principal component analysis of kinesin experimental structures and CNB+Latch mutant. Principal component sub-space (PC space) of the microtubule-binding surface of kinesin-1. See Table 4.3 for experimental structures utilized in defining the PC space. The position of each motor domain structure within the PCA map is indicated together with its nucleotide state (red, ATP; yellow, ADP-Pi; green, ADP), microtubule state (circle, no microtubule; triangle, bound to microtubule), and PDB code. PC1 represents the positioning of helix $\alpha 4$ as "down" in the ADP-like state and "up" in the ATP-like state. The ability of WT versus CNB+Latch mutant motors to sample these states was then analyzed by projecting trajectories from MD simulations starting from the 2KIN structure in the ADP-bound and microtubule-free state. The conformational space explored by each motor in the MD simulations is projected onto the PC space and displayed as contours (WT, blue; CNB+Latch, red).

4.3.6 Low- and high-load cargo transport in cells

Next, we proceeded to test whether, physiologically, the integrity of CNB forma-tion followed by NL latching is a critical determinant for kinesin motors working in teams to drive cargo transport in cells. Our collaborators performed inducible trans-

port assays to measure transport of low- and high-load cargo in COS7 cells. Please see (*Kapitein et al.*, 2010) and our manuscript (*Budaitis et al., 2019, in final revision at eLife*) for details regarding experimental setup. Briefly, the efficiency of transporting cargo within cells is measured by single-particle tracking of labelled cargo and motors. In these inducible assays, peroxisome are considered a low-load cargo, requiring teams of kinesin-1 motors to collectively transport against loads < 3 times greater than the force required to stall a single motor. Cargo dispersion is measured in the following two ways: qualitatively as clustered, partially dispersed, diffusely dispersed, or peripherally dispersed; and quantitatively as normalized distance distribution from nucleus, measured as radial intensity of labelled cargo across cells in all samples. Teams of WT kinesin-1 motors were able to transport peroxisomes to the cell periphery (quantitatively 81% of the peroxisome intensity at the cell periphery, Figure 4.8C). Although mutant motors are crippled in their ability to transport against load as single motors in an optical trap (Figure 4.5A), teams of CNB, Latch, or CNB+Latch mutant motors were able to disperse peroxisomes to the periphery of the cell (Figure 4.8B; and Figure 4.8C-E, quantitatively 84%, 81%, and 79% of the peroxisome intensity at the cell periphery, respectively). Statistical analysis indicates that peroxisome dispersion by the mutant motors was not significantly different than that of the WT motor (Figure 4.8C-E). Notable, these results suggest that impaired force generation by weakening CNB and/or N-latch formation can be overcome by teams of motors for efficient transport of low-load cargo in cells.

To address how motors cooperate in teams to transport high-load cargo in cells, our collaborators employed inducible dispersion assay as decribed above for measuring the transport of Golgi by kinesin-1 motors within cells. The Golgi is a compact organelle and its localization near the nucleus is maintained by a variety of mechanisms including microtubule minus-end directed activity of cytoplasmic dynein motors (*Brownhill et al.*, 2009). These mechanisms, including the additional resistence pro-
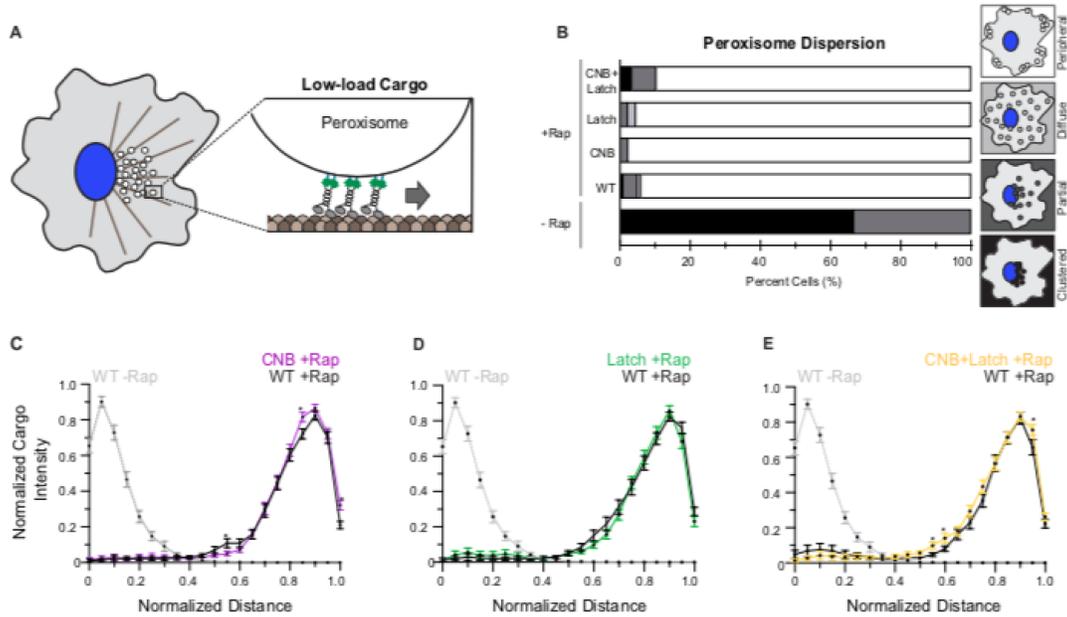
Figure 4.8: Low-load cargo transport by teams of kinesin-1 motors in cells. (**A**) Schematic of the inducible peroxisome dispersion assay utilized by our collaborators. Peroxisomes loosely clustered in the perinuclear region of COS7 cells serve as low-load cargo for transport by teams of motors recruited by addition of rapamycin (+Rap). (**B**) Qualitative analysis of peroxisome dispersion. Peroxisome localization in individual cells scored as clustered (black), partially dispersed (dark grey), diffusely dispersed (light grey), or peripherally dispersed (white) 30 minutes after recruitment of teams of WT, CNB, Latch, or CNB+Latch motors. Data summarized for $N \geq 50$ cells across three experiments. (**C-E**) Quantitative analysis of peroxisome dispersion. A radial profile of peroxisome intensity for each cell converted to an averaged and normalized distance distribution across all cells. Each data point indicates the mean normalized cargo intensity±SEM for $N \geq 50$ cells across three separate experiments. Gray dotted line: WT -Rap; Black line: WT +Rap; (*C*) Purple line: CNB +Rap; (*D*) Green line: Latch +Rap; (*E*) Yellow line: CNB+Latch +Rap; * indicates $P < 0.05$, as compared to the mean normalized cargo intensity of WT motors for each distance. Figure modified from (*Budaitis et al.*, 2019, *in final revision at eLife*).

vided by the minus-end directed pull of dynein motors in a direction against that of kinesin-1 motors, make Golgi a high-load cargo since teams of motors driving Golgi dispersion are required to cooperate to transport against forces ~ 30 times greater than the force required to stall a single kinesin-1 motor. Teams of WT kinesin-1 motors were able to transport Golgi to the cell periphery (Figure 4.9B, qualitatively 82% of cells have dispersed Golgi; Figure 4.9C, 50% of Golgi intensity at the cell

periphery). However, a significant fraction of Golgi elements remained clustered in the perinuclear region rather than accumulated at the cell periphery when teams of CNB or Latch mutants were recruited (Figure 4.9B, qualitatively only 64% of cells have dispersed Golgi for CNB and 42% for Latch mutant; Figure 4.9C, quantitatively only 34% of Golgi intensity at the cell periphery for CNB and Figure 4.9D, only 35% for Latch mutant). These results highlight the reduced capacity of motors with mutations that hinder either CNB or N-latch formation to transporting high-load cargo in cells.

Furthermore, the effects of the CNB and N-latch mutations were additive as teams of CNB+Latch mutants were even more impaired in their capacity to transport Golgi elements than the CNB and Latch mutant motors. Upon recruitment of CNB+Latch mutant motors, the majority of the Golgi elements remained clustered in the perinuclear region of the cell (Figure 4.9B, qualitatively only 13% of cells have dispersed Golgi; Figure 4.9E, quantitatively only 22% of Golgi intensity at the cell periphery). Collectively, these results suggest that while kinesin motors with impaired CNB formation and NL can cooperate for transport of a low-load, they are unable to work in teams when faced with a high-load cargo in cells, highlighting the critical physiological role of these structutral elements in kinesin-1 motor function.

4.4 Methods

4.4.1 Molecular modeling of kinesin-1 in complex with tubulin

The rat isoform of kinesin-1, *Rn*KIF5C, was selected for molecular modeling for use in molecular dynamics simulations as well as for concistency with later experimental assays. Initial models of *Rn*KIF5C motor domain in complex with tubulin were obtained from PDB 4LNU for the no nucleotide (apo) state (*Cao et al.*, 2014), and from PDB 4HNA for ATP-bound state (*Gigant et al.*, 2013). Since the motor domain
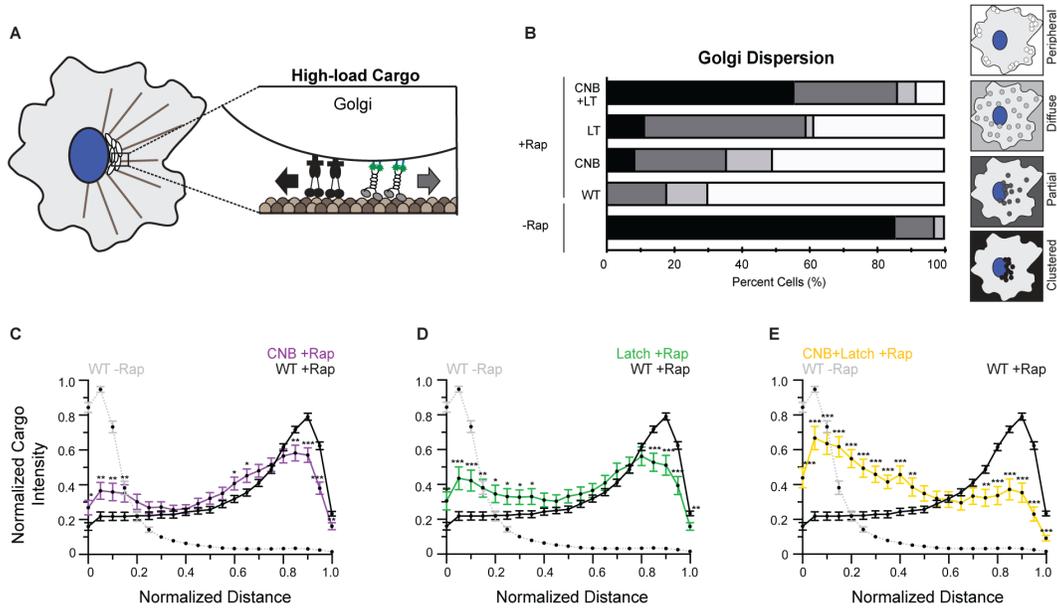
Figure 4.9: High-load cargo transport by teams of kinesin-1 motors in cells. (**A** Schematic of the inducible Golgi dispersion assay utilized by our collaborators. Golgi tighly clustered in near the nucleus serve as high-load cargo for transport by teams of motors recruited by addition of rapamycin (+Rap). Cytoplasmic dynein motors, moving in direction opposite to that of kinesin-1 motors, shown in black. (**B**) Qualitative analysis of Golgi dispersion. Golgi localization in individual cells scored as clustered (black), partially dispersed (dark grey), diffusely dispersed (light grey), or peripherally dispersed (white) 30 minutes after recruitment of teams of WT, CNB, Latch, or CNB+Latch motors. Data summarized for $N \geq 30$ cells across three experiments. (**C-E**) Quantitative analysis of Golgi dispersion. A radial profile of Golgi intensity for each cell converted to an averaged and normalized distance distribution across all cells. Each data point indicates the mean normalized cargo intensity±SEM for $N \geq 30$ cells across three separate experiments. Gray dotted line: WT -Rap; Black line: WT +Rap; (*C*) Purple line: CNB +Rap; (*D*) Green line: Latch +Rap; (*E*) Yellow line: CNB+Latch +Rap; * indicates $P < 0.05$; ** $P < 0.01$; and *** $P < 0.001$, as compared to the mean normalized cargo intensity of WT motors for each distance. Figure from (*Budaitis et al.*, 2019, *in final revision at eLife*).

in both template structures (PDBs 4LNU and 4HNA) is KIF5B, residues that differ were mutated to match the sequence of rat KIF5C (UniprotID: P56536). The tubulin dimer was left unmodified. Missing coordinates were modeled using MODELLER v9.18 (*Šali and Blundell*, 1993). The ATP-hydrolysis transition-state analog, ADP-AlF$_4^-$, in PDB 4HNA was converted to ATP. The resulting systems of motor domain

associated with tubulin dimer contained a total of ~170,000 atoms each. Models of ADP-bound wildtype and CNB+Latch mutant motor domains (not associated with the tubulin heterodimer) were prepared from PDB 2KIN (*Sack et al.*, 1997).

4.4.2 Molecular dynamics simulations

Energy minimization and molecular dynamics simulations were performed with AMBER14 (*Case et al.*, 2014) and the ff99SB AMBER force field (*Hornak et al.*, 2006). Nucleotide parameters were obtained from (*Meagher et al.*, 2003). Histidine protonation states were assigned based on the their pKa values calculated by PROPKA (*Li et al.*, 2005). Starting structures were solvated in a cubic box of pre-equilibrated TIP3P waters molecules, extending $12\,\text{Å}$ in each dimension from the surface of the solute. Sodium ions (Na^+) were added to neutralize the systems, followed by addition of sodium and chloride (Cl^-) ions to bring the ionic strength to $0.050\,\text{M}$. Energy minimization was performed in four stages, with each stage consisting of 500 steps of steepest descent followed by 4000 steps of conjugate gradient. First, minimization of solvent was performed by keeping positions of protein and nucleotides fixed. Second, side-chains and nucleotides were relaxed keeping the backbone positions fixed. Third, protein and nucleotide atoms were relaxed while keeping the solvent atoms fixed. Fourth, a last minimization stage was performed with no restraints. The system was gradually heated to 300K over $25\,\text{ps}$ of simulation time in constant-volume (NVT) and periodic boundary conditions (PBC), with restraint of $20\,\text{kcal/mol/Å}^2$ on backbone atoms. Constant-temperature (T= 300K) and constant-pressure (p= 1 bar) (NpT) equilibration was then performed in six stages. First, a $400\,\text{ps}$ NpT equilibration was performed with restraint of $20\,\text{kcal/mol/Å}^2$ on backbone atoms. Further stages involved gradually reducing restraints of 20, 10, 5, and $1\,\text{kcal/mol/Å}^2$ on α carbons over $5\,\text{ns}$ each. A final NpT equilibration was carried out without any restraints for $5\,\text{ns}$. Subsequent production phase molecular

dynamics simulations were then performed under NpT and PBC with random velocity assignments for each run. Particle-mesh Ewald summation was adopted for treating long-range electrostatics. A 12 Å cutoff for energy minimization, and a 10 Å cutoff for molecular dynamics simulations was used to truncate non-bonded interactions. A 2 fs time-step was adopted for all molecular dynamics simulations. Hydrogen atoms were constrained using the SHAKE algorithm. All simulations were performed in-house on NVIDIA GPU cards with the GPU version of PMEMD (*pmemd.cuda*). Molecular dynamics simulations were started from equilibrated structures with four independent runs of 100-300 ns each. Trajectory analysis was carried out in R using the Bio3D v2.3-3 package (*Grant et al.*, 2006; *Skjærven et al.*, 2014).

4.4.3 Residue-residue distance differences

Statistically significant residue-residue distance differences between apo, ATP-bound and mutant states were identified with ensemble difference distance matrix (eDDM) analysis routine (*Muretta et al.*, 2018). For this analysis, a total of 400 conformations were obtained for each state under comparison by extracting 100 equally time-spaced conformations from the last 20 ns of each simulation replicate. The details of obtaining the distance matrices from simulation trajectories, their processing, and the method of selecting significantly different residue-residue distances are discussed in Chapter III (see its *Methods* section). Briefly, the eDDM routine reduces the difference between long distances while difference between short distances are kept intact. The significance of residue distance variation between apo and ATP-bound states, and between ATP-bound and mutant states, were evaluated with the Wilcoxon test. Residue pairs showing a p-value$< 10^{-5}$ and an average masked distance difference > 1Å were considered statistically significant residue-residue distance differences for further analysis.

4.4.4 Principal component analysis

A set of 17 experimental structures from the RCSB protein data bank, nine in ADP-like state not associated with the microtubule and eight in ATP-like state bound to tubulin heterodimer, were selected for examining the major conformational differences of the kinesin motor domain in these two states (See Table 4.3). Principal component analysis (PCA) is a dimensionality reduction technique involving orthogonal transformation of the original data into a set of linearly uncorrelated variables termed principal components. Briefly, PCA involves diagonalization of the covariance matrix C, whose elements C_{ij} are calculated from the Cartesian coordinates of Cα atoms, r, after superposition:

$$C_{ij} = \langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \rangle \tag{4.1}$$

where i and j represent all pairs of 3N coordinates. The eigenvectors, or principal components (PCs), of the covariance matrix form a linear basis set of the distribution of structures. The variance of the distribution along each eigenvector is given by the corresponding eigenvalue. Projecting structures onto a sub-space defined by principal components with the largest variance (largest eigenvalues) provides a lower dimensional representation of the structure dataset.

PCA was performed on 112 equivalent, non-gap Cα atoms from each of the structures after superposition onto an invariant core comprising of structural elements $\beta 1$, $\beta 2$, $\beta 3$, P-loop, $\alpha 2$, $\beta 6$, $\beta 7$, and $\alpha 6$ (*Scarabelli and Grant*, 2013). The trajectories from MD simulations of ADP-bound wildtype and CNB+Latch mutant kinesin motor domains were projected on to the PC sub-space defined by the first two PCA eigenvectors to allow comparison of the conformational space spanned by the simulations and the experimental structures (Figure 4.7).

For rest of the methods pertaining to experimental results, please see our publi-

cation (*Budaitis et al.*, 2019) (*in final revision at eLife*).

4.5 Discussion and Conclusion

Structural studies of the nucleotide-dependent conformational changes in kinesin as well as studies at the single-molecule level have led to a model that CNB formation, and the subsequent "power-stroke", is the force-generating element for kinesin motors. Here we use molecular dynamics simulations, in combination with optical trapping studies and single-molecule assays carried out by our collaborators, to show that both CNB and N-latch formation are critical for single kinesin-1 motors to transport against force.

Analysis of the structure and dynamics of kinesin-1 in the nucleotide-free (apo, starting from PDB 4LNU (*Cao et al.*, 2014)) and ATP-bound, post-power stroke (stating from PDB 4HNA (*Gigant et al.*, 2013)) states highlighted residues critical for CNB formation (β9 with CS) and NL latching (N-latch and β10 with β7) (Figure 4.1). Single-molecule assays showed that disrupting CNB or N-latch formation did not impair the ability of mutant motors to undergo processive motility, indicating that the mutations are tolerated by the motor when stepping under no load. On the contrary, these mutations resulted in enhanced motilty properties under single-molecule, unloaded conditions (Figure 4.6). We attribute the enhanced motility properties to allosteric effects on the nucleotide coordinating and microtubule-binding elements of the motor domain. First, the allosteric effects of NL docking on core motor regions that coordinate and bind nucleotide (α0, S1, S2, PL; Figure 4.4D-F and H-I) could result in enhanced catalytic site closure that would favor ATP hydrolysis in the mutant motors. Our findings are consistent with previous structural and enzymatic studies suggesting that NL docking allosterically gates ATPase activity (*Hahlen et al.*, 2006; *Cao et al.*, 2014). Our findings are also consistent with previous time-resolved TR^{2}FRET studies of kinesin-1 and kinesin-5 motors demonstrating that NL docking

is allosterically coupled to active site closure (See *Chapter III* for results pertaining kinesin-5 mitotic motor Eg5) (*Muretta et al.*, 2018, 2015).

Second, the allosteric effects on the microtubule-binding surface of kinesin-1 could explain the enhanced processivity of CNB, Latch, and CNB+Latch mutant motors (Figures 4.7 and 4.6C). PCA analysis demonstrated that the major structural difference between motors in the ADP-bound and microtubule-free state and motors in the ATP- bound and microtubule-bound state is the positioning of $\alpha 4$ (Figure 4.7). MD simulations of CNB+Latch mutant motors indicated an increased ability of this motor's microtubule-binding surface to sample a conformation compatible with strong microtubule binding (Figure 4.7, $\alpha 4$ up). Collectively, these results highlight how perturbing the CNB forming and NL docking elements (CS, $\beta 9$, N-latch/$\beta 10$) can act as a molecular gearshift, a mechanism applied across kinesin motor families to modulate motor speed and processivity.

The enhanced motility properties under single-molecule, unload conditions, however, do not correlate with force production by single motors under optical trapping experiments (Figure 4.5). Disruption of either CNB or N-latch formation resulted in motors unable to stall under load and and were more likely to detach when subjected to low forces (average detachment force $\sim 1\,\mathrm{pN}$, Figure 4.5). We note that mutation of equivalent residues of CNB mutant (A5G/S8G) in fly kinesin-1 motor (*Dm*KHC) resulted in detachment for of $\sim 3\,\mathrm{pN}$ (*Khalil et al.*, 2008) (compared to $\sim 1\,\mathrm{pN}$ in *Rn*KIF5C in this study). This highlights an additional role of CS length (9 residues long in *Rn*KIF5C versus 13 in *Dm*KHC) on motor force output.

Our collaborator's inducible cargo assays suggest that both CNB formation and NL latching are essential for teams of kinesin-1 motors to collectively transport high-load cargo under physiological conditions (Figure 4.9). However, teams of motors carrying the same mutations were able to transport low-load cargo (peroxisomes) with the same efficiency as that of WT motors. This suggests other compensatory mecha-

nisms, such as enhanced motility properites as seen under single-molecule conditions (Figure 4.6), are sufficient for motor function under low load, but their cripped force output is insufficient for transport of high-load cargo in cell. Teams of CNB+Latch mutant motors were more impaired in their ability to drive Golgi dispersion as compared to the CNB and N-Latch mutants (Figure 4.9). This result differs from optical trapping experiments, which show that the force output of all three mutant motors was equally disrupted (Figure 4.5A). These results highlight the need for characterizing motility properties of kinesin motors with physiologically relevant cell-based assays, in addition to force output measurements. We have shown in Chapter III that acetylation of lysine residue 146 (helix α2b) disrupts its ionic interaction with residue D91 (helix α1) in kinesin-5 mitotic motor Eg5, confering enhanced motility properties. However, under physiological conditions, modified Eg5 motors (carrying acetylation mimetic mutation K146Q) act as a mitotic "brake", slowing centrosome separation by generating drag force against unmodified Eg5 motors (see *Chapter III* and (*Muretta et al.*, 2018)).

Complementing previous computational and experimental studies (*Rice et al.*, 1999b; *Khalil et al.*, 2008; *Hwang et al.*, 2008; *Hesse et al.*, 2013), we have shown that CNB and N-latch formation serve as mechanical elements for force generaion in kinesin-1. Our collaborators also provide first evidence, to our knowledge, that a power stroke mechanism and force generation are critical for multi-motor driven transport under physiological conditions. Further studies are required to explore the role of CNB and N-latch formation in force generation in other kinesin families, including the role of CS length in family-specific force and motility properties.

Dynamical effects of mutations in Kinesin-3 KIF1A

5.1 Abstract

KIF1A is a molecular motor of the kinesin-3 family of motor proteins, with primary function of anterograde transport of synaptic vesicle precursors along axonal microtubules. Recently, a number of disease-associated genetic variants and de novo mutations have been identified from clinical studies. These mutations have been linked to neurodevelopmental disorders including cognitive disability, spasticity, and cerebellar and optic nerve atrophy. The mechanism of how these mutations disrupt motor function is not understood. We used molecular dynamics simulations to characterize dynamical effects of mutations localized to the motor domain of KIF1A. We focused on the V8M mutation, which is localized on the first beta strand ($\beta 1$), and hypothesized that this V8M mutation would affect the velocity of KIF1A. We observed significant changes in residue-residue interactions for three functionally important domains. First, V8M impacted residues involved the power stroke i.e., neck linker (NL), cover strand (CS), $\beta 7$, Loop 13, and suggest an unexpected output of enhanced NL docking. Second, differences in interactions between nucleotide coordinating residues in $\alpha 0$, P loop, Loop 9/Switch 1, Loop 11/Switch 2, indicate an impact on ATPase activity. Third, decreased interactions for residues interacting with the microtubule in $\alpha 4$, Loop 11/Switch 2, Loop 13, $\alpha 6$ suggest the V8M motor is slower and less pro-

cessive, consistent with previous experimental work (*Scarabelli et al.*, 2015). Current efforts are focused on characterizing the dynamics of additional mutations—V8A ($\beta 1$), Y89D ($\alpha 1/\beta 3$) and E146D/A ($\beta 4$)—as well as experimentally measuring their kinetic and motility properties. Our results demonstrate the allosteric effects of KIF1A neurological disease-associated mutations and help us better understand how mutations in kinesin motors can lead to neurodevelopmental disorders.

5.2   Introduction

KIF1A is part of an important class of neuronal plus-end directed transport motors, primarily transporting synaptic vesicle precursors in cells (*Hirokawa et al.*, 2009). While certain isoforms of kinesin-1 are also neuron-specific transport motors (such as KIF5A and KIF5C), studies have shown that it is the kinesin-3 family of motor proteins that are involved in long-range transport, with strikingly different motility properties—kinesin-3 motors are called "superprocessive", with 10-fold more processive than kinesin-1 motors (*Soppina et al.*, 2014).

Recently, a number of disease-associated genetic variants and *de novo* mutations have been identified from clinical studies (*Lee et al.*, 2015; *Esmaeeli Nieh et al.*, 2015; *Hotchkiss et al.*, 2016; *Iqbal et al.*, 2017). These mutations have been linked to neurodevelopmental disorders, typically with a progressive course. The disease, collectively termed KIF1A Associated Neurological Disorder (KAND), is associated with cognitive disability, spasticity, and cerebellar and optic nerve atrophy (*Okamoto et al.*, 2014; *Ohba et al.*, 2015). The mechanism of how these mutations disrupt motor function is still an area of active research.

The diease-associated genetic variants and *de novo* mutations identified through clinical sequencing studies span the entirety of the KIF1A gene. The variants and mutations occuring in the core motor domain are shown linearly and also mapped onto the kinesin structure in Figure 5.5. In this study, we focus on variants and

mutations occurring in regions not directly interacting with the microtubule (MT), or involved directly in coordinating/hydrolyzing ATP. The effect of mutations occurring in these two important regions is relatively straightforward to predict: they either disrupt MT-binding, or cripple the ability of kinesin motor to bind/hydrolyze ATP for processive runs. We hypothesized that the allosteric mutants indirectly affect kinesin motor function: by affecting these regions critical for kinesin mechanochemistry and MT-binding. Here, we combine molecular dynamics simulations, residue distances analyses, and single-molecule motility assays to characterize the allosteric effects of allosteric mutatiosn in KAND mutations on kinesin-3 function and regulation. The preliminary results presented here are our first set of analyses in this study, with further investigations underway. Our results show that these mutants have dynamical effects on functionally important elements for kinesin motility: elements involved in NL docking (NL, CS, $\beta 7$, L13), and elements coordinating or hydrolyzing the nucleotide ($\alpha 0$, PL, L9/S1, L11/S2).

5.3 Results

5.3.1 Sequence and structure mapping of KIF1A variants

Disease-associated variants and mutations spanning the motor domain of KIF1A are shown in Figure 5.1. The mutations span the enterity of the motor domain, with seemingly no obvious localization on helices, sheets, or loops. However, coloring the mutations mapped onto the KIF1A structure by their known involvement in either microtubule (MT)-binding or their role in coordinating/hydrolyzing ATP reveals a function-associated clustering (Figure 5.1C,D). MT-interacting residues with mapped mutations include R167 (Loop 8), R307 (Loop 12), R316 (helix $\alpha 5$), and R350 (helix $\alpha 6$) (shown as circles or spheres in cyan in Figure 5.1A,C,D). The importance of these residues in kinesin motor function have been tested through mutagenesis of either

these residue directly or their interacting parter-residues in MTs (*Scarabelli et al.*, 2015; *Uchimura et al.*, 2010). Almost all of the MT-binding associated mutations result in a loss of charge, predicted to disrupt ionic interactions with the MT and hence affecting motor processivity. Mutations ocurring in residues associated with ATP binding or hydrolysis are predicted to affect function by disrupting nucleotide capture and/or hydrolysis (show as circles or spheres in red in Figure 5.1A,C,D).

Mutations not associated with residues involved in either of these critical elements also span the motor domain (shown as uncolored circles or spheres Figure 5.1A,C,D). We hypothesize these mutations have an allosteric effect on the MT-binding or ATP-coordinating/hydrolyzing regions of the motor domain. We tested the effect of mutating three of such residues on kinesin dynamics: V8M in $\beta 1$, Y89 in $\alpha 1/\beta 3$, and E148 in $\beta 4$. Molecular dynamics (MD) simulations of mutantions tested are listed in Figure 5.1B. A valine residue at position 8 in KIF1A is situated in $\beta 1$), in continuation from the N-terminal coverstrand (CS). Residue V8 is highly conserved across the kinesin superfamily, with a valine at this position in most families except for kinesin-10 (*Richard et al.*, 2016). However, the role of a valine residue at this position in $\beta 1$ in kinesin function or mechanochemistry has not been studied in pan-family manner. A tyrosine residue at position 89 at the intersection of helix $\alpha 1$ and strand $\beta 3$, Y89, forms non-polar interactions with a patch of hydrophobic residues, 358-360, in the NL. We hypothesize that gain-of-charge mutation Y89D will disrupt hydrophobic interactions between $\alpha 1/\beta 3$ and the NL, causing NL to undock from the kinesin motor core domain in the ATP-state.

5.3.2 Allosteric mutant V8M predicted to affect ATPase activity

The global and local impact of V8M mutation on KIF1A motor domain dynamics was characterized by residue-residue distance analysis (see *Methods* section). Significant differences ($P < 10^{-5}$ in residue-residue distances between wildtype (WT) and

Figure 5.1: Sequence and structure mapping of KIF1A variants. (**A**) Variants associated with neurodevelopmental disorders span the KIF1A motor domain. (**B**) Mutations tested in this study localized to $\beta 1$, $\alpha 1/\beta 3$, and $\beta 4$. (**C** and **D**) Mutations tested mapped onto structure of KIF1A motor domain. Colors: microtubule-binding (cyan), nucleotide-coordinating (red), allosteric or other effects (pink). Nucleotide and divalent cation shown as spacefill. Tubulin dimer not shown for clarity.

V8M mutant simulations are highlighted in Figure 5.2A,B. The V8M mutation in $\beta 1$ disrupts the coupling (longer residue-residue distance in V8M, shorter in WT) between P loop (PL), switch 1 (S1), and switch 2 (S2), elements critical for nucleotide binding and hydrolysis (blue squares marked S1-PL and S2-S1 in Figure 5.2). For example Q98 (PL) - R254 (Loop 11/S2) distance increases from 4.00 ± 0.94Å in WT to 7.54 ± 0.60Å in V8M; and the R216 (Loop 9/S1) – N270 (Loop 11/S2) distance increases from 3.45 ± 0.89Å in WT to 7.19 ± 0.89Å in V8M. All pairs of significantly differetnt inter-residue distances for the WT and V8M mutant are listed in Appendix D, Table D.1.

Figure 5.2: V8M (β1) mutant allosterically impacts ATPase activity. (**A**) Differences in residue-residue distances between WT KIF1A and V8M mutant motors in ATP- and microtubule-bound states as determined from MD simulations. Secondary structures elements: α helices (black), β strands (grey). Interactions that are significantly closer ($P < 10^{-5}$) in the WT (blue) or V8M mutant (red) are indicated. The magnitude of the distance change is indicated by size and color intensity. (**B**) Cartoon representation of residue distances changes. Blue lines depict residue-residue distances that are shorter in the WT motor versus V8M mutant motor ($P < 10^{-5}$). The magnitude of the distance change is indicated by line color intensity. (**C**, **D**, and **E**) Motility properties as measured from kymographs under single-molecule, unloaded conditions. The V8M mutant motors are slower (*C*) and less likely to engage with microtubule to start a processive run (*E*). Data colored as: KIF1A WT (blue), V8M mutant (red), with *Rn*Kif5C (gray).

5.3.3 V8M mutant motors are slower in motility assays

We predict the allosteric impact of V8M mutation on the PL and switch regions, indicated by increased residue-residue distances between S1-PL and S2-S1, will impair

the ATPase activity of the motor domain. To measure the impact of V8M mutation of kinesin motility properties, our collaborators[1] characterized velocity and processivity of WT and V8M mutant KIF1A motor under single-molecule, unloaded conditions. The experimental results show that indeed the V8M mutant motors are slower than the WT (969 ± 8.2 nm·s^{-1} for V8M mutant motor versus 1687 ± 11.2 nm·s^{-1} for WT motor) (Figure 5.2C). No appreciable differences in processivity, as measured from kymographs under single-molecule, unloaded conditions, were noted when comparing V8M mutant to WT motor (Figure 5.2D). We note that results pertaining to processivity may be a function of MT length due to the superprocessive nature of KIF1A motors under *in vitro* conditions. Our collaborators also noted that the V8M mutant motors exhibited higher number of non-processive and diffuse events (marked by cyan arrows in Figure 5.3). This is also reflected in the lower apparent on-rates, quantified from kymographs, of V8M mutants, indicating that V8M mutant motors are less likely to engage with microtubule to start a processive run (Figure 5.2E).

5.3.4   Y89D mutation predicted to affect force generation

A second KAND-related mutation we analyzed in this study was a Y89D mutation. A large aromatic side-chain at this position appears to be a conserved feature of most N-terminally situated, plus-end directed kinesin motors. The most frequently observed residue at this position is a tyrosine followed by phenylalanine, with a notable exception of a gluatmine residue in kinesin-6 motors KIF20A/B. In *Hs*KIF1A, tyrosine 89, situated between helix $\alpha 1$ and strand $\beta 3$, makes hydrophobic interactions with a residues (A358, V359, and I360) in the NL (Figure 5.4B). The disease-associated mutation observed in clinical studies, Y89D, introduces a negative-charge. We predict the change of hydrophobic to polar charged residue will disrupt the hydrophobic interactions with residues $358 - 360$ in the NL. We hypothesize that this

---

[1]Experimental work, shown here as well as under progress, was carried out by Breane G. Budaitis in Kristen Verhey laboratory at the University of Michigan
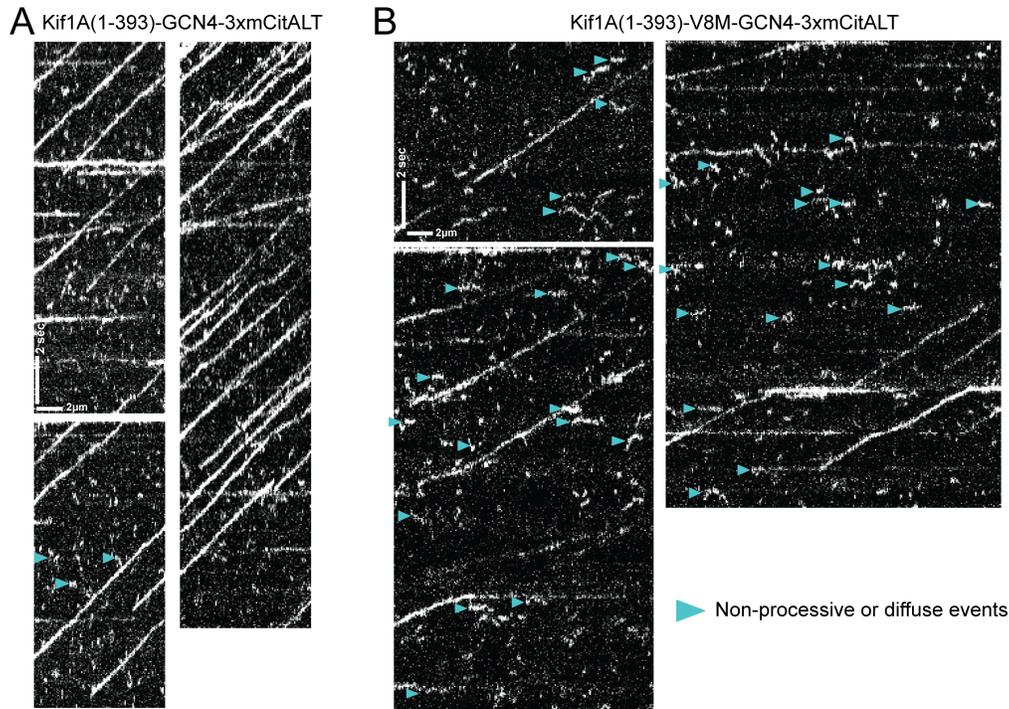
Figure 5.3: KIF1A V8M mutants undergo non-processive and diffuse events. (**A**) KIF1A WT and (**B**) V8M mutants under single-molecule, unloaded conditions. Blue arrowheads indicate single motor non-processive or diffuse events. Bars as: x-axis (2 s), y-axis (2 μm).

disruption of hydrophobic interactions will cause the NL to dock from the motor domain core. Comparison of inter-residue distances in the WT and Y89D mutant simulations reveals increased distances between residue (Y)89D ($\alpha1/\beta3$) and residues $358 - 360$ (NL) (indicated by blue lines in Figure 5.4B). The largest difference in distance was between Y89 ($\alpha1/\beta3$) and I360 (NL): an increase from 5.57 ± 3.49Å in WT to 10.04 ± 1.64Å in the Y89D mutant (complete list of significantly different ($P < 10^{-5}$) inter-residue distances in Appendix D, Table D.2). The negatively charged aspartate residue now flips inwards towards the central β sheet, and forms backbone and side-chain interactions with V421, S242, and K243 in $\beta7$ (red square between $\beta7$-$\alpha1/\beta3$ in Figure 5.4A, inter-residue distances listed in Appendix D, Table D.2).

Inspection of MD simulation trajectories also show NL undocking after a few nanoseconds in all four trajectories of Y89D mutant, and remains undocked towards the end of the simulations (representative snapshots shown in Figure 5.5A,B). The NL, along with the CS and L13, is responsible for generating the "power stroke" for force generation in kinesin (*Hwang et al.*, 2008; *Khalil et al.*, 2008). The NL moves in response to ATP binding between "undocked" (not force-generating) and "docked" (force-generating) orientations. Disruption of NL docking in the Y89D mutation is predicted to impair the force-output of the kinesin motor domain, as reported by us in the case of a mutant disrupting NL docking in kinesin-1 motor *Rn*KIF5C (See Chapter IV). Furthermore, in-line with coupling observed between NL docking and the active-site regions of kinesin-1 and kinesin-5 motors in our previous studies (*Muretta et al.*, 2015, 2018), the Y89D mutant has an allosteritc effect on the nucleotide coordinating/hydrolyzing regions in KIF1A. However, unlike a clear effect as seen in kinesin-1 and kinesin-5 motors, we see a mixed effect on the active site regions: the interactions between nucleotide coordinating PL and ATP-gating helix $\alpha 0$ (*Hwang et al.*, 2017) are enhanced (shorter) in the Y89D mutant (red squares between S1-$\alpha 0$ and S1-PL in Figure 5.4A), while the interactions between nucleotide hydrolysis and exchange-related switch regions (*Kull and Endow*, 2002) are weakened (longer) in the mutant (blue squares between S2-PL and S2-S1 in Figure 5.4A). The effect of Y89D mutation on the motility properties of KIF1A motors are currently being charazterized by our collaborators.

5.4 Methods

5.4.1 Structural model preparation of KIF1A motor complex

Initial coordinates of KIF1A kinesin motor domain in the ATP-bound state (ATP analogue, ANP-PNP), in complex with the tubulin heterodimer were taken from PDB
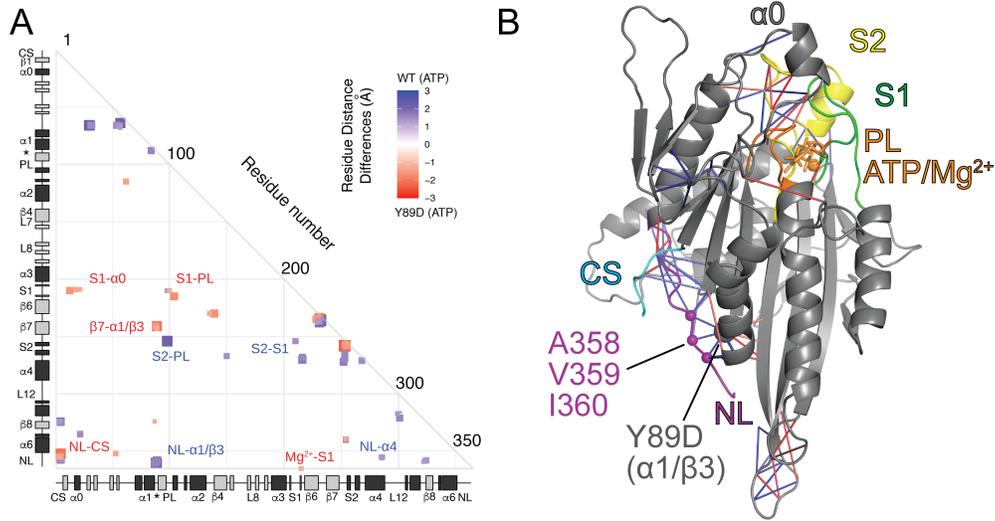
Figure 5.4: KIF1A Y89D ($\alpha 1/\beta 3$) mutant impairs neck linker docking. (**A**) Differences in residue-residue distances between WT KIF1A and Y89D mutant motors in ATP- and microtubule-bound states as determined from MD simulations. Secondary structures colored as: α helices (black), β strands (grey), coverstrand (CS, cyan), P loop/ATP/Mg^{2+} (orange), Switch 1 (green), Switch 2 (yellow), and neck linker (NL, purple). Residue-residue interactions that are significantly closer ($P < 10^{-5}$) in the WT (blue) or Y89D mutant (red) are indicated. The magnitude of the distance change is indicated by size and color intensity. (**B**) Cartoon representation of residue distances changes. Blue lines depict residue-residue distances that are shorter in the WT motor versus Y89D mutant motor ($P < 10^{-5}$). The magnitude of the distance change is indicated by line color intensity. Residues 358-360 in NL shown to highlight the loss of hydrophobic contacts in the Y89D mutant. Colors according to (A).

4UXP (*Atherton et al.*, 2014). The kinesin motor domain sequence was already that of *Hs*KIF1A (Uniprot ID Q12756). Missing coordinates, where applicable, were modeled using MODELLER v9.18 (*Šali and Blundell*, 1993). A total of 100 models were generated with the following options in MODELLER: variable target function method (VTFM) was set to *slow* with associated conjugate gradient set to 150 iterations; MD with simulated annealing option was set to *slow*; and the entire optimization process was repeated twice. The top-scoring model was selected with discrete optimized protein energy (DOPE) score (*Shen and Sali*, 2006) for loop refinement. The top-scoring model was selected for MD simulations.
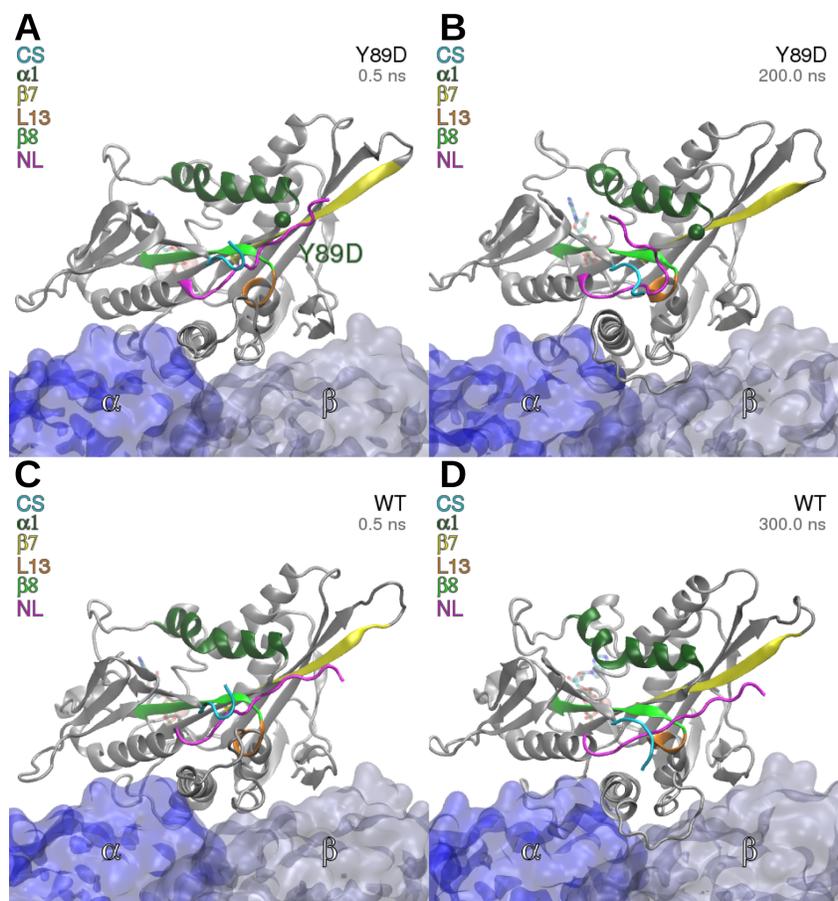
Figure 5.5: NL undocking in Y89D ($\alpha1/\beta3$) mutant. Representative snapshots from simulations of (**A,B**) Y89D ($\alpha1/\beta3$) mutant in ATP-state, and (**C,D**) wildtype (WT) in ATP-state. Snapshots on the left (A,C) are from the initial starting-point, and snapshots on the right (B,D) from the end time-point of MD simulations (MD time-point shown in top-right corner). The Y89D mutation in $\alpha1/\beta3$ (dark green sphere) introduces a negative charge causing loss of hydrophobic contacts with residues 358-360 in NL (magenta), causing the NL to undock (A,B). Secondary structure elements are colored: coverstrand (CS, cyan), $\alpha1$ (dark green), $\beta7$ (yellow), Loop13 (L13, orange), $\beta8$ (light green), neck liker (NL: $\beta9$ and $\beta10$, magenta), with α-tubulin (blue), β-tubulin (light blue). The changes in residue-residue distances in the Y89D mutant compared to the WT are quantified in Appendix D, Table D.1.

5.4.2 Molecular dynamics simulations of KIF1A motor complex

Energy minimization and molecular dynamics (MD) simulations were performed with AMBER 18 (*Case*, 2018) and the ff99SB AMBER force field (*Hornak et al.*, 2006). Nucleotide parameters were obtained from (*Meagher et al.*, 2003). Histi-

dine protonation states were assigned based on the their pKa values calculated by PROPKA (*Li et al.*, 2005). The simulation setup and procedures were adopted as described (see *Methods* section in *Chapter VI*). MD simulations were started from equilibrated structures with at least four independent runs of at least 200 ns each. All simulations were performed in-house on NVIDIA GPU cards with the GPU version of PMEMD (*pmemd.cuda*). We thank NVIDIA for their gift of GPU card through their Academic GPU seed grant. Trajectory analyses were carried out in R using the Bio3D v2.3-3 package (*Skjærven et al.*, 2014).

5.4.3 Residue-residue distance differences

Statistically significant residue-residue distance differences between wiltype (WT) and mutant ATP-bound kinesin motor domain in complex with tubulin heterodimer were identified with ensemble difference distance matrix (eDDM) analysis routine (*Muretta et al.*, 2018). For this analysis, a total of 400 conformations were obtained for each state under comparison by extracting 100 equally time-spaced conformations from the last 20 ns of each simulation replicate. The details of obtaining the distance matrices from simulation trajectories, their processing, and the method of selecting significantly different residue-residue distances are discussed in Chapter III (see its *Methods* section). Briefly, the eDDM routine reduces the difference between long distances while difference between short distances are kept intact. The significance of residue distance variation between apo and ATP-bound states, and between ATP-bound and mutant states, were evaluated with the Wilcoxon test. Residue pairs showing a p-value$< 10^{-5}$ and an average masked distance difference > 1Å were considered statistically significant residue-residue distance differences for further analysis.

5.5 Discussion and Conclusion

In this chapter, we have presented our preliminary findings on the dynamical effects of an autosomal dominant mutation V8M ($\beta1$) and a *de novo* mutation Y89D ($\alpha1/\beta3$) in KIF1A associated with neurodevelopmental disorders (*Ohba et al.*, 2015; *Esmaeeli Nieh et al.*, 2015; *Iqbal et al.*, 2017). Inter-residue distance analysis of MD simulations of wildtype (WT) and V8M mutant motors in ATP- and microtubule-bound state reveal allosteric effects on motor active site. Single-molecule motility assays showed that motors with V8M mutation were slower than WT motors under unloaded conditions. Interestingly, a recent study uploaded on bioRxiv report opposite effects of this mutation: they observed higher velocity and on-rates for KIF1A and *C. elegans* homolog unc-104 mutant motors (*Chiba et al.*, 2019). We note that our computational and experimental setup utilized a truncated version of the KIF1A motor lacking the tail domain, whereas (*Chiba et al.*, 2019) utilized full-length proteins in their study with possibility of an additional role of autoinhibition on motor function. This highlights the need for a combination of studies, including single-molecule experiments, force-measurements through optical trapping, and cell-based dispersion assays in characterizing effects of mutations on motor function. Our previous studies have highlighted the differences in motor function for the same given mutation under different conditions. For example, in kinesin-5 mitotic motor Eg5, an acetylation mimetic mutation, K146Q, enhanced single-molecule motility properties, but exhibited "breaking" properties when working in teams, slowing sindle separation in cells during mitosis (see Chapter III and (*Muretta et al.*, 2018)). Furthermore, mutations disrupting cover-neck bundle formation and/or NL docking also cause enhanced motility properties in single-molecule assays, but crippled force production and impaired cellular transport of high-load cargo under physiological conditions (see Chapter IV and (*Budaitis et al.*, 2019)). Experiments characterzing the motility properties of $\alpha1/\beta3$ mutant Y89D, as well as force measurements for V8M and Y89D

mutants through optical trapping experiments are currently underway.

CHAPTER VI

Dynamics of *C. elegans* tubulin

Contents of this chapter were partially published in (*Chaaban et al.*, 2018). Citation: Chaaban, S., S. Jariwala, C.-T. Hsu, S. Redemann, J. M. Kollman, T. Müller-Reichert, D. Sept, K. H. Bui, and G. J. Brouhard (2018), The Structure and Dynamics of C. elegans Tubulin Reveals the Mechanistic Basis of Microtubule Growth, *Developmental Cell*, pp. 1–14, doi: 10.1016/j.devcel.2018.08.023. This has been a collaborative effort. The computational work pertaining molecular dynamics simulations and subsequent analysis was done by myself. Most of the experimental work was done by S.C. in G.J.B. lab (McGuill University, work related to Figures 6.1, 6.6, 4.7A, **??**), with additional experimental data from S.R. (University of Virginia), T.M.R. (Technische Universität Dresden), C.T.H. and K.H.B. (McGill University), and J.K. (University of Washington).

6.1 Abstract

Microtubules are dynamic biopolymers composed of stacks of structurally homologous α- and β-tubulin subunits, cylindrically arranged in about 10 to 16 protofilaments. They play a crucial role in the cytoskeleton of the cell, providing structural support as well as facilitating cellular transport via molecular motors, beating of cilia and flagella, and separation of chromosomes during the cell cycle. Microtubule

97

polymerization is central to their biological functions, with varying growth rates in vitro and in vivo. Recent studies have shown that microtubules derived from *C. elegans* polymerize at a rate three times faster than those from vertebrates. The mechanism for achieving faster growth rates is not clear, especially given that the α- and β-tubulins share $\sim 88\%$ and $\sim 89\%$ sequence identity with bovine (as candidate vertebrate) tubulins, respectively. Here, we investigated the dynamical properties of nematode and vertebrate tubulins using molecular dynamics (MD) simulations of isolated dimers in explicit solution. The simulations of isolated dimers indicate that one of the lateral contact loops in nematode tubulins is more ordered, potentially explaining a role in faster polymerization rates. These results, complemented by *in vitro* experiments carried out by our collaborators, provide direct evidence that amino acid substitutions in *C. elegans* tubulin lead to new secondary structure formation, and that the *C. elegans* tubulin has higher free energy in solution. We also propose a model wherein the ordering of lateral contact loops activates tubulin for growth, improving our understanding of microtubule polymerization.

6.2   Introduction

The microtubule (MT) "dynamic instability" is central to its functional regulation in cells (*Mitchison and Kirschner*, 1984), wherein microtubules grow by polymerizing tubulin subunits, undergo catastrophe events, and rescue back to growth stage. All eukaryotic microtubules studied thus far exhibit dynamic instability under *in vitro* conditions, perhaps a feature of their highly conserved sequence and structure (*Little and Seehaus*, 1988; *Fygenson et al.*, 1994). How do subtle changes in tubulin sequence affect differences in observable parameters such as MT polymerization rates? Also, why do some MT lattices undergo longitudinal compaction following GTP-hydrolysis but some do not, even though they do not differ wildly in sequence? (*Alushin et al.*, 2014; *Howes et al.*, 2017; *von Loeffelholz et al.*, 2017). Guided by these questions, we

sought to address of the mechanistic differences in MT dynamics using tubulin from *C. elegans* and comparison its dynamics to those of well-studied mammalian tubulin from *B. taurus*.

Our collaborators observed that *C. elegans* tubulin grows very fast ($\sim$ 3 times faster than *B. taurus* tubulin, under microtuble reconstition assays, indicating that the dimer itself is adapted to fast growth. Tubulin from yeast *S. cerevisiae* also exhibit higher on-rates under *in vitro* conditions. However, unlike the long lifetimes of yeast MTs (*Geyer et al.*, 2015), *C. elegans* combined gast growth rates with frequent catastrophe events *in vitro*. What is the mechanism behind the highlyl dynamic nature of *C. elegens* tubulin? In thus study, we combine molecular modeling, molecular dynamics simulations, sequence analysis, cryo-electron microscopy and temperature-dependent microtubule reconstition assays to answer this question. We find that the subtle sequence changes in *C. elegans* tubulin, as compared to *B. taurus* tubulin, occcur in lateral contact loops of the tubulin dimer. Using molecular dynamics simulations, we show that these lateral contact loops in the *C. elegans* dimer have an increased tendency to form secondary structures, which was confirmed by CryoEM images which showed that the lateral contact loop H1-S2 was structured in *C. elegans* MTs. This indicates the structuring of lateral contact loops may be a key rate-limiting step in MT growth. Finally, with temperature-dependent reconstitution assays, our collaboratos are able to directly address our predictions and observations, and support our hypothesis that the *C. elegans* tubulin has a higher free energy than *B. taurus* tubulin in solution.

6.3 Results

6.3.1 *C. elegans* tubulin combines fast growth and frequent catastrophe *in vitro*

Before embarking on structural characterization of *C. elegans* tubulin, we wanted to understand their growth and catastrophe properties. Our collaborators used reconstition assays with tubulin purified from mixed life-stage liquid cultures of *C. elegans* (*Widlund et al.*, 2012). The purified tubulin is primarily composed of tbb-1, tbb-2, and tba-2 (see (*Chaaban et al.*, 2018) for details). Next, microtubule growth was reconstitued *in vitro* by nucleation from GMPCPP-stabilized microtubule templates and imaged the microtubules by differential interference contrast (DIC) microscopy Figure 6.1A). Surprisingly, our collaborators found that *C. elegans* microtubules nucleated from the templates at tubulin concentrations as low as $2\,\mu$M, well below the nucleation threshold for *B. taurus* (bovine brain) microtubules (critical concentration, $C_c = 9.5 \pm 0.1\,\mu$M (*Wieczorek et al.*, 2015)). Even at these low concentrations, *C. elegans* microtubules grew remarkably fast, reaching $1.65 \pm 0.02\,\mu$m·min^{-1} at $5\,\mu$M tubulin, a growth rate that requires nearly $20\,\mu$M of *B. taurus* tubulin (Figure 6.1B). The plot of the microtubule growth rate versus *C. elegans* tubulin concentration gave an apparent on-rate constant of $k_a = 9.21 \pm 0.46$ dimers·μM^{-1}·sec^{-1}, which is more than $3\times$ higher than *B. taurus* (2.90 ± 0.17 dimers·μM^{-1}·sec^{-1}) (Figure 6.1E and F). These results suggest that *C. elegans* tubulin has a higher affinity for microtubule ends than *B. taurus* tubulin, resulting in templated nucleation at low concentrations and fast growth.

It is commonly assumed that faster microtubule growth leads to more stable polymers that are less prone to catastrophe. This assumtion is based on observations that as microtubules grow faster, their stabilizing caps (a region enriched in GTP-tubulin) become larger (*Bieling et al.*, 2007). Contrary to our expectations that the fast
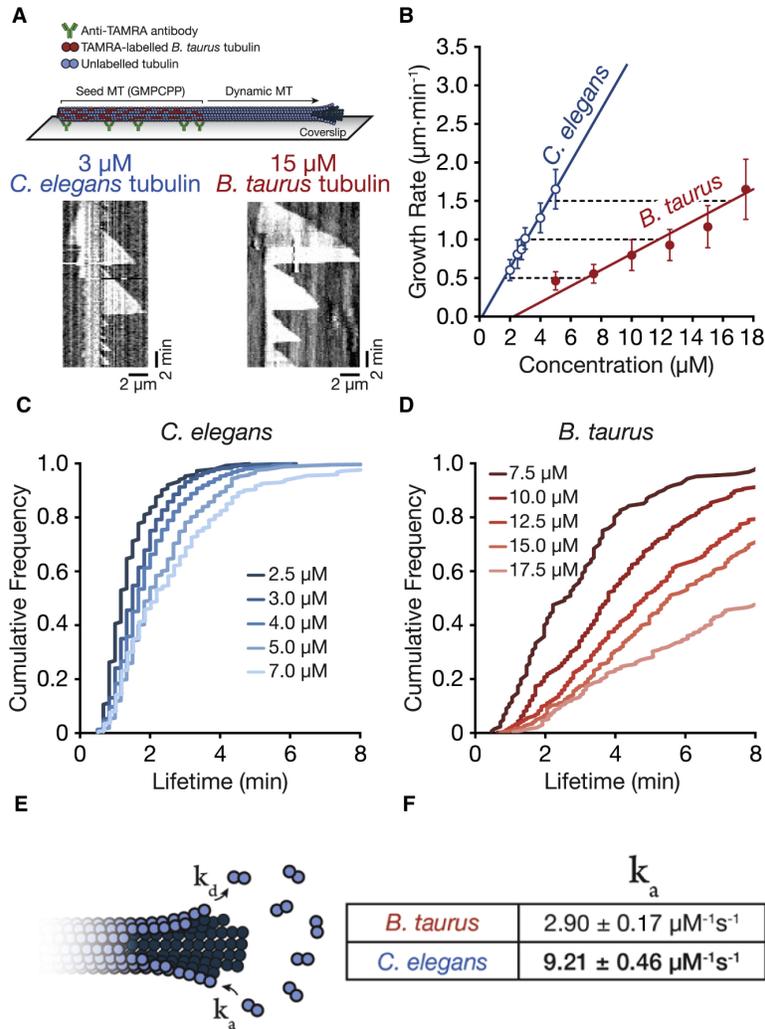
Figure 6.1: Dynamics of *C. elegans* tubulin *in vitro*. (**A**) Schematic of the microtubule growth reconstitution assay. Dynamic microtubules are grown from the seeds, and microtubule dynamics are analyzed from kymographs (bottom). (**B**) Plot of microtubule plus end growth rates as a function of tubulin concentration for *C. elegans* tubulin (blue; n $= 32, 68, 114, 200, 332, 199$ from ≥ 2 replicates) and *B. taurus* tubulin (red; n $= 112, 119, 605, 387, 738, 220$ from ≥ 2 replicates). Error bars represent the propagated SD. (**C** and **D**) Plot of the cumulative distribution of microtubule lifetimes for increasing concentrations of (*C*) *C. elegans* tubulin (n $= 153, 905, 611, 293, 236$ from ≥ 2 replicates) and (*D*) *B. taurus* tubulin (n $= 119, 329, 300, 222, 191$ from ≥ 2 replicates). (**E**) Schematic of tubulin dimer association and dissociation at the microtubule plus end. (**F**) Apparent on-rate constants (k_a) obtained from linear least-squares fitting to growth rate data from microtubule reconstitution assays (*B*). Figure adapted from (*Chaaban et al.*, 2018).

101

growing microtubules should have long lifetimes, our collaborators found that the *C. elegans* microtubules had very short lifetimes (Figure 6.1C). Comparing *C. elegans* and *B. taurus* at the same growth rate of $\sim 1\,\mu\text{m}\cdot\text{min}^{-1}$, the mean lifetime for *C. elegans* microtubules was $1.71 \pm 0.03\,\text{min}$ versus $5.43 \pm 0.15\,\text{min}$ for *B. taurus*. Over the range of concentrations where growth rates overlapped, *C. elegans* microtubules aged at about four times the rate of *B. taurus*. These results demonstrate that *C. elegans* microtubules combine very fast growth with very short lifetimes *in vitro*. This combination contrasts with *H. sapiens* microtubules reconstituted from tubulin mutants or from different isoforms, where faster growth correlates with longer lifetimes (*Ti et al.*, 2016; *Vemu et al.*, 2017). Similarly, *S. cerevisiae* microtubules have the highest reported on-rate constant but extremely long lifetimes (*Geyer et al.*, 2015). Thus, *C. elegans* microtubules are uniquely dynamic among microtubules studied to date, in that they exhibit fast growth and frequent catastrophe *in vitro*.

6.3.2 Lateral contact loops are ordered in *C. elegans*

Comparison of *C. elegans* and canonical tubulin sequences reveal sequence changes clustered in loops involed in inter-protofilament lateral contacts (more on this in section 6.3.5 below). Of note was deletion of G45 and D46, and a D47 to E substitution in *C. elegans* α-tubulin (residue number as per *B. taurus* (Figure 6.2. For alignment of β-tubulins, see Figure 6.3). While the D47E substitution preserves charge, a glutamate residue at the N-terminus has been observed to have a slightly higher α helix propensity, i.e. glutamate forms an N-terminal cap to an α helix more readily (*Williams et al.*, 1987; *Nick Pace and Martin Scholtz*, 1998). The structural effects of these substle changes, if any, were not known. Structured H1-S2 loops may correlate with fast growth by lowering the entropic cost of incorporating a tubulin dimer within the microtubule (MT) lattice. We hypothesized that these lateral loops might be structured in solution for *C. elegans*. In order to test this hypothesis, we per-

formed all-atom molecular dynamics of the *C. elegans* and *B. taurus* tubulin dimers (see *Methods* for details). The total simulated time for each dimer was $3.2\,\mu s$, and the root-mean-square deviation (RMSD) stabilized within the first 50-100 ns. To analyze our simulations, we plotted the frequency of secondary structure formation for each residue in the dimer (Figure 6.4A and B). Although the *C. elegans* dimer adopted the same overall structure as *B. taurus*, we found an increased probability of secondary structure formation in the lateral loops, namely the H1-S2 loops and the M loops. The largest difference was found at the H1-S2 loop of α-tubulin (Figure 6.5A), which forms a short α helix and β sheet (Figure 6.5B). Similarly, the M loops of both α- and β-tubulin formed short α-helices. These secondary structures are almost entirely absent in the *B. taurus* simulations (Figure 6.5A). The α helix in β-tubulin's M loop matches the one induced by zampanolide and epothilone A; (*Prota et al.*, 2013) interpreted this ordering of the M loop as "tubulin activation", or the pre-conditioning of the dimer to facilitate polymer formation. This concept of tubulin activation was used to explain the mechanism of action of microtubule-stabilizing drugs. Our results suggest that the *C. elegans* dimer may adopt an activated form in solution. This activation may explain why the *C. elegans* microtubule grows faster than the canocial microtubules.

Following the secondary strucure analyis of MD simulations, we wanted to confirm whether the lateral contact loops in *C. elegans* tubulin dimer are indeed ordered. To this end, our collaborators generated a near-atomic resolution reconstruction of dynamic *C. elegans* microtubules previously imaged with Cryo-EM (Figure 6.6). In brief, our collaborators first nucleated *C. elegans* microtubules in the absence of templates (such as γ-TuRCs). Further details regarding identificaltion of α- and β-tubulin, and the "seam-search" processing protocol used can be obtained from (*Zhang and Nogales*, 2015) and (*Chaaban et al.*, 2018). Our tba-2/tbb-2 homology model was then docked into the final density map ($4.8\,\text{Å}$ overall resolution) and refined (Figure 6.6C).
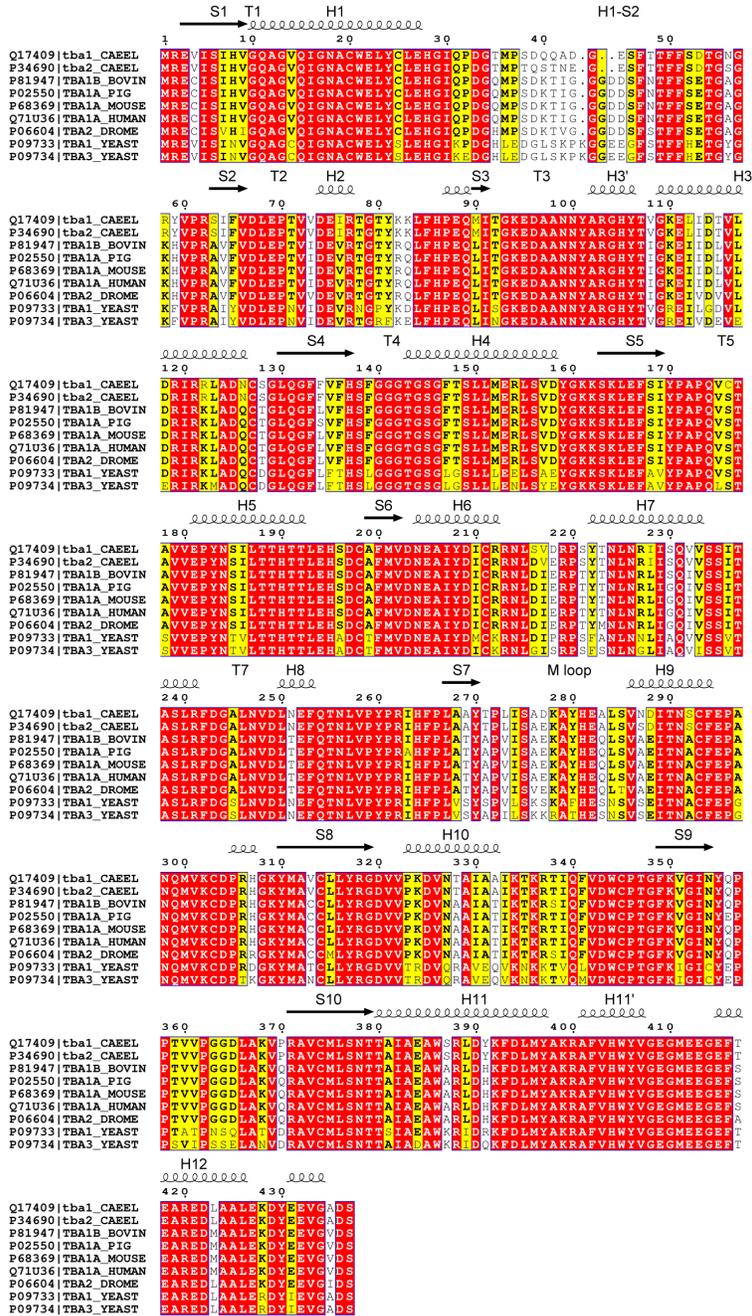
Figure 6.2: Multiple sequence alignment of *C. elegans* and canonical α-tubulin sequences. Uniprot IDs and common gene names are listed on the left for each sequence (row). Invariant residues are shown in red. Conserved residues are shown in yellow, with the most frequent residue observed for a given position (column) highlighted in bold. Secondary structures (sheets/strands and helices) annotated on top. C-terminal tail excluded from alignment. Figure rendered with ESPript 3.0 (*Robert and Gouet*, 2014) and modified locally.
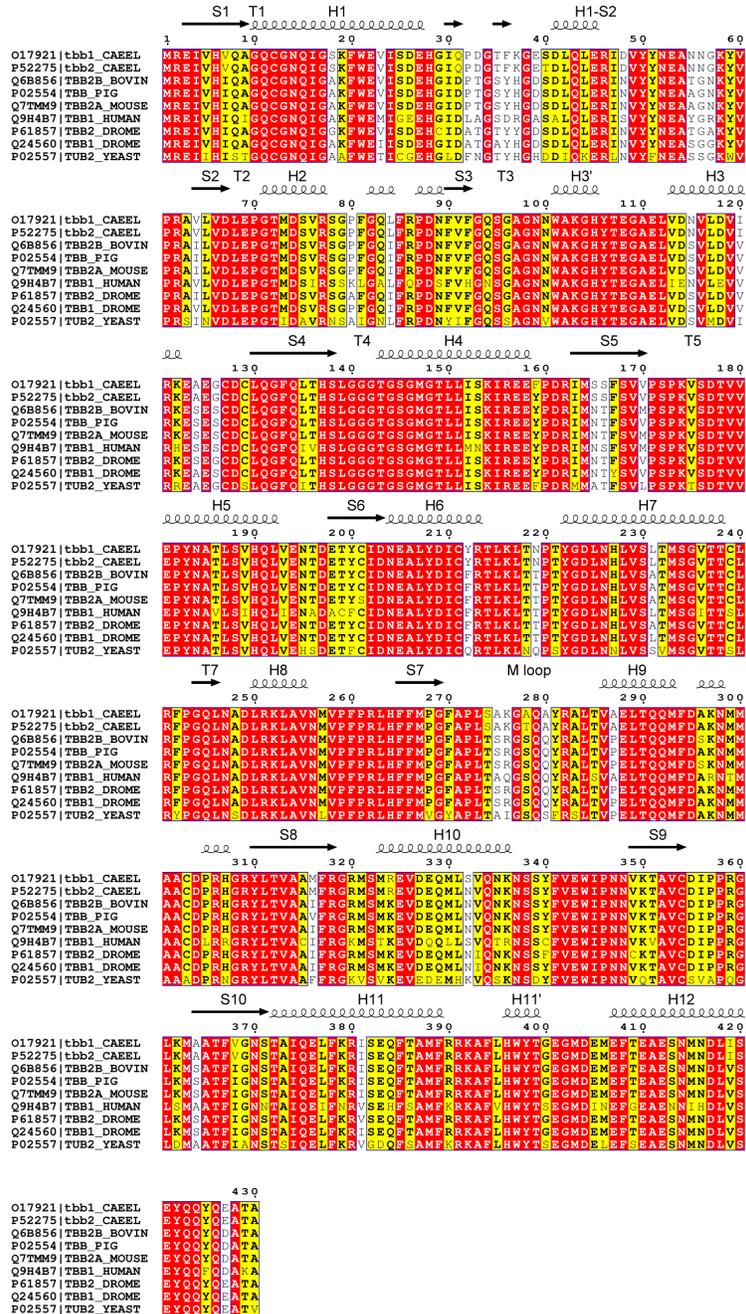
Figure 6.3: Multiple sequence alignment of *C. elegans* and canonical β-tubulin sequences. Uniprot IDs and common gene names are listed on the left for each sequence (row). Invariant residues are shown in red. Conserved residues are shown in yellow, with the most frequent residue observed for a given position (column) highlighted in bold. Secondary structures (sheets/strands and helices) annotated on top. C-terminal tail excluded from alignment. Figure rendered with ESPript 3.0 (*Robert and Gouet*, 2014) and modified locally.
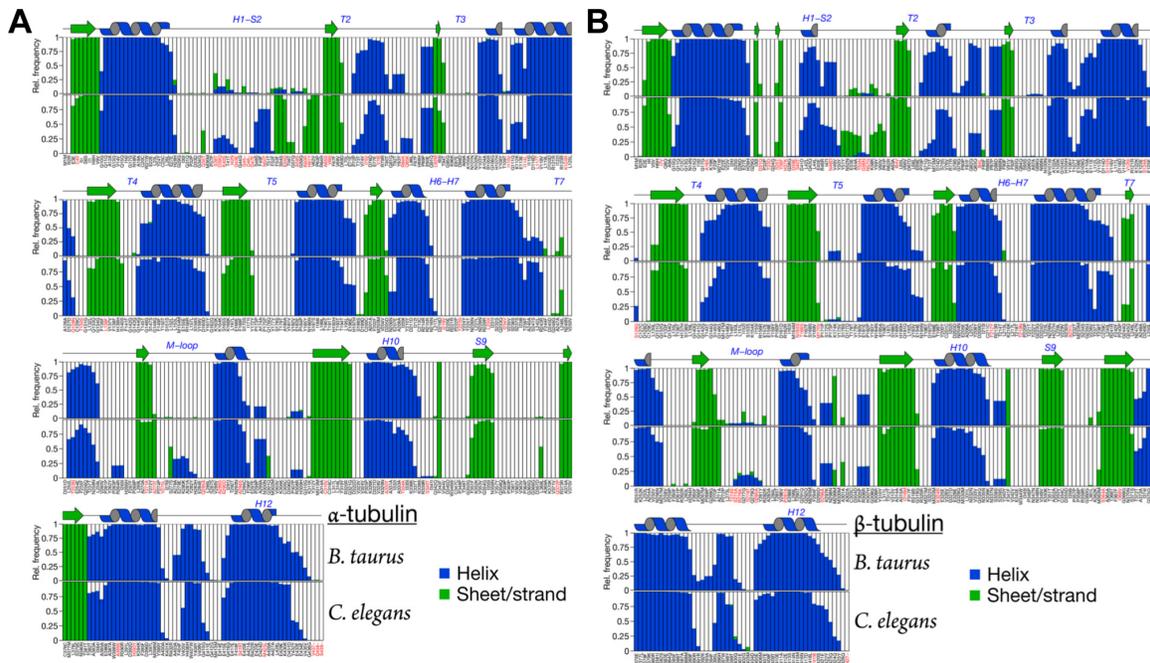
Figure 6.4: Relative frequency of secondary structure elements. (**A**) α-tubulin and (**B**) β-tubulin frequency of secondary structure elements for *B. taurus* (*top*) and *C. elegans* (*bottom*). Sheets and strands are shown in green, helices in blue. Residues are numbered according to the *B. taurus* tubulin sequence and red represents dissimilar residue between *B. taurus* and *C. elegans* tubulin.

Visual analysis revealed structuring of residues 38–45 in the H1-S2 loop of α-tubulin, which are typically disordered and therefore unresolved in other structures (e.g., *S. scrofa* and *H. sapiens*) (Figure 6.7) (*Zhang et al.*, 2015; *Vemu et al.*, 2016). Taken together, these results support the idea that the pre-ordering of lateral loops may be responsible for the faster growth rates of *C. elegans* microtubules.

6.3.3 Reduced flexibility of lateral loops in *C. elegans*

To characterize the global trends in flexibility of *C. elegans* and *B. taurus* tubulin dimers, we applied root mean-sqaure fluctuation analysis to the trajectories from MD simulations. Both states were found to display an overall common trend of per-residue root mean-square fluctuations (RMSF, Figure 6.8). This trend consisted of
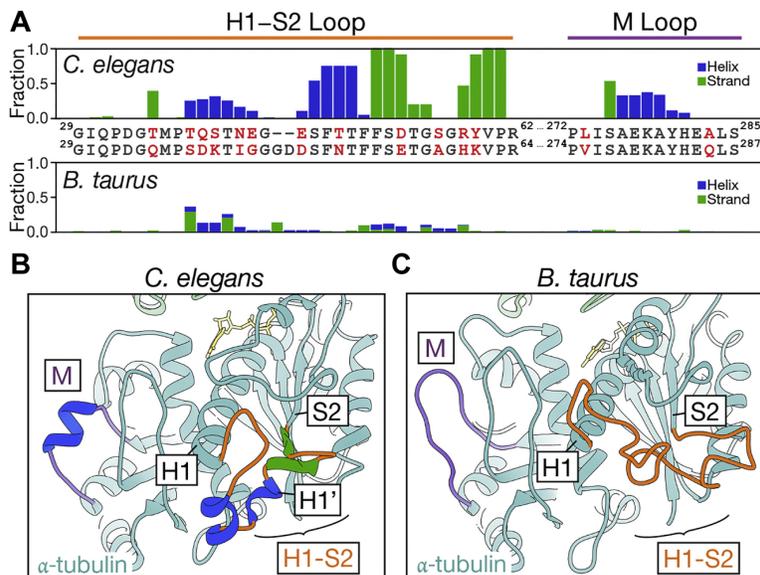
Figure 6.5: Lateral contact loops are ordered in the *C. elegans* dimer. (**A**) α-tubulin and (**B**) β-tubulin frequency of secondary structure elements for *B. taurus* (*top*) and *C. elegans* (*bottom*). Sheets and strands are shown in green, helices in blue. Residues are numbered according to the *B. taurus* tubulin sequence and red represents dissimilar residue between *B. taurus* and *C. elegans* tubulin.

comparatively low RMSF values for the buried central core β sheets in the intermediate domains and the core helices H7 in each of the tubulin monomers, with larger values evident for solvent-exposed regions, including the H1-S2 and M loops. Structural regions with statistically significant differences between the dimer states (shaded regions in Figure 6.8, $P < 0.01$ with Wilcoxon rank sum test and ΔRMSF > 0.2) primarily included the lateral contact loops H1-S2 (residues 28-64) and M loop (residues 273-287) in α-tubulin. We note no significant flexibility differences for nucleotide-contacting regions (marked with black vertical bars in Figure 6.8). The reduced flexibility of lateral contact loops, specifically residues 38-61 (38-63 in *B. taurus*), can be attributed to the structured state of H1-S2 loop in *C. elegans* α-tubulin in solution (Figures 6.5, 6.7A, and 6.8A). This reduced flexibility of lateral contact loops further complements data shown above regarding "pre-ordered" state of *C. elegans* tubulin dimer in solution.
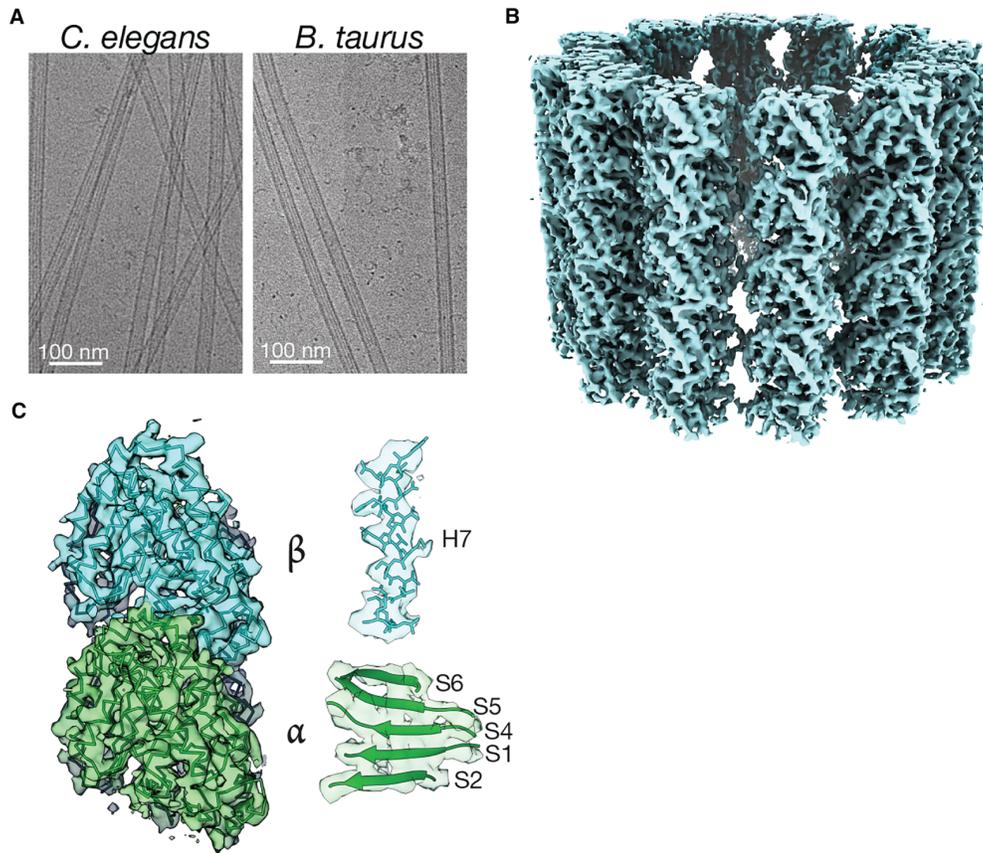
Figure 6.6: Cryo-EM structure of *C. elegans* microtubule. (**A**) Spontaneously nucle-ated *C. elegans* (left) and *B. taurus* (right) microtubules visualized by cryo-EM. (**B**) A reconstruction of 12-pf *C. elegans* microtubules at 4.8 Å resolution. (**C**) Trace of homology model of *C. elegans* tba2/tbb2 tubulin dimer fitted in the electron density map (left) and a close up of the fit showing a resolved core α helix (H7 of β-tubulin, top-right) and β sheet (S1-S2, S4-S6 of α-tubulin, bottom-right). Figure adapted from (*Chaaban et al.*, 2018).

6.3.4 Principal component analysis of *C. elegans* tubulin dynamics

Principal component analysis (PCA) is a dimensionality reduction technique in-volving orthogonal transformation of the original data into a set of linearly uncorre-lated variables termed principal components. The application of PCA to experimental structures and MD trajectories to map conformational differences in a range of pro-tein families has been previously discussed (*van Aalten et al.*, 1997; *Gorfe et al.*, 2008; *Scarabelli and Grant*, 2013). The initial PCA involved constructing a linear
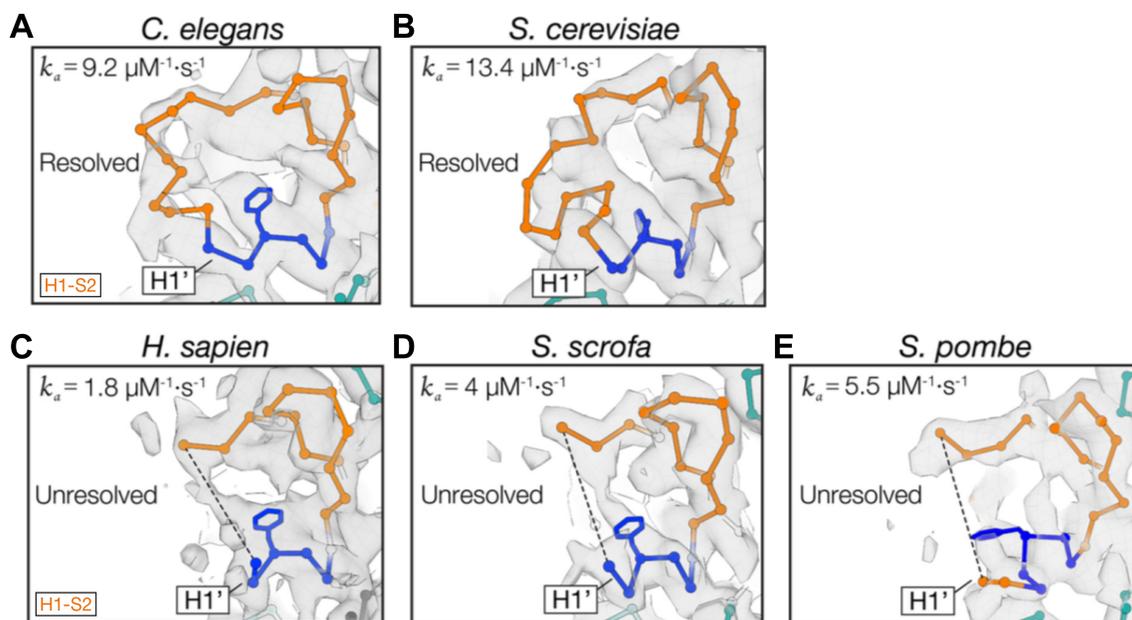
Figure 6.7: Structural comparison of α-tubulin H1-S2 loop. The H1-S2 loop shown in stick (orange) and electron density (gray), highlighting α-tubulin helix H' (blue) of (**A**) *C. elegans* (PDB:6E88, EMD:9004) (*Chaaban et al.*, 2018), (**B**) *S. cerevisiae* (PDB:5W3F, EMD:8755) (*Howes et al.*, 2017), (**C**) *H. sapiens* (PDB:5JCO, EMD:8150 (*Vemu et al.*, 2016), (**D**) *S. scrofa* (PDB:3JAS, EMD:6353) (*Zhang et al.*, 2015), (**E**) *S. pombe* (PDB:5MJS, EMD:3522) (*von Loeffelholz et al.*, 2017). The apparent on rate constant, k_a, is also shown in the top-left corner of each panel. Figure adapted from (*Chaaban et al.*, 2018).

basis set— a principal component subspace (PC-space)— from coordinates obtained from simulations of *B. taurus* tubulin dimer (showin in red for the first two PCs, PC1 and PC2, in Figure 6.9). The first two PCs together explain $\sim 39\%$ of variance in the simulations of *B. taurus* tubulin dimer. The simlations of *C. elegans* tubulin dimer were then projected onto this PC-space (blue, Figure 6.9A). It is apparent from these projections that the *C. elegans* dimer samples a comparatively restricted region of the conformational space. To interpret the conformational space defined by this PC-space, we projected selected experimentally-determined structures of tubulin dimers onto this space (Figure 6.9A, tubulin dimers shown as circles, dimers extracted from Cryo-EM images of microtubules shown as triangles). The region corresponding to
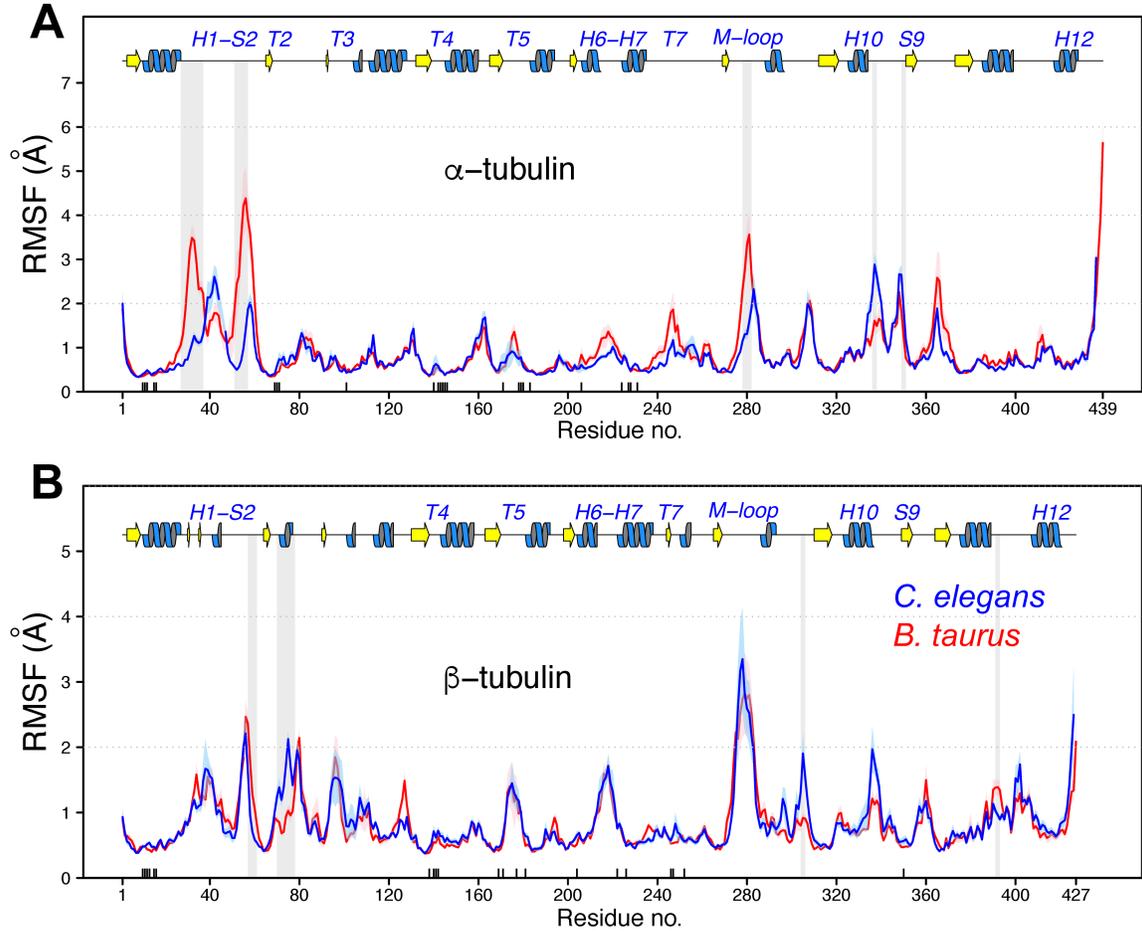
Figure 6.8: Comparison of residue-wise flexibilities of tubulin heterodimer. Distinc flexibilites of *C. elegans* and *B. taurus* tubulin dimer, measured as root mean-square fluctuations (RMSF) of (**A**) α-tubulin, and (**B**) β-tubulin. Regions that show statistically significant differences between states ($P < 0.01$ and ΔRMSF$ > 0.2$) and are composed of at least three consecutive residues are highlighted with a gray-shaded background. The consensus secondary structure is reported above the plots with β strands in yellow and α helices in blue.

the hihgest density of conformers from simulations of *C. elegans* dimer correspond to a "MT-lattice-like" conformation. While conformers from simulations of *B. taurus* also sample this conformation, they sample a much wider region of the conformational space. However, we note two issues with this prilimnary analysis: PCA maximizes the variance, in this case, of simulations of *B. taurus* tubulin dimer. So our PC-space, by design, maximizes the variance in the conformational space sampled by *B. taurus*

tubulin dimer. Furthermore, the cumulative variance explained by the first two PCs only amounts to $\sim 39\%$. Hence, we then improved our PCA to define a PC-space built from experimental structures, as a reference point for all further comparisons of the dynamics of tubulin dimers.

To construct a new PC-space of the conformational variability within tubulin dimers, a set of 350 experimental structures were first obtained from the RCSB Protein Data Bank (*Berman et al.*, 2002). The structures were selected after a HMMER search (*Eddy*, 2011) of sequences (α-tubulin: TBA1B, P81947 and β-tubulin: TBB2B, Q6B856) against the PDB database using Bio3D-web (*Jariwala et al.*, 2017; **?**). A final set of 127 crystallographic and EM structures was obtained after selecting unique dimers from structures with resolution better than 5 Å (see Table E.1 in Appendix E). PCA was performed on 425 equivalent, non-gap Cα atoms from each of the dimers after superposition (Figure 6.9A). The first two principal components (PCs) account for over 84% of the total varance in atomic positions of the equivalent Cα atoms. The major features described by first two PCs, and details of PCA, are described in Chapter VI. In brief, the two clusters evident from this analysis refer to tubulin dimers in "MT-lattice-like" and "bent" conformations. The conformers in the "bent" cluster are typically tubulin oligomers associated with microtubule dynamics regulating proteins such as stathmin. The conformers in the "MT-lattice-like" cluster are structures obtained from Cryo-EM images of microtubules bound with GDP, GMPCPP, or GTPγS in the nucleotide-binding pocket of β-tubulin. Conformers from simulations of *C. elegans* and *B. taurus* tubulin dimers were then projected onto this PC-space, shown in contours, as described above. It is apparent from these projections that the *C. elegans* and *B. taurus* tubulin dimers both sample "bent" and "MT-lattice-like" conformations, consistent with both experimental and computational studies showing that both GTP- and GDP-tubulin sample bent conformations in solution (*Andreu et al.*, 1989; *Rice et al.*, 2008; *Gebremichael et al.*, 2008; *Grafmüller and Voth*, 2011).

The majority of the conformers from simulations of *C. elegans* tubulin dimer sample the "MT-lattice-like" conformation, as shown by larger density of contours near the "MT-lattice-like" experimental structures. The *B. taurus* tubulin dimer, in comparison, have a larger density corresponding to a "slightly-bent" conformation. Thus, the PCA complements our other analyses, supporting the idea that free *C. elegans* tubulin dimer is in a "primed" state in solution.
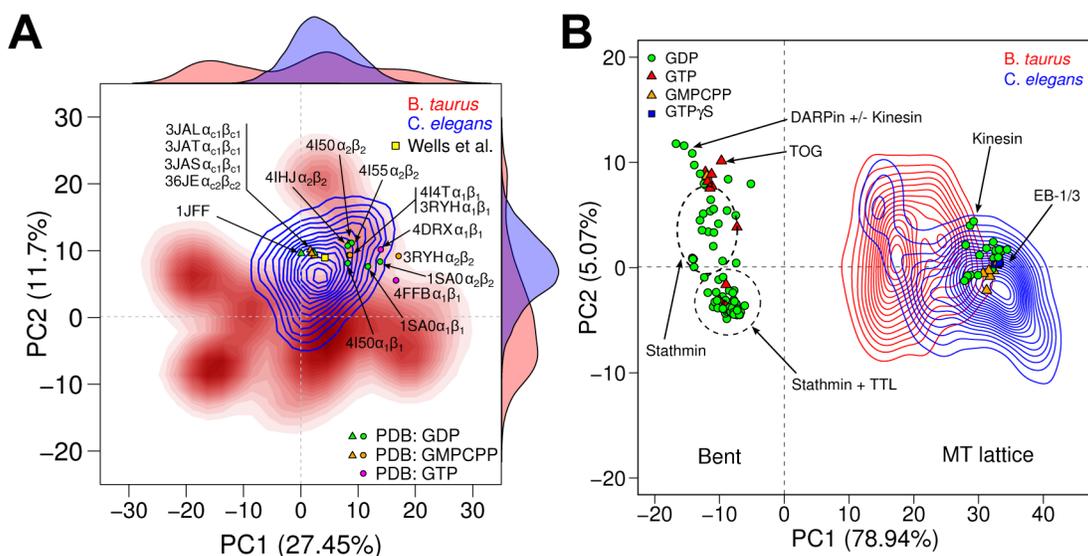
Figure 6.9: Principal component analysis of tubulin dimer simulations. (**A**) Conformer plot of *C. elegans* simulation trajectories projected onto the principal planes defined by *B. taurus* simulations, showing the two most significant PCs (PC1 and PC2). Projections of tubulin dimers from experimental structures are labelled with their PDBs IDs and colored as per the nucleotide state of β-tubulin (green: GDP, orange: GMPCPP, magenta: GTP. The yellow square is projection of average-coordinates of tubulin dimers from explicit-solvent simulations of a single microtuble ring (*Wells and Aksimentiev*, 2010). (**B**) *C. elegans* and *B. taurus* simulation trajectories projected onto PC-space defined by 127 tubulin dimers from experimental structures. The two clusters highlighted by PCA are "MT-lattice-like" and "bent" tubulin dimer conformations. Color according to (*A*) with additional GTPγS tubulin dimers shown as blue squares. The values within brackets in the axis labels represent the percentage of the total variance captured by the corresponding eigenvector in the respective PCA.

6.3.5 Sequence divergence in *C. elegans* tubulin

The computational and experimental data above show the unique, more dynamic nature of *C. elegans* as compared to *B. taurus* tubulin. After obtaining the structure of *C. elegans* tubulin dimer, we wanted to map the sequence changes onto our atomic model to highlight the global and local evolutionary divergence. We applied two similar methods to study the sequence divergence. We started with a curated Pfam seed alignment as a reference profile for aligning *C. elegans* tubulin sequences (Pfam ID PF00091 with title Tubulin/FtsZ family, GTPase domain). Our collaborators, on the other hand, used the multiple sequence alignments to generate two groups of position weight matrices (PWM)—one for *C. elegans* and the other for canonical tubulins seperately for α- and β-tubulin sequences (see Uniprot IDs in Figures 6.2 and 6.3). The results discussed below are equivalent from both the methods, but the figures shown here are from analysis employing PWM comparison (Figure 6.10). In both approaches, we assigned each residue a "similarity score" by comparing *C. elegans* tubulins (tba-1: Q17409, tba-2: P34690, tbb-1: O17921, and tbb-2: P52275), with those of diverse "canonical" tubulins: *H. sapiens* (TBA1A: Q71U36, TBB1: Q9H4B7), *B. taurus* (TBA1B: P81947, TBB2B: Q6B856), *S. scrofa* (TBA1A: P02550, TBB: P02554), *M. musculus* (TBA1A: P68369, TBB2A: Q7TMM9), *X. laevis*, *D. melanogaster* (TBA2: P06604, TBB1: Q24560, TBB2: P61857) and *S. cerevisiae* (TUB1: P09733, TUB3: P09734, TUB2: P02557). The divergent residues were not randomly distributed. Rather, we found clusters of divergence within the loops that form lateral bonds, especially in the M loops and H1-S2 loops (Figure 6.10A). While the central residues that are buried at the lateral bond interface were conserved (Y281 of β-tubulin and H281 of α-tubulin), the residues adjacent to these contact points showed significant divergence (Figure 6.10B). This divergence in the biochemical environment surrounding lateral contacts may explain the increased dynamics of *C. elegans* microtubules. Note that our analysis excluded the C-terminal tails (CTT)

of both tubulins. The CTTs would be the most divergent, and are so even among the isoforms of the same species (*Luedeña and Banerjee*, 2008). We also note that the CTTs of *C. elegans* microtubules were not resolved in our collaborators' cryo-EM structures. Interestingly, we did not find divergence around the nucleotide binding pockets (Figures 6.10C and D), despite the fact that *C. elegans* microtubules age faster (frequent catastrophe). This result suggests that the rates of GTP hydrolysis and P_i release are not governed directly by the nucleotide pocket but rather by long-range allosteric effects (*Brouhard and Rice*, 2018).
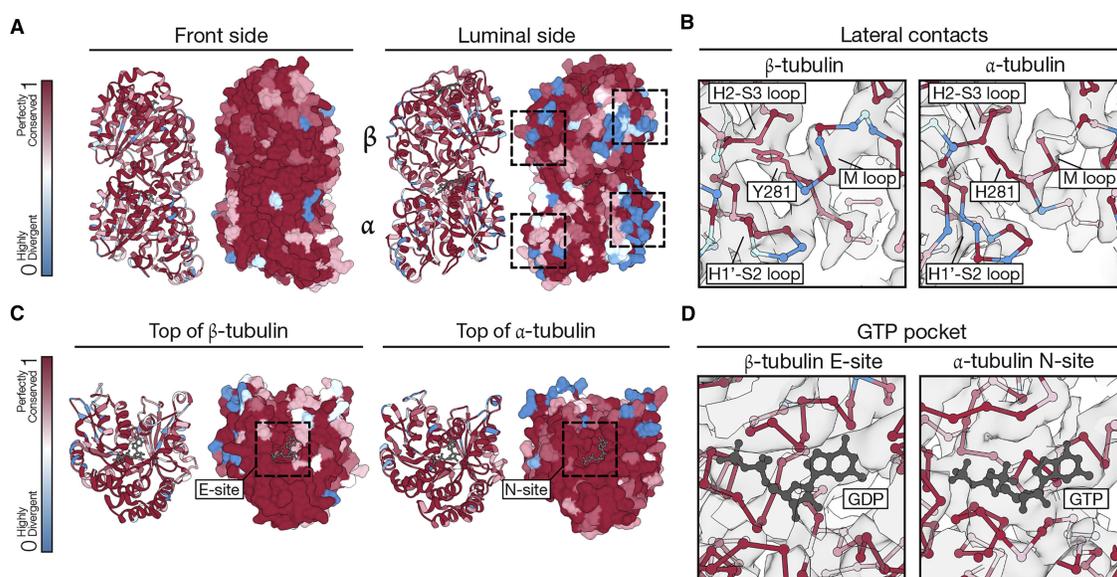
Figure 6.10: Sequence divergence in *C. elegans* tubulin. (**A**) *C. elegans* tubulin dimer depicted in cartoon- and molecular surface representation. Residues in red are more conserved whereas residues in blue are more divergent. Boxes indicate lateral contact loops (H1-S2, H2-S3, and M loop). (**B**) Close up of lateral contact loops marked in (*A*). (**C**) No evolutionary divergence around the nucleotide-binding pocket in α- and β-tubulin. (**D**) Close up of the nucleotide-binding pockets in the electron density map. Figure adapted from (*Chaaban et al.*, 2018).

6.3.6   *C. elegans* tubulin has a lower activation energy for polymerization

The computational and experimental data shown above supports the idea that *C. elegans* tubulin adopts an activated form in solution. We hypothesized that the *C.*

elegans tubulin would require less energy to make it through the structural transitions in the microtubule growth reaction. To test this, we sought to measure the activation energy (E_a or $\Delta G^{\ddagger}$) for a simple model of linear polymer growth (*Oosawa and Asakura*, 1975). To do so, our collaborators measured the apparent on-rate constant, k_a, as a function of temperature for both *C. elegans* and *B. taurus* microtubules (Figures 6.11A and B). For both types of microtubules, increases in temperature drive the reaction forward, leading to faster growth and increased k_a. Using these data, an Arrhenius plot can be generated for the forward reaction, which shows how the reaction rate constant (here, $\ln k_a$) changes with temperature ($1/T$, see Figure 6.11C). In simple systems, the slope of $\ln k_a$ versus $1/T$ directly measures the activation energy, $\Delta G^{\ddagger}$), of the forward reaction (*Arrhenius*, 1889). We note that even though microtubule growth is not a simple system, we can see that *C. elegans* tubulin has a smaller slope than *B. taurus* tubulin (Figure 6.11C). That is, increases in temperature do not drive the reaction forward as strongly for *C. elegans* tubulin. A smaller slope means a smaller $\Delta G^{\ddagger}$ in this model, implying *C. elegans* tubulin requires less energy to enter some transition state of the growth reaction. A smaller $\Delta G^{\ddagger}$ in turn requires that the *C. elegans* dimer has a higher free energy in solution (Figure 6.11D). The faster forward reaction might be due to lowering the height of the activation barrier, or higher free energy of the dimer in solution. As shown in Figure 6.1B, the critical concentration of *C. elegans* tubulin is $10\times$ lower than that of B. taurus tubulin, implying a faster transition from (templated) nucleation to elongation phase. Therefore, the smaller $\Delta G^{\ddagger}$ of the forward reaction in *C. elegans* comes from an increased free energy of the dimer in solution, and the fast growth of *C. elegans* microtubules is driven by an activated dimer.

115
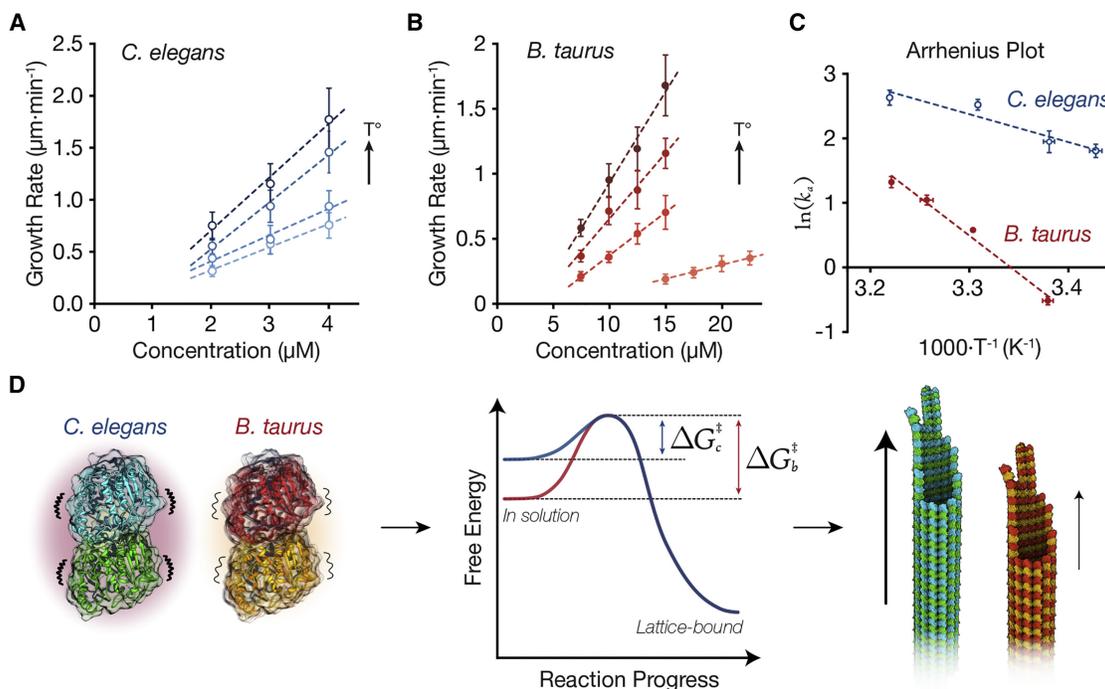
Figure 6.11: *C. elegans* tubulin activation energy. (**A**) *C. elegans* microtubule growth rate as a function of tubulin concentration for multiple temperatures ($17.7 \pm 0.5°C$, $21.8 \pm 0.5°C$, $28.2 \pm 0.2°C$, 36.5C, respectively). (**B**) *B. taurus* microtubule growth rate as a function of tubulin concentration for multiple temperatures ($21.9 \pm 0.5°C$, $28.6°C$, $32.9 \pm 0.6°C$, 36.3C, respectively). (**C**) Arrhenius plot for the forward reaction showing data for *C. elegans* (blue) and *B. taurus* (red). Apparent $E_a = 8.7 \pm 2.2 \, kcal \cdot mol^{-1}$ and $23.6 \pm 1.8 \, kcal \cdot mol^{-1}$, respectively). (**D**) Model for the mechanistic basis of fast growth of *C. elegans* tubulin. The *C. elegans* dimer is more active (left), which leads to a higher free energy in solution. This higher free energy translates into a lower activation energy for the growth reaction (middle, $\Delta G^{\ddagger}$). This lower activation energy translates into faster growth (right). Figure from (*Chaaban et al.*, 2018).

6.4 Methods

6.4.1 Structural model preparation of tubulin dimers

Structural models of tubulin dimers from *C. elegans* and *B. taurus* were prepared as follows for starting molecular dynamics simulations. Initial coordinates were extracted from the central dimer of GMPCPP-*S. scrofa* structure (PDB ID 3JAT) (*Zhang et al.*, 2015). Missing coordinates of α-tubulin residues 38-46 were modeled using MODELLER v9.18 (*Šali and Blundell*, 1993). To change GMPCPP to GTP,

the carbon atom linking α- and β-phosphates was changed to oxygen. Residues were mutated, where applicable, to match sequences of *B. taurus* tubulin (TBA1B: P81947 and TBB2B: Q6B856) or *C. elegans* tubulin (tba-2: P34690 and tbb-2: P52275). A total of 100 models were generated for each with the following options in MOD-ELLER: variable target function method (VTFM) was set to *slow* with associated conjugate gradient set to 150 iterations; MD with simulated annealing option was set to *slow*; and the entire optimization process was repeated twice. The top-scoring model was selected with discrete optimized protein energy (DOPE) score (*Shen and Sali*, 2006) for loop refinement. The top-scoring model for each was selected for MD simulations.

6.4.2 Molecular dynamics simulations of tubulin dimers

Energy minimization and molecular dynamics simulations were performed with AMBER 14 (*Case et al.*, 2014) and the ff99SB AMBER force field (*Hornak et al.*, 2006). Nucleotide parameters were obtained from (*Meagher et al.*, 2003). Histidine protonation states were assigned based on the their pKa values calculated by PROPKA (*Li et al.*, 2005). Starting structures were solvated in a cubic box of pre-equilibrated TIP3P waters molecules, extending 12 Å in each dimension from the surface of the solute. Sodium ions (Na^+) were added to neutralize the systems, followed by addition of sodium and chloride (Cl^-) ions to bring the ionic strength to 0.050 M. Energy minimization was performed in four stages, with each stage consisting of 500 steps of steepest descent followed by 4000 steps of conjugate gradient. First, minimization of solvent was performed by keeping positions of protein and nucleotides fixed. Second, side-chains and nucleotides were relaxed keeping the backbone positions fixed. Third, protein and nucleotide atoms were relaxed while keeping the solvent atoms fixed. Fourth, a last minimization stage was performed with no restraints. The system was gradually heated to 300 K over 25 ps of simulation

117

time in constant-volume (NVT) and periodic boundary conditions (PBC), with restraint of $20\,\mathrm{kcal/mol/\mathring{A}}^2$ on backbone atoms. Constant-temperature (T $=300\,\mathrm{K}$) and constant-pressure (p $=1\,\mathrm{bar}$) (NpT) equilibration was then performed in six stages. First, a $400\,\mathrm{ps}$ NpT equilibration was performed with restraint of $20\,\mathrm{kcal/mol/\mathring{A}}^2$ on backbone atoms. Further stages involved gradually reducing restraints of 20, 10, 5, and $1\,\mathrm{kcal/mol/\mathring{A}}^2$ on α carbons over $5\,\mathrm{ns}$ each. A final NpT equilibration was carried out without any restraint for $5\,\mathrm{ns}$. Subsequent production phase molecular dynamics simulations were then performed under NpT and PBC with random velocity assignments for each run. Particle-mesh Ewald summation was adopted for treating long-range electrostatics. A $12\,\mathring{A}$ cutoff for energy minimization, and a $10\,\mathring{A}$ cutoff for molecular dynamics simulations was used to truncate non-bonded interactions. A $2\,\mathrm{fs}$ time-step was adopted for all molecular dynamics simulations. Hydrogen atoms were constrained using the SHAKE algorithm. All simulations were performed in-house on NVIDIA GPU cards with the GPU version of PMEMD (*pmemd.cuda*). Molecular dynamics simulations were started from equilibrated structures with at least four independent runs of $200\,\mathrm{ns}$ each. Trajectory analyses were carried out in R using the Bio3D v2.3-3 package (*Skjærven et al.*, 2014).

6.4.3 Secondary structure from MD simulations

The frequency of secondary stucture formation for both *C. elegans* and *B. taurus* tubulin dimers were obtained from MD simulations. Equally spaced conformations, $1\,\mathrm{ns}$ apart, were extracted as frames from each simulation replicate for each system. This was done to ensure independent sampling for further calculations. Secondary structure annotations were then obtained using DSSP (*Kabsch and Sander*, 1983). The conventional α helix, 3_{10} helix, and π helix were considered as one group of helices. Finally, the frequency of secondary structure formation was then obtained by averaging and normalizing by number of frames in each system. See Figure 6.4 and

118

Figure 6.5A for the output.

For rest of the methods pertaining to experimental results, please see our publication (*Chaaban et al.*, 2018).

6.5   Discussion and Conclusion

In vitro reconstitution assays showcase the highly dynamic nature of *C. elegans* microtubules (MTs) (fast growth coupled with frequent catastrophe events). Sequence analysis of *C. elegans* and canonical tubulin sequences highlight sequence divergence around lateral contact loops (Figure 6.10) leading us to hypothesize their role in different MT dynamics when compared to *B. taurus* tubulin. Indeed, MD simluations show secondary structure formation in lateral contact loop H1-S2 (Figure 6.5), which is unstructured (and hence, unresolved) in structures of mammalian MTs (Figure 6.7C-E). CryoEM imaging confirmed additional density and structuring of this H1-S2 loop (Figure 6.7A), in addition to observation that the same loop is also structured in MTs from *S. cerevisiae* (Figure 6.7B) which also display higher on-rates (*Geyer et al.*, 2015; *Howes et al.*, 2017).

Temperature-dependent reconstitution assays carried out by our collaborators show that the *C. elegans* dimer has a higher free energy in solution. In terms of energetics, the longitudinal bond energy between tubulin subunits in lattice ranges between $-16 \sim -25kT$, while the lateral bond energies are estimated to be between $-4kT \sim -12kT$ (*VanBuren et al.*, 2002, 2005; *Gardner et al.*, 2011; *Castle and Odde*, 2013; *Kononova et al.*, 2014). Entropically, both these lateral- and longitudinal-interactions would release water molecules (solvation shell) from the incoming tubulin heterodimer Both our collarborator's experiments and prior studies (*Fygenson et al.*, 1994) have shown that the MT the growth rate increases with temperature, that is, MT growth is entropically driven. Our MD simulation studies and temperature-dependent reconstitution assays, comparing *C. elegans* tubulin to canonical mam-

119

malian tubulin, indicates that the "pre-ordering" of later contact loops in *C. elegans* reduces the entropic penalty for constraining the dimer and its disordered loops into the lattice (Figures 6.5, 6.8, and 6.11). Thus, we show that the faster growth rates of *C. elegans* can be explained by an activated dimers, with pre-ordering of later contact loops as a possible rate-limiting step in MT growth. We note that in our studies, we did not consider the divergent C-terminal tails (CTTs) in our MD simulations. Whether the CTTs have a dynamical role in *C. elegans* MT polymerization, or their role is simply related to the tubulin code (*Verhey and Gaertig*, 2007; *Janke*, 2014; *Yu et al.*, 2015; *Sirajuddin et al.*, 2014), remains to be uncovered.

CHAPTER VII

Dynamics of Tubulin Heterodimer

7.1 Abstract

Microtubules are dynamic biopolymers composed of stacks of structurally homologous α- and β-tubulin heterodimers, cylindrically arranged in about 10 to 16 protofilaments. They play a crucial role in the cytoskeleton of the cell, providing structural support as well as facilitating cellular transport via molecular motors, beating of cilia and flagella, and separation of chromosomes during the cell cycle. Microtubule (MT) polymerization is central to their biological functions. However, the question—why does GTP-tubulin polymerize?—remains a fundamental one in the MT field. With the advent of CryoEM resolution evolution, a large number of structures of microtubules from divergent sources, polymerized under various conditions and with different protein binding partners are now becoming accessible. Coupled with existing large number of X-ray crystal structures, a detailed study of the inter-conformer relationships of microtubule and tubulin dimer structures is now possible. In this chapter, I briefly discuss our efforts in characterizing the conformational dynamics of tubulin heterodimer using principal component analysis and molecular dynamics simulations.

7.2 Introduction

In this chapter, I will mention some highlights of our work in characterizing the conformational dynamics of tubulin heterodimers. Most of the work done in this area in the laboratory has been in the form of prelimiary analysis for grant support, and for setting up a framework for comparing future work, such as dynamics of tubulin heterodimer under posttranslational modifications, interactions of carboxy-terminal tails with the tubulin dimer, and effect of disease-associated mutations such as those encountered in tubulinopathies (*Bahi-Buisson et al.*, 2014).

Microtubules (MT) play a crucial role in the cytoskeleton of the cell, with microtubule polymerization forming the basis to their biological function. However, we still do not fully understand what gives rise to the varying rates of MT polymerization, both *in vivo* and *in vitro*. Tubulin subunits, α- and β-tubulin, are one of the most conserved family of proteins. One of the lowest α-tubulin homology ($\sim 62\%$) is between a very divergent *Drosophila* α-tubulin and an α-tubulin from the yeast *S. cerevisiae*, and one of the lowest β-tubulin homology ($\sim 63\%$) is between a yeast (*S. cerevisiae*) β-tubulin and a mouse β-tubulin (*Little and Seehaus*, 1988). In contrast, some mammalian tubulins are almost identical. While we have shed light on mechanisms behind different MT polymerizaton rates of nematode and mammalian tubulin (nematode MTs polymerize three times as fast as mammalian tubulin even though the tubulin subunits share $\sim 88\%$ sequence identity, see *Chapter VI*), we still do not fully understand how subtle changes in sequence give rise to considerably different MT dynamics. Understanding the intrinsic dynamics of tubulin heterodimers, the basic building blocks of MTs, is the first step in answering such questions. Here, I will briefly mention efforts in understanding some of the intrinsic dynamics of mammalian tubulin heterodimer as the prototypical tubulin dimer. We combine principal component analysis of experimentally available structures, molecular dynamics simulations, and subsequent analyses to better understand the intrinsic dynamics of

tubulin heterodimer.

7.3 Results

7.3.1 Principal component analysis of experimental structures

In total, 127 tubulin dimers, mostly from mammalian sources, were extracted from experimentally available strucutures from RCSB protein data bank (*Berman et al.*, 2002). For pre-processing and list of structures selected, see *Methods* and Appendix E, Table E.1. Principal component analysis (PCA) was used to examine the major conformational differences between the structures of tubulin dimers. $\sim 79\%$ of the variance in atomic positional displacements was captured by the first principal component (PC1), with PC2 accounting for $\sim 5\%$ of the variance. The first two PCs, hence, provide an informative description of the conformational space spanned by structures of tubulin dimers and that of dimers within a microtubule lattice. The original 127 structures of tubulin dimers were then projected onto the first two PCs (PC1 and PC2) to provide a low dimensional representation, hightlighting the differences between the structures (Figure 7.1A).

The major feature described by the PC1 is the concerted displacement of N-terminal and intermediate domains (I-domain), akin to "breathing" motion of the tubulin heterodimer. The N- and I-terminal domains comprise of GTP-binding elements and lateral contact loop H1-S2, and taxol-binding and lateral contact loop M-loop (*Nogales et al.*, 1998; *Löwe et al.*, 2001), respectively. The helix β-H10 and strand β-S9 also form intra-dimer contacts with the α-tubulin. The PCA maximizes the separation along this linear coordinate, separating the 127 tubulin structures used into: 1) a "bent" conformation, comprising primarily of X-ray structures of 1-2 tubulin heterodimers in association with MAPs such as stathmins, tubulin-tyrosine ligase (TTL), and DARPins; and 2) "MT lattice-like" conformation, comprising primarily
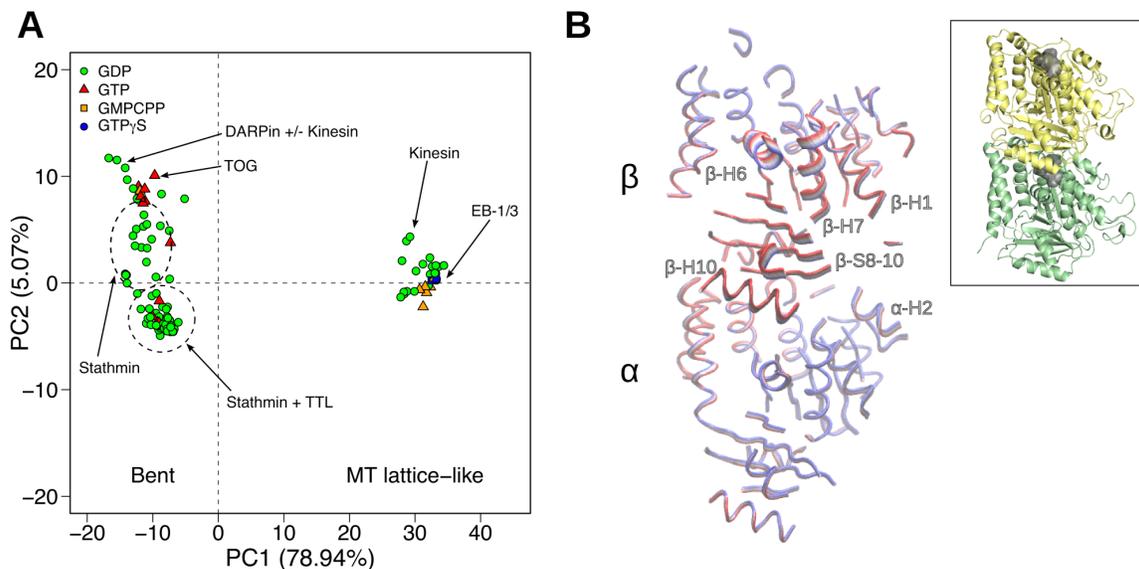
Figure 7.1: Principal component analysis of tubulin dimers. (**A**) Projection of experimental structures of tubulin dimers on first two, most significant PCs (PC1 and PC2). Structures are colored by the nucleotide state of β-tubulin, GDP (green), GTP (red), GMPCPP (orange), GTPγS (blue). Dashed ovals present groups of structures with the same microtubule associated protein(s) (MAPs), indicated by arrow and accompanying text. Two major groupings obtained from the PCA are marked as "bent" and "MT lattice-like". Percentage of total variance captured by corresponding PC is marked along the axes. (**B**) Interpolation between the most dissimilar structures in the distribution along PC1, colored red for "straight" and blue for "bent" structures, capturing a "breathing" motion of the tubulin dimer. Inset, cartoon representation of tubulin dimer as a reference orientation for PC1 trajectory shown in (*B*).

of CryoEM and electron crystallography structures of microtubules or zinc-stabilized tubulin sheets. We note that the conformational clustering obtained from PCA does not coincide with the nucleotide state of β-tubulin. A separate PCA analysis of microtubule structures in different nucleotide states may reveal differences in nucleotide-associated conformational differences, such as a lattice compaction induced by GTP hydrolysis (*Alushin et al.*, 2014; *Zhang et al.*, 2015). The current PCA highlights differences in the intra-dimer bend of the tubulin dimer, and provides a conformational map for comparing differences impacting the intrinsic dynamics of the tubulin heterodimer, such as nucleotide states (*Gebremichael et al.*, 2008; *Grafmüller and Voth*,

2011; *Grafmüller et al.*, 2013; *Igaev and Grubmüller*, 2018) and sequence divergence (Chapter VI) (*Chaaban et al.*, 2018).

7.3.2 Nucleotide-dependent conformational sampling

Replicate MD simulations of GDP-bound (GDP-tubulin) and GTP-bound (GTP-tubulin) tubulin heterodimers were performed, for a total of 3.6μs and 2.7μs respectively, to characterize their intrinsic dynamics (see *Methods*). Projection of the simulation trajectories onto the PCs determined from analysis of experimental structures (Figure 7.1) was used to evaluate the conformational space sampled (Figure 7.2). Simulations of both the GDP- and GTP-tubulin were initiated from a relatively "straight" conformation (see *Methods*). However, it is apparent from these projections that each set of simulations also sampled a slightly bent conformation, and also overlap in the conformational space sampled. While the GTP-tubulin sampled two distinct regions: one near the MT lattice-like straight conformation and another bent conformation, the GDP-tubulin was restricted in its sampling near the MT lattice-like starting conformation. A recent study by (*Igaev and Grubmüller*, 2018) showed similar results, with GDP-tubulin more confined in its sampling. This study also estimated the free energy of dimer kinking (ΔG_{kink}) stored in GTP- and GDP-tubulin, estimating these to be $\sim 2.0\,kT$ for GTP-tubulin, and $\sim 6.6\,kT$ for GDP-tubulin, suggesting that it costs less free energy for GTP-tubulin to adopt the straight MT conformation (*Igaev and Grubmüller*, 2018).

Additional sets of MD simulations of the nucleotide-free state of the β-tubulin (apo) were also carried out for a total of 1.5μs. Starting structure for MD simulations of the apo-state was obtained by first removing the GDP and Mg^{2+} ion from the E-site of β-tubulin in the equilibrated structure from GDP-tubulin simulations. Projection of apo-state trajectories reveal a restricted sampling around the starting conformation (Figure 7.2B). Although the free energy of kinking stored in apo-tubulin is not known,
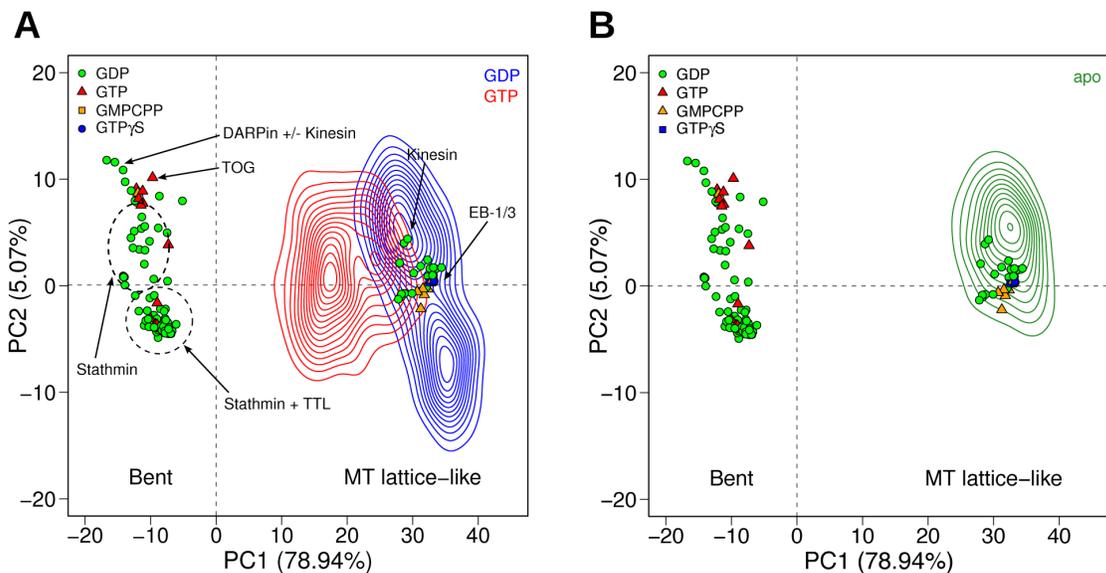
Figure 7.2: Analysis of MD simulations of tubulin dimers. Projection of trajectories from MD simulations from (**A**) GDP- and GTP-tubulin, and (**B**) apo-tubulin onto the first two PCs (PC1 and PC2) defined by experimental structures (see Figure 7.1). Projections shown are contours and colored according to β-tubulin nucleotide state, GDP (blue), GTP (red), apo (green). Experimental structures annotated as in Figure 7.1.

we can hypothesize it to be even larger than that of GDP- and GTP-tubulin (*Igaev and Grubmüller*, 2018).

7.3.3 Tubulin intra-dimer bend angles

The tubulin intra-dimer rotation has been extensively studied before, both in terms of free-energy associated with intra-dimer rotation (*Peng et al.*, 2014), and nucleotide-dependent intra-dimer bend (*Gebremichael et al.*, 2008; *Grafmüller et al.*, 2013) and twist (*Grafmüller and Voth*, 2011). A recent study has estimated a high free energy barrier of $\sim 8.6\,kT$ between the two most common intra-dimer motions observed—bending orthogonal to the MT wall (splay), and twisting of β-tubulin relative to the α-tubulin subunit (twist)—in GTP-tubulin (*Igaev and Grubmüller*, 2018). However, the study did not report the range and distribution of the bend and twist angles. Our approach to calculating these angles was slightly different than

126

those applied previously (which used least squares fit between tubulin domains or those between the core helix H7 (*Peng et al.*, 2014)). For a given tubulin dimer, we first aligned the α-subunit along the MT wall. Then, the β-subunit was aligned onto the α-subunit using the Kabsch algorithm (*Kabsch*, 1978), using equilivalent Cα residues used in the PCA discussed previously (see Figure 7.3A-D). The rotation matrix of this routine was extracted using the Bio3D v2.3-3 package (*Skjærven et al.*, 2014). This rotation matrix was further decomposed into corresponding Euler angles (*Arfken*, 1985). The rotation angles of selected structures, along with comparison to angles reported in (*Peng et al.*, 2014) are listed in Table 7.1. The reference orientations of the decomposed angles, R_x, R_y, and R_z, are shown in Figure 7.3E. We note that the spurious high values of rotation reported for PDB ID 4EB6 (dimer chains A,B and C,D) by (*Peng et al.*, 2014) in Table 7.1 are due to partial unfolding of core helix H7, which is used for measuring the intra-dimer rotation in their study.

Table 7.1: Tubulin intra-dimer bend angles. Intra-dimer bend angles, calculated from rotation matrix component of fitting β-subunit onto α-subunit of the same dimer are reported in degrees (°). PDB IDs, the respective chain IDs of the tubulin dimer, and associated MAP(s) are listed.

PDB ID	Chains	Nucleotide	PPI	Angle (°)	R_x(°)	R_y(°)	R_z(°)	Brouhard & Rice, 2014	Peng et al., 2014
1TVK	A,B	GDP	–	0.85	0.8	0.19	0.22	–	1.9
4I4T	A,B	GDP	Stathmin, TTL	9.71	2.79	2.41	8.92	–	10.5
4I4T	C,D	GDP	Stathmin, TTL	8.93	3.31	3.54	7.4	–	5.8
4IHJ	A,B	GDP	Stathmin, TTL	9.46	2.68	2.77	8.58	–	9.7
4IHJ	C,D	GDP	Stathmin, TTL	8.74	3.33	3.42	7.23	–	6.1
4I55	A,B	GDP	Stathmin, TTL	9.57	3.04	2.81	8.55	–	9.7
4I55	C,D	GDP	Stathmin, TTL	9.01	3.62	3.53	7.36	–	6.1
3RYC	A,B	GDP	Stathmin	10.14	1.37	8.15	5.99	–	6.9
3RYC	C,D	GTP	Stathmin	12.02	2.46	7.18	9.48	–	6.2
4DRX	A,B	GTP	DARPin	11.36	8.28	7.57	1.37	11	6.8
4DRX	C,D	GTP	DARPin	11.95	9.16	7.58	0.73	–	6.8
4FFB	A,B	GTP	TOG	12.76	6.44	10.72	2.04	13	–

PDB ID	Chains	Nucleotide	PPI	Angle (°)	R_x(°)	R_y(°)	R_z(°)	Brouhard & Rice, 2014	Peng et al., 2014
4I50	A,B	GDP	Stathmin, TTL	9.5	2.55	2.79	8.66	–	10
4I50	C,D	GDP	Stathmin, TTL	8.72	3.47	3.59	7.04	–	6.1
3RYI	A,B	GDP	Stathmin	10.2	1.33	8.36	5.79	–	11.7
3RYI	C,D	GDP	Stathmin	12.55	2.69	7.46	9.91	–	6.3
3RYF	A,B	GTP	Stathmin	9.96	0.99	7.98	5.94	–	10.6
3RYF	C,D	GTP	Stathmin	12.28	2.13	7.59	9.56	–	6.1
3UT5	A,B	GDP	Stathmin, Vinca	10.96	0.38	9.69	5.13	–	7.7
3UT5	C,D	GDP	Stathmin, Vinca	11.14	1.64	6.19	9.21	–	7.7
3RYH	A,B	GMPCPP	Stathmin	9.68	0.45	7.78	5.77	9.5	10.3
3RYH	C,D	GMPCPP	Stathmin	12.45	1.93	7.32	10.02	–	6.3
4EB6	A,B	GDP	Stathmin	10	1.57	8.83	4.54	–	20.4
4EB6	C,D	GDP	Stathmin	11.44	2.6	6.68	9.07	–	31.7
1SA0	A,B	GDP	Stathmin	12.5	4.84	9.25	7.3	13	12
1SA0	C,D	GDP	Stathmin	11.04	2.42	8.81	6.38	–	9.2
3HKB	A,B	GDP	Stathmin	12.67	5.53	8.82	7.66	–	7.7
3HKB	C,D	GDP	Stathmin	11.11	3.1	8.17	7.09	–	16.5
1SA1	A,B	GDP	Stathmin	11.58	3.64	8.32	7.46	–	12.1
1SA1	C,D	GDP	Stathmin	10.95	1.66	8.62	6.68	–	4.8

7.3.4 Interaction of tubulin C-terminal tails

Another dynamical aspect of the tubulin heterodimer we wanted to investigate was the interaction of the tubulin C-terminal tails (CTTs) with the globular core domains. CTTs are highly acidic stretches of 9–25 residues, overrepresented by glutamate residues. Their highly flexible nature has kept them elusive to structure determination through crystallography or CryoEM. Most MD simulations-based studies have excluded the CTTs in modeling. Previous studies have described the interactions of CTTs with the tubulin dimer using MD simulations (*Luchko et al.*, 2008; *Freedman et al.*, 2011), however, these studies were limited in the time-scales achieved in their simulations ($\sim$ 5ns). Here, we modelled CTTs from Bovine tubulin (α-tubulin: TBA1B, P81947 and β-tubulin: TBB2B, Q6B856) as extended stretches orthagonal
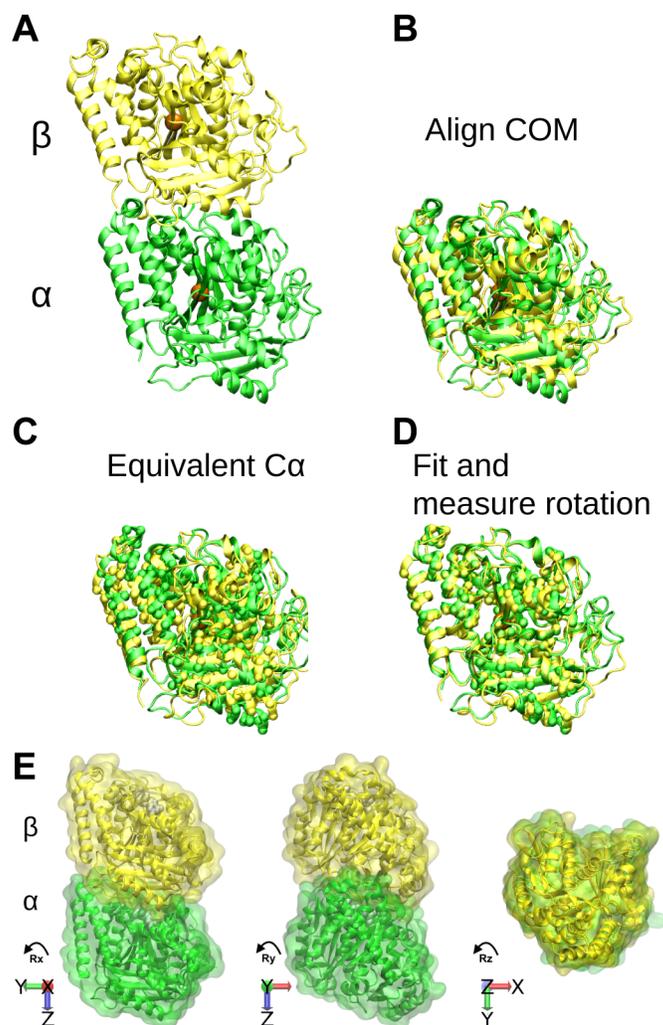
Figure 7.3: Measuring tubulin intra-dimer angles. (**A-D**) Steps for aligning and fitting tubulin dimer subunits, followed by measuring rotation angles using the Kabsch algorithm (*Kabsch*, 1978). Center of mass (COM) of each subunit shown in orange, with equivalent Cα atoms used for superposition shown as mini-spheres. α-subunit in green, β-subunit in yellow. (**B**) Orientation of tubulin dimer, shown in cartoon and surface representations, relative to the MT wall, in reference to angles listed in Table 7.1. Tubulin subunits colored as per (*A*).

to the MT wall-forming helices H11 and H12 of the tubulin heterodimer. MD simulations were carried as described for other systems (*Methods*), for a total of 2.75μs in GDP-state (GDP-CTT), 1.5μs in GTP-state (GTP-CTT), and 1μs for detyrosinated tubulin in GDP-state (tyrosine 451 of α-tubulin was removed, GDP-deTyr).

Although CTTs are disordered in structural studies, we observe transient forma-

tion of secondary structure elements in our long-timscale MD simulations (Figure 7.4). We also characterized the intra- and inter-subunit contacts of the CTTs in each state. Custom *Tcl* scripts in VMD (*Humphrey et al.*, 1996) were used to calculate contact frequencies of residues for each state, with a residue-residue contact defined as any heavy-atoms within 5Å of each other. The CTTs in GDP-CTT simulations primarily formed intra-subunit contacts with positively charged residues (Figure 7.5A and B). Furthermore, residues V440-E443 of α-CTT in the GDP state also interacted with loop H10-S9 of α-tubulin as a β strand. Loop H10-S9 and the adjacent helix H10 of the α-tubulin are the inter-dimer interface, interacting longitudinally with the β-tubulin of the next dimer in the protofilament. CTT interactions with this region could act as a transient steric block in the incoming tubulin dimer during polymerization. The wildtype CTT (WT-CTT) in GTP-CTT and deTyr-CTT in GDP-deTyr also interact with loop H10-S9 and helix H10, however, not as a β strand (Figure 7.4B and C). Interestingly, the β-CTT in GTP state makes extensive contacts with helices H11 and H12 from both the tubulin subunits (Figure 7.5D), suggesting possibility of nucleotide-dependent dynamics of the β-CTT. However, we note that even at these timescale (microsecond), the complete conformational sampling and interactions of CTTs are not covered, possibily due to kinetic trapping because of the highly charged nature of the CTTs and their ability form strong ionic interactions with charged residues on the surface of tubulin subunits (*Debiec et al.*, 2014).

7.4 Methods

7.4.1 Principal component analysis

A set of 350 experimental structures were obtained from the RCSB Protein Data Bank (*Berman et al.*, 2002) after a HMMER search (*Eddy*, 2011) of sequences (α-tubulin: TBA1B, P81947 and β-tubulin: TBB2B, Q6B856) against the PDB database

130
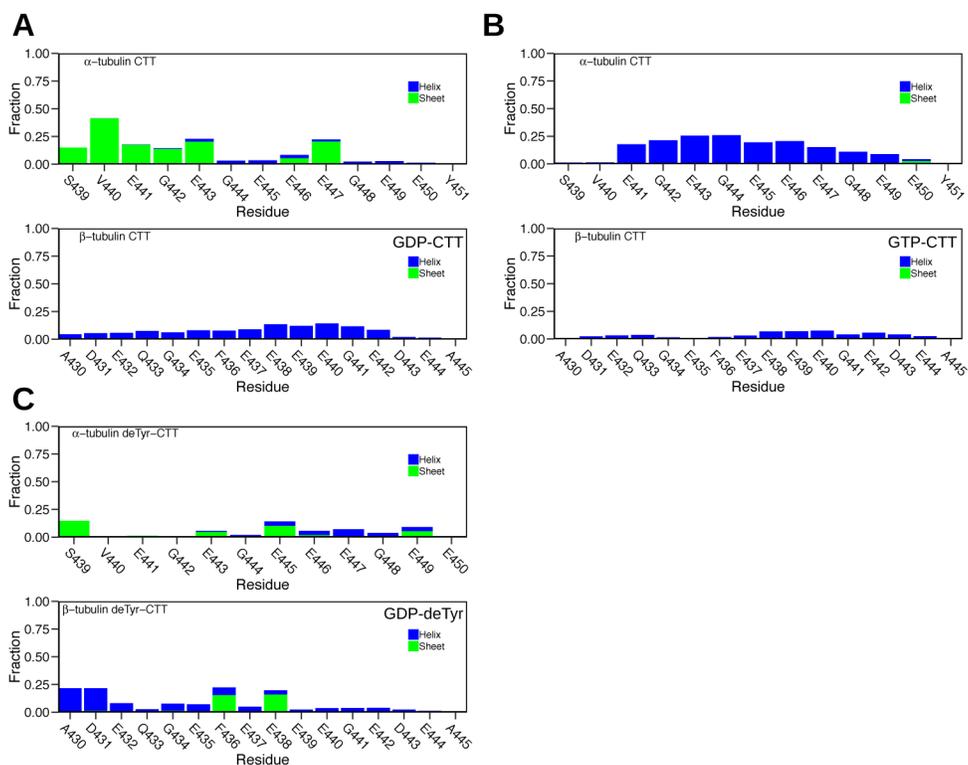
Figure 7.4: Secondary structure formation in tubulin CTTs. Secondary structure formation in CTTs for (**A**) GDP-state, (**B**) GTP-state, and (**C**) deTyrosinated CTT in GDP-state. Formation of secondary structure elements was determined with DSSP (*Kabsch and Sander*, 1983), with atleast two consecutive residues considered for forming a β strand, and atleast three consecutive residues for a helix. Data shown as normalized frequency of secondary structure element formation, with helix shown in blue, and sheet in green.

using Bio3D-web (*Skjærven et al.*, 2016; *Jariwala et al.*, 2017). A final set of 127 crystallographic and cryo-EM structures was obtained after selecting unique dimers from structures with resolution better than 5Å (See Appendix E, Table E.1). Principal component analysis (PCA) is a dimensionality reduction technique involving orthogonal transformation of the original data into a set of linearly uncorrelated variables termed principal components. Briefly, PCA involves diagonalization of the covariance matrix C, whose elements C_{ij} are calculated from the Cartesian coordinates of Cα
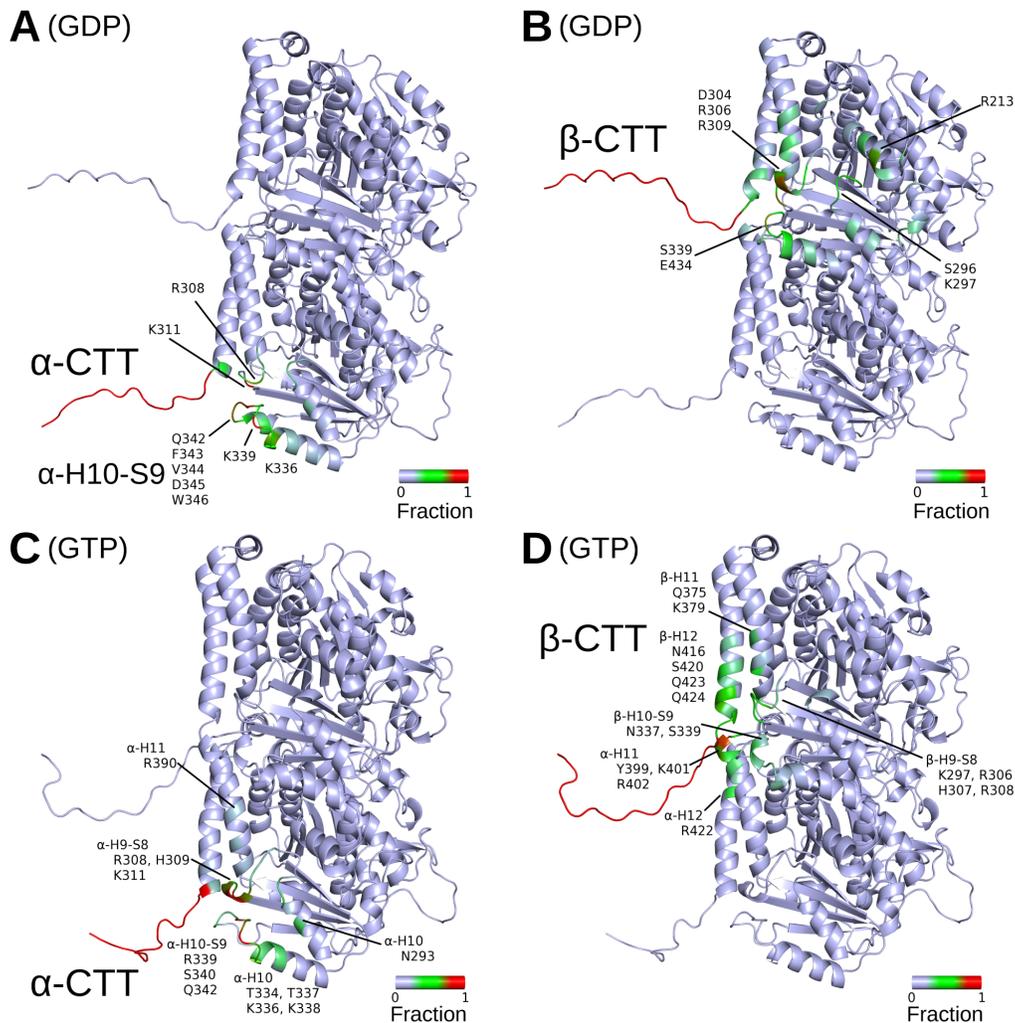
Figure 7.5: Interactions of CTTs with tubulin dimer. Contact frequency of CTT-tubulin dimer residue contacts, shown as normalized frequency in a blue-green-red scale, red implying a residue within the tubulin dimer is in contact with at least one CTT residue. Contact frequency shown for (**A**) α-CTT in GDP-state, (**B**) β-CTT in GDP-state, (**C**) α-CTT in GTP-state, and (**D**) β-CTT in GTP-state. Nucleotides hidden for clarity since CTTs did not contact the nucleotides in either subunits in GDP- or GTP-states.

atoms, r, after superposition:

$$C_{ij} = \langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \rangle \tag{7.1}$$

where i and j represent all pairs of 3N coordinates. The eigenvectors, or principal components (PCs), of the covariance matrix form a linear basis set of the distribution of structures. The variance of the distribution along each eigenvector is given by the corresponding eigenvalue. Projecting structures onto a sub-space defined by principal components with the largest variance (largest eigenvalues) provides a lower dimensional representation of the structure dataset.

PCA was performed on 425 equivalent, non-gap Cα atoms from each of the dimers after superposition (Figure 7.1). Trajectories from MD simulations were projected on to the PCA space defined by the first two PCA eigenvectors to allow comparison of the conformational space spanned by the simulations and the available tubulin structures (Figure 7.1).

7.4.2 Models of tubulin dimers for MD simulations

GDP-tubulin

Initial coordinates of tubulin dimer were taken from PDB 1JFF (*Löwe et al.*, 2001) after removing bound Taxol molecule. Missing coordinates of residues α:35-60 were modeled using MODELLER v9.18 (*Šali and Blundell*, 1993) using PDB 1TUB (*Nogales et al.*, 1998) as template. Residues were mutated, where applicable, to match sequences from UniProt (α-tubulin: TBA1B, P81947 and β-tubulin: TBB2B, Q6B856). A final model was selected from 100 models after loop refinement and evaluation with DOPE score. A Mg^{2+} ion was added to the GDP-bound β-tubulin in the final model.

GTP-tubulin

Initial coordinates of tubulin dimer were taken from the central dimer of GMPCPP-bound structure (PDB 3JAT). Missing coordinates of residues α:38-46 were modeled using MODELLER v9.18 (*Šali and Blundell*, 1993). The carbon atom linking α- and

β-phosphates was changed to oxygen. Residues were mutated, where applicable, to match sequences from UniProt (α-tubulin: TBA1B, P81947 and β-tubulin: TBB2B, Q6B856). A total of 100 models were generated with the following options in MOD-ELLER: variable target function method (VTFM) was set to *slow* with associated conjugate gradient set to 150 iterations; MD with simulated annealing option was set to *slow*; and the entire optimization process was repeated twice. The top-scoring model was selected with discrete optimized protein energy (DOPE) score (*Shen and Sali*, 2006) for loop refinement. The top-scoring model was selected for MD simulations.

7.4.3 Molecular dynamics simulations

Energy minimization and molecular dynamics (MD) simulations were performed with AMBER 18 (*Case*, 2018) and the ff99SB AMBER force field (*Hornak et al.*, 2006). Nucleotide parameters were obtained from (*Meagher et al.*, 2003). Histidine protonation states were assigned based on the their pKa values calculated by PROPKA (*Li et al.*, 2005). The simulation setup and procedures were adopted as described previously (see *Methods* section in *Chapter VI*). MD simulations were started from equilibrated structures with at least four independent runs of at least 500 ns each. All simulations were performed in-house on NVIDIA GPU cards with the GPU version of PMEMD (*pmemd.cuda*). We thank NVIDIA for their gift of GPU card through their Academic GPU seed grant. Trajectory analyses were carried out in R using the Bio3D v2.3-3 package (*Skjærven et al.*, 2014).

7.5 Discussion and Conclusion

We have combined principal component analysis (PCA) of experimentally available structures, molecular modeling, and molecular dynamics (MD) simulations to set up a conformational map for further investigations into the dynamics of tubulin het-

erodimer. The PCA describes inter-conformer relationships of available experimentally determined structures of tubulin dimer. This conformational space (PC-space) describes the relationship between straight "MT lattice-like" and bent conformations of tubulin dimers (Figure 7.1). We note that the conformational clustering obtained from PCA do not coincide with the nucleotide state of β-tubulin, similar to observations in myosin (*Grant et al.*, 2010), kinesin (*Scarabelli and Grant*, 2013), and F_1ATPase (*Okazaki and Takada*, 2011). Structures selected for PCA include both free dimers and dimer-of-dimers in association with large MAP domains, as well as microtubules in mixed nucleotide states (GDP, GMPCPP, and GTPγS). The conformational variability in this structure-set is dominated by the "intra-dimer" bend, as captured by the first principal component (PC1). A separate PCA study of only microtubule structures may reveal the nucleotide state-associated conformation differences, such as the MT lattice compaction upon hydrolysis (*Alushin et al.*, 2014; *Zhang et al.*, 2015). With the increasing availablility of structures of microtubules and tubulin dimers of higher orders from different organisms, isoforms of same species, with different binding partners and in different nucleotide states, a thoughrough investigation into structural relationships within protofilaments is also warranted.

When comparing the intrinsic dynamics of the free tubulin heterodimers in solution, we observe that tubulin heterodimer in both GDP- and GTP-states sample overlapping conformations. Projection of trajectories of GTP- and GDP-state tubulin dimer highlights the higher conformational variability of GTP-tubulin, with the GDP-tubulin restricted in its sampling (Figure 7.2A). This result has been recapiculated in other work and extended to describe a lower free energy of kinking for GTP-tubulin ($\Delta G_{kink}^{GTP} \sim 2.0\,kT$ versus $\Delta G_{kink}^{GDP} \sim 6.6\,kT$), suggesting that it costs less free energy for GTP-tubulin to adopt the straight MT conformation (*Igaev and Grubmüller*, 2018). Further investigations, however, are required in characterizing the intra-dimer rotation and twist angles along the first two principal components

(PC1 and PC2). Our approach of decomposing angles obtained from the rotation matrix of superimposing the β-subunit onto the α-subunit (Table 7.1 and Figure 7.3) do not form a one-to-one correspondence with PC1 and PC2 of the unsupervised PCA.

Finally, we have also combined molecular modeling and MD simulations to describe interactions of the highly acidic C-terminal tails (CTTs) with the tubulin dimer. While the α-CTT, in both wildtype (WT) and detyrosinated (ΔY451) α-tubulin, forms comparable interactions in GDP- and GTP- states, there are differences in the frequency of secondary structures adopted (Figure 7.4). Of note are the interactions of α-CTT with loop S9-H10 of the α-subunit (Figure 7.5), which along with helix H10, longitudinally contacts the β-subunit of the neighbouring dimer. There are also differences in the interactions of β-CTT in GDP- and GTP-state. While β-CTT in GDP-state primarily contacts charged residues in intermediate- and C-terminal domain of the same β-subunit, β-CTT in the GTP-state forms extensive interactions with helices H11 and H12 in both the tubulin subunits, suggesting nucleotide-dependent differences in interactions of the β-CTT (Figure 7.5B and D). In summary, characterizing the intrinsic dynamics of tubulin dimer with CTTs as well as the interactions of the CTTs with the tubulin dimer will help provide a deeper understanding of the "tubulin code" (*Janke*, 2014; *Yu et al.*, 2015; *Sirajuddin et al.*, 2014), and their functional significance in cells.

CHAPTER VIII

Concluding Remarks and Future Directions

8.0.1 Open science considerations

In Chapter II, I discussed the development of Bio3D-web, an online application for analyzing the sequence, structure and conformational heterogeneity of protein families (*Skjærven et al.*, 2016; *Jariwala et al.*, 2017). A wide variety of bioinformatics tools and online servers are already available for exploring biomolecular structures, performing pairwise structural alignment, and biophysics based tools including various normal mode analysis servers (*Eyal et al.*, 2015; *Suhre and Sanejouand*, 2004; *Tiwari et al.*, 2014). However, detailed quantitative analysis of protein dynamics require the use of local tools, often requiring computational expertise and hence, accessible only to expert users with relevant programming skills. For example, Bio3D, a package extensively used throughout this dissertation requires knowledge of R, Maven requires Matlab (*Zimmermann et al.*, 2011), and ProDy (*Bakan et al.*, 2011) requires knowledge of Python. Furthermore, many tools, while free to use and accessbile online as webservers, do not have their source-code available. This limits their usage to use cases designed by their author(s), and prevents researchers from modifying the tools and incorporating them in their local analyses pipelines. To this end, we developed Bio3D-web, a web application that implements a complete workflow for user customized investigation of protein sequence-structure-dynamic

relationships. Bio3D-web requires no programming knowledge and thus decreases the entry barrier to performing advanced comparative sequence, structure and dynamics analysis. An online, free to use, and interactive server is made available at `http://bio3d.ucsd.edu/pca-app`. Furthermore, we have made the source-code available under GPLv2 license for modification, or for incorporating in local analyses pipelines at `http://bitbucket.org/Grantlab/bio3d`. The accessbile, interactive, and reproducible feature of my application has not only helped researchers, it is also used for classroom teaching and workshops—for example by Dr. Patrick Fleming at the Biophysics Department at Johns Hopkins University for teaching undergraduate courses. We hope that researchers continue to make their tools freely accessible since code-sharing and open-source licensing are essential for continued progress in this domain.

8.0.2 Improving ensemble distance difference analysis

Analysis of residue-residue distances have been used for studying functional dynamics of biomolecules, both in form of distances matrices (*Elber and Karplus*, 1987) and in terms of changes in residue contact maps generated from such distance matrices (*Doshi et al.*, 2016). Instead of averaging residue-residue differences or converting them into contact maps (adjacency matrices) by use of a cutoff, we modify longrange distances before comparison between systems by applying a smooth function. The significance of residue-residue distance differences was then evaluated with the Wilcoxon test (see Chapters III, IV, and V). We note, however, that this approach of potentially identifying allosteric residues, or highlighting residues critical for protein function, can be further improved. Firstly, the Wilcoxon test can compare only two groups (such as wildtype versus a mutant state of a protein). A Kruskal-Willis test, another non-parametric test, can first be applied for comparing multiple states, as it can accomodate more than two groups (*Kruskal and Wallis*, 1952; *Kruskal*, 1952).

138

Once we know that there is a significant difference between groups, further post hoc tests can be applied to find wich pairs of groups are different. Furthermore, corrections for multiple testing are also necessary for simultaneous tests (*Kutner et al.*, 2005), as is the case when comparing residue-residue distance differences of proteins. This can be achieved by application of Bonferroni correction, or by controlling the false discovery rate (*Benjamini and Hochberg*, 1995).

8.0.3 Combining computational predictions with experiments, or: How I learned to stop worrying and love kinesin motors

In Chapter III, Chapter IV, and Chapter V, I analyze the dynamics of three representative families of kinesin motor proteins. I identify residues critical for force production in kinesin-1, and characterize the effects of a posttranslational modification (PTM) and of disease-associated variants on kinesin-5 and kinesin-3 motors, respectively. In all three studies, the predictions made from computational analyses were directly transferable to experiments *in vitro*. In all three studies, however, the results under physiological conditions were unexpected. In the case of kinesin-5 mitotic motor Eg5, acetylation of a lysine residue disrupts a salt bridge between K146 (helix $\alpha 2$) and D91 (helix $\alpha 1$). Simulation studies of both acetylated K146 (K146Ac) and acetylcated mimetic mutation (K146Q) predicted that the disruption of this ionic bond increases the allosteric coupling between neck linker (NL), a critical element in kinesin force production, and the ATP binding/hydrolyzing active site. The predicted enhanced interactions within the active site led us to predict enahnced motility properties of Eg5 motors carrying the K146Q mimetic mutation. Single-molecule motility assays confirmed this prediction, with mutant motors behaving more akin to faster and more processive kinesin-1 motors (Chapter III). However, cell-based assays clarified the effect of this acetylation mimetic mutation. Enhanced coupling between NL and the active site made the mutant motors more resistant to detachment from

microtubules (MT) and increased their propensity to stall. Teams of modified Eg5 expressed in cells acted like a "brake", slowing spindle pole separation in mitotically active cells (*Muretta et al.*, 2018). In the case of kinesin-1 neuronal transport motor KIF5C, I identified residues critical for cover neck bundle formation and NL docking; assistive mechanisms for force production in kinesin motors. Analysis of MD trajectories of kinesin motor domains carrying mutantations at these critical sites revealed an allosteric effect on the active site. Both the mutant systems—N334A (NL latch-disrupting) and A5G/S8G/N334A (CNB- and NL-latch-disrupting)—showed enhanced residue-residue interactions within the active site, which led us to predict an enhancement of their motility properties. Single-molecule motility assays under unloaded conditions supported these predictions, with mutant kinesin motors exhibiting faster velocity, longer run lengths, and enhanced MT-binding. Force measurements with optical trapping experiments and cell-based transport assays, however, revealed that these mutant motors were severely crippled in their ability to generate force and to transport high-load cargo in cells (Chapter IV) (*Budaitis et al., 2019, in final revision at eLife*). Finally, in Chapter V, I study the dynamical effects of disease-associated variants and mutations in the "superprocessive" motor kinesin-3. In my analyses, an autosomal dominant mutation associated with neurological disorders, V8M in strand $\beta 1$, is predicted to impair ATPase activity of the kinesin motor domain. Motility assays support this hypothesis, showing that mutant KIF1A motors are slower and undergo more non-processive and diffuse events than wildtype (WT) motors (Chapter V). However, a recent manuscript on bioRxiv reports opposite effects of this mutation: they observed higher velocity and on-rates for KIF1A and *C. elegans* homolog unc-104 mutant motors (*Chiba et al.*, 2019). We note that our computational and experimental setup utilized a truncated version of the KIF1A motor lacking the tail domain, whereas (*Chiba et al.*, 2019) utilized full-length proteins in their study. The use of full-length kinesin introduces an additional variable of au-

toinhibition and autoregulation. All the examples listed above highlight the need for a combination of studies, including single-molecule experiments, force-measurements with optical trapping, and cell-based dispersion assays in characterizing effects of mutations or PTMs on motor function. Furthermore, a rigorous assessment of enzymes involved in PTMs are critical, since a considerable portion of experiments studying PTMs make use of PTM-mimetic mutations, which may not fully capture the size, charge distribution or dynamics of a PTM (*Barber and Rinehart*, 2018).

8.0.3.1 Dynamics of tubulin oligomers

In Chapter VI, we analyzed the highly dynamic *C. elegans* tubulin using sequence comparison, flexibility analysis, and secondary structure formation for a dimer in solution. The mechanism for *C. elegans* achieving faster growth rates was not clear (Figure 6.1E and F; apparent on-rate constant of $k_a = 9.21 \pm 0.46$ dimers·μM^{-1}·sec^{-1} for *C. elegans* versus 2.90 ± 0.17 dimers·μM^{-1}·sec^{-1} for *B. taurus*). These results were intriguing, especially since the α- and β-tubulins share $\sim 88\%$ and $\sim 89\%$ sequence identity with bovine (taken as reference vertebrate) tubulins, respectively. Comparison with canonical tubulin sequences revealed divergence hotspots in the lateral contact loops. Molecular dynamics (MD) simulations of dimers indicate that one of the lateral contact loops in *C. elegans* tubulin dimer is more ordered, potentially explaining a role in faster polymerization rates. In Chapter VII, we mapped the conformational relationship of tubulin dimer using principal component analysis (PCA). PCA of experimental X-ray and CryoEM structures revealed a nucleotide-independent intra-dimer bend, differentiated (clustered) by their conformational states: that is, bent dimers or dimer-of-dimers associated with regulatory proteins, or straight dimers part of the MT lattice. The PCA highlights this difference in the intra-dimer bend of the tubulin dimer, and provides a comformational map for comparing differences impacting the intrinsic dynamics of the tubulin heterodimer, such as nucleotide states (*Ge-*

bremichael et al., 2008; *Grafmüller and Voth*, 2011; *Grafmüller et al.*, 2013; *Igaev and Grubmüller*, 2018), sequence divergence (*Chaaban et al.*, 2018), or disease-associated mutations (*Bahi-Buisson et al.*, 2014). These analyses were all based on the tubulin heterodimer as an independent unit. However, MT functions such as serving as tracks for molecular motors, and providing structural support within the cell depend on higher-order interactions of the dimer (*Nogales*, 2000). As part of exploratory analysis into the effects of diease-associated mutations and PTMs on MTs, we modelled two-protofilament systems—eight tubulin subunits, four in each protofilament, made infinitely long by use of period boundary conditions. However, we faced issues with convergence of MD simulations of these systems (data not shown). Further analysis and equilibration treatments are required for such studies. A useful approach for studying such systems, and even larger ones such as multiple turns of a MT helix, is the use of coarge-graining (CG). CG-based strategies for simulation of biomolecules compromise on atomistic details for gain in simulation speed-up. In commonly used CG-based models such as MARTINI and AWSEM-MD, roughly four atoms are grouped into a single bead, enabling simulations of large systems at long time scales (*Monticelli et al.*, 2008; *Davtyan et al.*, 2012). Such CG-based approaches will be useful in studying the dynamical effects of PTMs and in measuring mechanical properties of such modified MTs.

APPENDICES

APPENDIX A

Additional Report for Chapter II

Bio3D-web Summary Report

Lars Skjærven, Shashank Jariwala & Barry J. Grant

March 15, 2019 (Report template updated: April 12, 2016)

Contents

Overview

This report was generated on Friday, March 15 2019 at 11:30 (EST) by the *Bio3D principal component analysis and ensemble normal mode analysis web application* (Bio3D-web, http://bio3d.ucsd.edu/pca-app) version **0.1**. For complete version information of all dependencies please see the session and software version information section below.

All included figures and report values in this document reflect those displayed in the online app with user supplied options (including actual input, graph type, clustering and similarity thresholds etc.). Further customization of analysis protocols and all resulting figures is possible with Bio3D (http://thegrantlab.org/bio3d/index.php) (*Skjærven et al.*, 2014) itself. Please see the conventional usage example section together with our collection of tutorials (http://thegrantlab.org/bio3d/tutorials) for further details about using Bio3D directly on your own computers.

We also automatically save your data as you proceed through the analysis. To revisit your session, please click the following link:

http://bio3d.ucsd.edu/pca-app/?SSUID=2019-03-15_dbb7db4169.

1 SEARCH: Structure Search Summary

User input consisted of a single PDB structure code: **1AKE**.

This structure is annotated as: **ADENYLATE KINASE (Escherichia coli)** in the RCSB PDB database (*Berman et al.*, 2002). The user selected chain for further analysis was chain id: **A**. Pfam database annotation of this chain can be found in **Table** A.**1** (*Finn et al.*, 2014) along with a simplified structure visualization in **Figure** A.**1** and structural composition log below.

Table A.1: Pfam database annotation.

ID	PFAM	Annotation	eValue
1AKE_A	ADK (PF00406.20)	Adenylate kinase	0.0
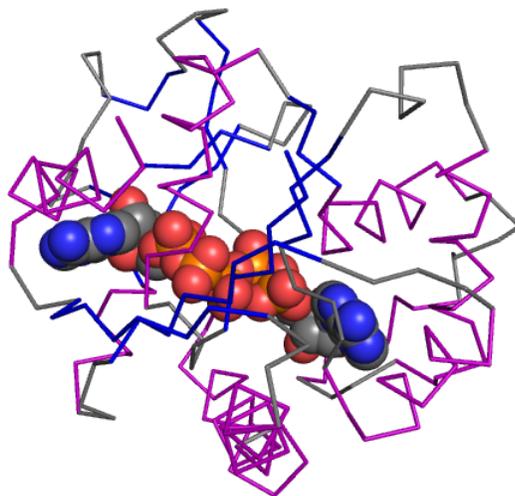
Figure A.1: PDB overview.

Input PDB composition log

```
##
##  RCSB PDB ID:1AKE
##  Selected Chain ID:A
##
##  Total Models#: 1
##  Total Atoms#: 1954
##  Chains#: 1  (values: A)
##
##  Protein Atoms#: 1656
##   (residues/Calpha atoms#: 214)
##  Nucleic acid Atoms#: 0
##   (residues/phosphate atoms#: 0)
##
##  Non-protein/nucleic Atoms#: 298
```

```
##    (residues: 242)
##  Non-protein/nucleic resid values:
##    [ AP5 (1), HOH (241) ]
```

A sequence based HMMER (3.1b2 (February 2015); http://hmmer.org) (*Eddy*, 2011) search identified **167** sequences similar hits in the RCSB PDB database. The distribution of alignment bitscores to the input sequence is shown in **Figure** A.**2**. From these hits **26** were selected for further analysis based on bitscore cutoff of **457** (encompassing **26** structures above this cutoff) and inclusion limit of **26** structures. See Table 2 for selected structures for further analysis.
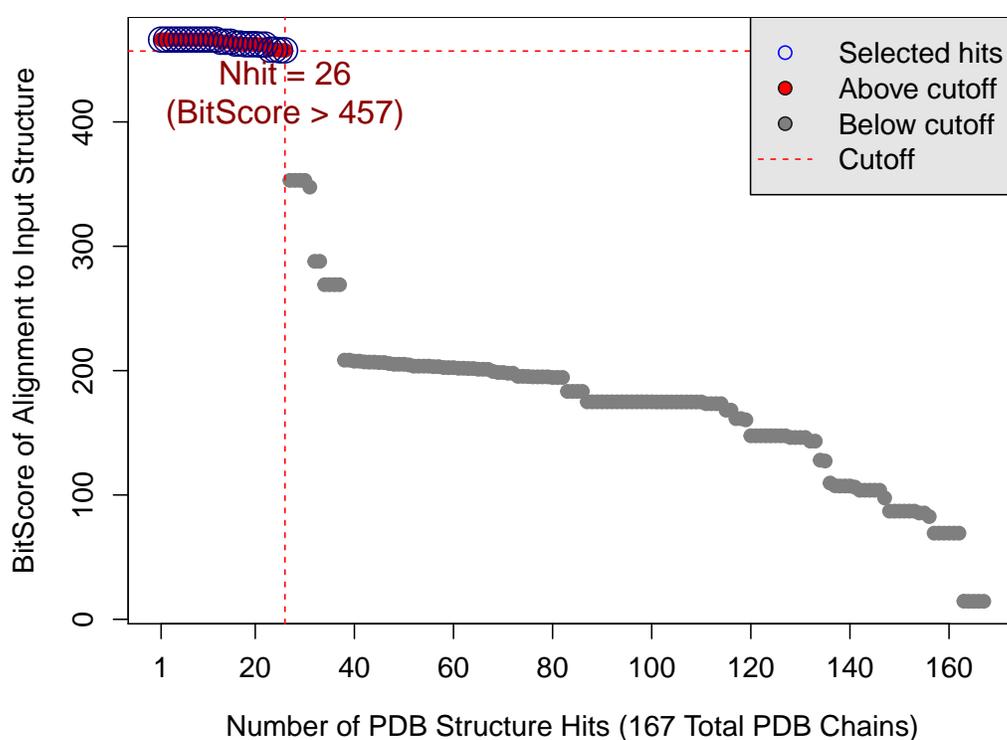
Figure A.2: Summary of search results for user input query structure sequence against the RCSB PDB chain database.

3   FIT: Structure Superposition

Structures were superposed on their **83** invariant core positions. See **Figure** A.**3**. Superposed coordinate sets and PyMol session files are available from the FIT tab of the web-app.
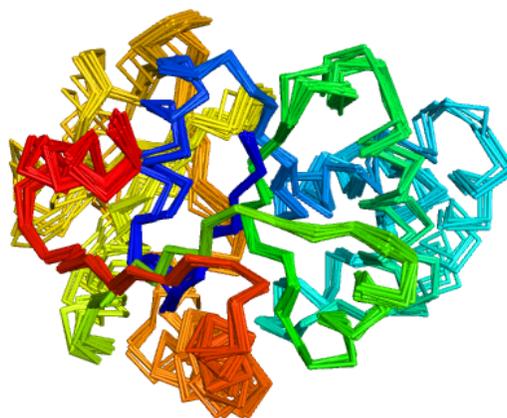
Figure A.3: Superposed PDBs colored from N-terminal (blue) to C-terminal (red) alignment position.

A hierarchical cluster analysis of these RMSD values was performed with the **ward.D2** method yielding a dendrogram that was partitioned into **2** major cluster groups. See **Figure** A.**4**.

4   PCA: Principal Component Analysis

Principal component analysis (PCA) was used to provide a lower dimensional representation of the superposed structure set that usefully summarizes inter-conformer relationships (*Grant et al.*, 2006). Applying PCA to all **26** structures revealed that **98.71%** of the total coordinate variance can be captured in three dimensions (**97.01%** in the first PC, **1.22%** in the second, and **0.48%** in the third; see **Figure** A.**5B** for a so-called scree plot or egienvalue spectrum of these values).

A projection of the structures onto **PC1** (X-axis) and **PC2** (Y-axis) (that collectively account for **98.23%** of mean square displacements in the original coordinate data) is also shown in **Figure** A.**5A**. Points represent individual structures and are colored by user defined cluster groups from either **PC subspace**, RMSD or sequence identity clustering (user selection in bold) as selected in the PCA tab of the web-app. Both sequence identity and RMSD clustering have been described separately above. PC subspace clustering is described separately further below.

Residue contributions to the first 3 PCs are shown in **Figure** A.**6**.
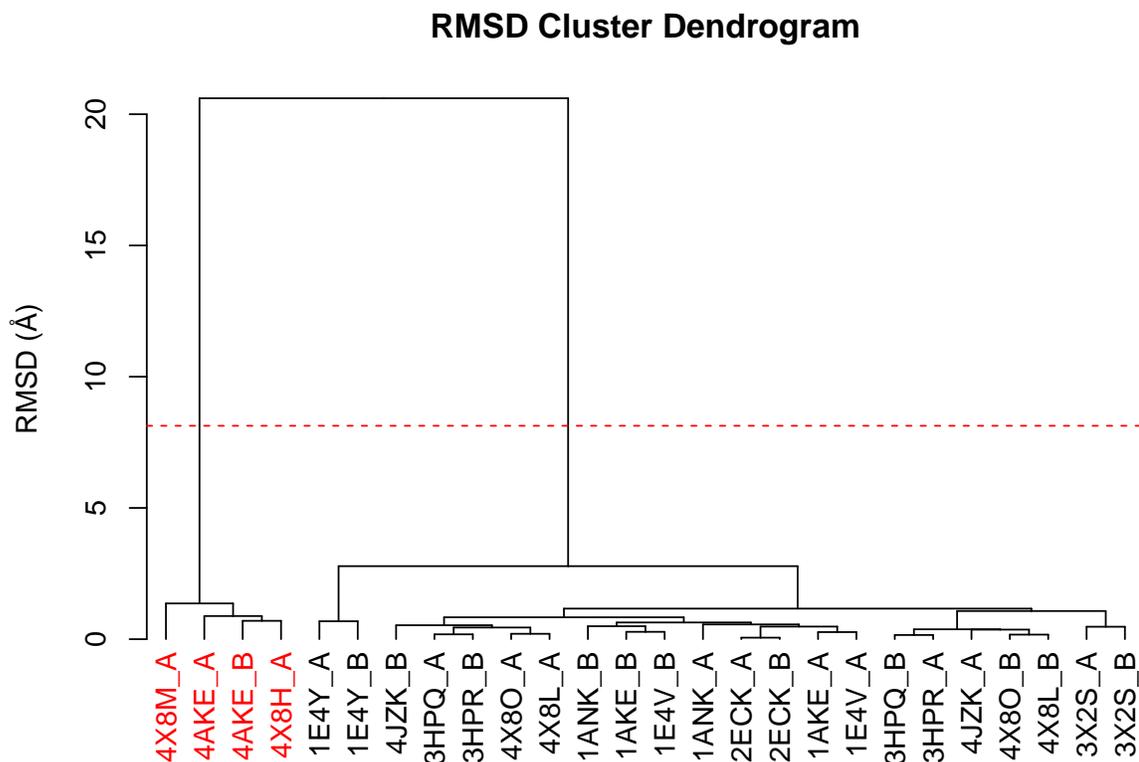
Figure A.4: RMSD Cluster Dendrogram.

Conventional Usage Example

To read your selected input structure (with PDB ID: PDBCODE) into Bio3D directly you can use the following command sequence:

```
# NOTE: Replace PDBCODE with your chosen 4 character PDB ID
library(bio3d)
pdb <- read.pdb("PDBCODE")
summary(pdb)
```

To search the online RCSB PDB database with the sequence of your query structure you could use the following commands:

```
# Use hmmer or blast
blast <- blast.pdb(pdbseq(pdb))
hits <- plot(blast)
```
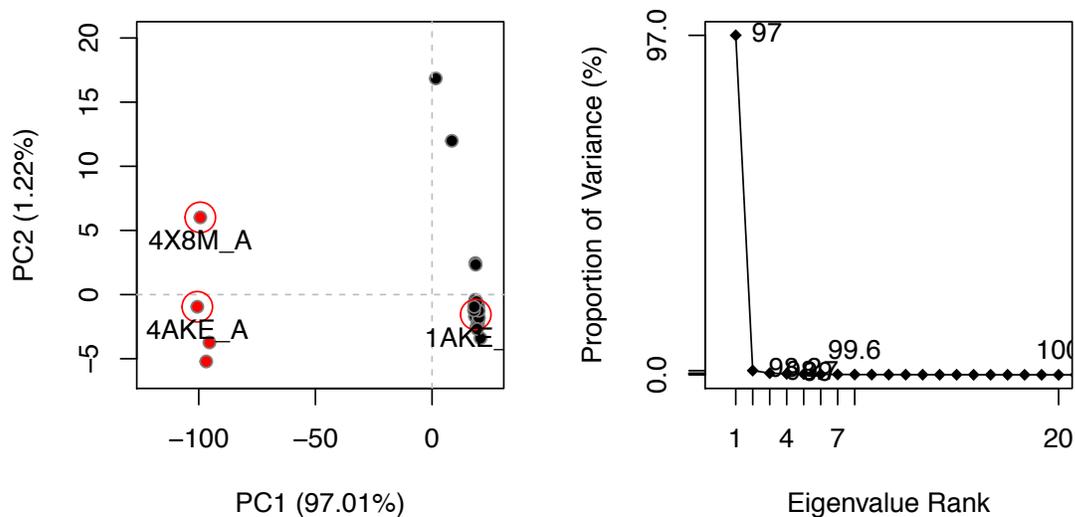
Figure A.5: Principal component analysis of structures. (**A**) Conformer plot: projection of all structures onto the principal planes defined by the user selected principal components (termed PCs). (**B**) Eigenvalue spectrum: results obtained from diagonalization of the covariance matrix of superposed coordinates. The magnitude of each eigenvalue is expressed as the percentage of the total variance (mean-square fluctuation) captured by the corresponding eigenvector. Labels beside each point indicate the cumulative sum of the proportion of the total variance accounted for in all preceding eigenvectors.

To download and align the identified structures you can use the following commands:

```
# Use the optional 'path' input argument
# to set a specific a download location
files <- get.pdb(hits$pdb.id, split=TRUE)
pdbs <- pdbaln(files)
```

For rigid core identification and structural superposition use:

```
core <- core.find(pdbs)


# Use the optional 'outpath' argument
# to write superimposed PDBs to disk
xyz <- pdbfit(pdbs, core)
```

Investigate pairwise structural deviations and perform cluster analysis with:
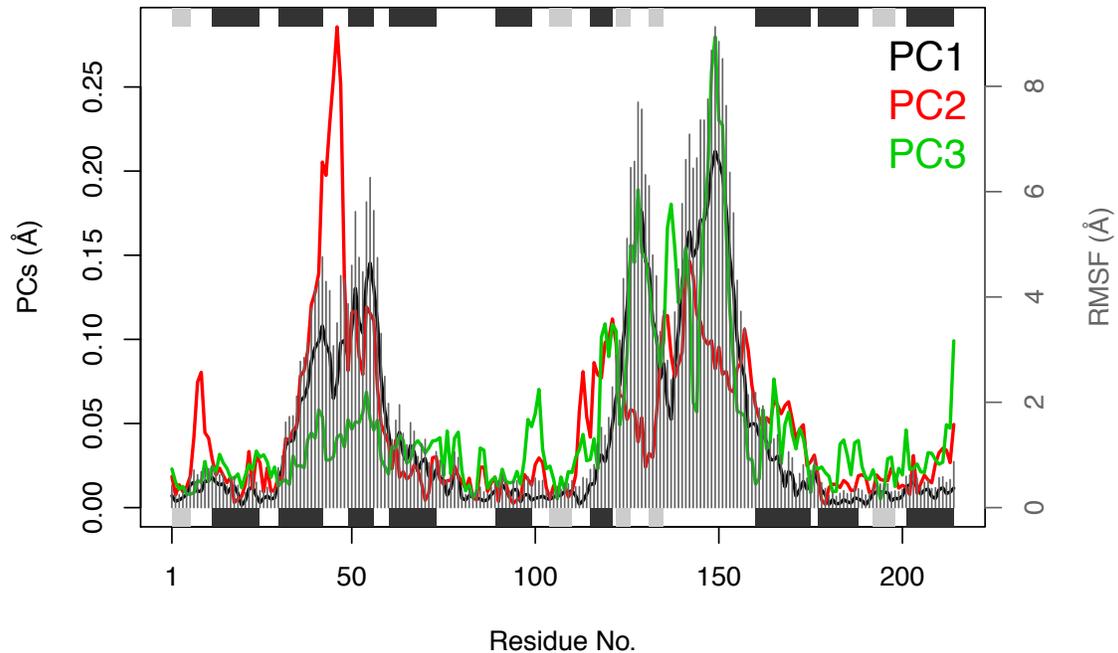
Figure A.6: The contribution of each residue to selected principal components (PCs).

```r
rd <- rmsd(xyz)
hc <- hclust(as.dist(rd))
hclustplot(hc, k=2)
```

Perform principal component analysis with:

```r
pc <- pca(xyz)
plot(pc)
mktrj(pc)
```

To perform normal mode analysis use:

```r
modes <- nma(pdbs)
plot(modes)
mktrj(modes)
```

Citation Information

This Bio3D web-app should be referenced with the URL http://thegrantlab.org/

151

bio3d/webapps and the following citations: Skjærven, L., S. Jariwala, X.-Q. Yao, and B. J. Grant (2016), Online interactive analysis of protein structure ensembles with Bio3D-web, *Bioinformatics*, (July), btw482, doi: 10.1093/bioinformatics/btw482; and Jariwala, S., L. Skjærven, X.-Q. Yao, and B. J. Grant (2017), Investigating Protein Sequence-structure-dynamics Relationships with Bio3D-web, *Journal of Visualized Experiments*, doi: 10.3791/55640.

Session and Software Version

This report was auto-magically generated by Bio3D along with the additional R packages noted below.

You can install and run **Bio3D-web** locally by following these instructions.

```
## R version 3.5.1 (2018-07-02)
## Platform: x86_64-redhat-linux-gnu (64-bit)
## Running under: CentOS Linux 7 (Core)
##
## Matrix products: default
## BLAS/LAPACK: /usr/lib64/R/lib/libRblas.so
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] knitr_1.20          data.table_1.11.8    RCurl_1.95-4.11
##  [4] bitops_1.0-6        rmarkdown_1.10       shinyBS_0.61
```

```
##  [7] abind_1.4-5           threejs_0.3.1          igraph_1.2.2
## [10] maptools_0.9-4        sp_1.3-1               reshape2_1.4.3
## [13] rgl_0.99.16           rCharts_0.4.5          lattice_0.20-35
## [16] DT_0.4                bio3d.view_0.1.0.9000 bio3d_2.3-4.9000
## [19] shiny_1.1.0
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.19          highr_0.7
##  [3] compiler_3.5.1        later_0.7.5
##  [5] plyr_1.8.4            shinyjs_1.0
##  [7] base64enc_0.1-3       tools_3.5.1
##  [9] digest_0.6.18         evaluate_0.12
## [11] jsonlite_1.5          rlang_0.3.0
## [13] pkgconfig_2.0.2       crosstalk_1.0.0
## [15] yaml_2.2.0            parallel_3.5.1
## [17] stringr_1.3.1         htmlwidgets_1.3
## [19] rprojroot_1.3-2       manipulateWidget_0.10.0
## [21] grid_3.5.1            webshot_0.5.1
## [23] R6_2.3.0              foreign_0.8-71
## [25] RJSONIO_1.3-0         pander_0.6.2
## [27] magrittr_1.5          whisker_0.3-2
## [29] backports_1.1.2       promises_1.0.1
## [31] htmltools_0.3.6       mime_0.6
## [33] xtable_1.8-3          httpuv_1.4.5
## [35] stringi_1.2.4         miniUI_0.1.1.1
```

References

Berman, H. M., et al. (2002), The protein data bank, *Acta Crystallographica Section D: Biological Crystallography*, doi: 10.1107/S0907444902003451

Eddy, S. R. (2011), Accelerated profile HMM searches, *PLoS Computational Biology*, doi: 10.1371/journal.pcbi.1002195

Finn, R. D., et al. (2014), Pfam: The protein families database, doi: 10.1093/nar/gkt1223

Grant, B. J., A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon, and L. S. D. Caves (2006), Bio3d: An R package for the comparative analysis of protein structures, *Bioinformatics*, doi: 10.1093/bioinformatics/btl461

Jariwala, S., L. Skjærven, X.-Q. Yao, and B. J. Grant (2017), Investigating Protein Sequence-structure-dynamics Relationships with Bio3D-web, *Journal of Visualized Experiments*, doi: 10.3791/55640

Skjærven, L., X. Q. Yao, G. Scarabelli, and B. J. Grant (2014), Integrating protein structural dynamics and evolutionary analysis with Bio3D, *BMC Bioinformatics*, *15*(1), 1–11, doi: 10.1186/s12859-014-0399-6

Skjærven, L., S. Jariwala, X.-Q. Yao, and B. J. Grant (2016), Online interactive analysis of protein structure ensembles with Bio3D-web, *Bioinformatics*, (July), btw482, doi: 10.1093/bioinformatics/btw482

APPENDIX B

Additional Tables for Chapter III

Table B.1: Average inter-residue distances from MD simulations. All residue pairs listed, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference between wild type (WT) and mutant (K146Q) simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop.

Residue position	WT distance (Å)	K146Q distance (Å)	K146Q − WT (Å)
K11 (CS) – Q367 (NL)	17.38 (8.57)	8.29 (4.77)	−9.09
K12 (CS) – E364 (NL)	13.36 (5.06)	6.14 2.43	−7.22
K12 (CS) – V365 (NL)	14.11 (6.87)	5.87 (2.72)	−8.24
E13 (CS) – R329 (β8)	8.44 (6.02)	10.22 (1.93)	1.78
E13 (CS) – E364 (NL)	10.55 (4.58)	5.25 (2.18)	−5.3
E13 (CS) – V365 (NL)	12.17 (5.59)	5.88 (3.2)	−6.29
E14 (CS) – R327 (L13)	10.48 (1.93)	5.76 (3.45)	−4.72
E14 (CS) – K362 (NL)	6.77 (2.19)	4.43 (1.36)	−2.34
E14 (CS) – E364 (NL)	10.31 (2.19)	5.53 (2.17)	−4.78
K15 (CS) – R329 (β8)	6.27, (3.59)	8.56, (2.37)	2.29
G16 (CS) – R329 (β8)	6.17 (2.52)	7.78 (0.98)	1.61
K17 (CS) – E92 (α1)	5.57 (1.63)	4.26 (1.55)	−1.31
Q20 (β1) – R329 (β8)	5.74 (1.71)	3.96 (1.26)	−1.78
R24 (β1) – A74 (β2b)	4.97 (1.15)	6.18 (0.39)	1.21
R26 (β1) – M229 (L9)	9.93 (2.58)	5.04 (0.63)	−4.89
F28 (α0) – A37 (L2a)	3.79 (0.56)	6.01 (2.23)	2.22

Residue position	WT distance (Å)	K146Q distance (Å)	K146Q − WT (Å)
F28 (α0) – H38 (L2a)	3.94 (0.58)	6.72 (2.16)	2.78
F28 (α0) – S39 (L2a)	4.02 (0.61)	6.03 (2.02)	2.01
F28 (α0) – P338 (L14)	3.87 (0.37)	5.12 (1.1)	1.25
N29 (α0) – M228 (L9)	8.9 (3.49)	4.37 (2.47)	−4.53
N29 (α0) – N229 (L9)	10 (2.56)	6.55 (2.54)	−3.45
A31 (α0) – L227 (L9)	13.45 (5.41)	5.38 (3.02)	−8.07
A31 (α0) – M228 (L9)	9.73 (5.19)	5.4 (2.9)	−4.33
E32 (α0) – M228 (L9)	9.41 (3.75)	5.28 (1.56)	−4.13
A35 (α0) – L341 (L14)	9.4 (5.36)	4.69 (2.02)	−4.71
A37 (L2a) – S340 (L14)	6.22 (2.89)	3.99 (1.84)	−2.23
A37 (L2a) – L341 (L14)	7.96 (4.1)	4.71 (2.43)	−3.25
H38 (L2a) – A339 (L14)	4.92 (1.69)	3.57 (0.75)	−1.35
H38 (L2a) – S340 (L14)	4.34 (1.76)	3.13 (0.85)	−1.21
K77 (α1) – P131 (L5)	4.47 (1)	5.78 (0.36)	1.31
K77 (α1) – L132 (L5)	5.45 (1.69)	7.51 (0.51)	2.06
Q78 (α1) – R138 (α2b)	4.99 (1.07)	3.83 (0.49)	−1.16
I79 (α1) – P131 (L5)	4.39 (1.31)	5.42 (0.36)	1.03
D91 (α1) – K146 (α2b)	3.56 (1.49)	4.88 (1.21)	1.32
Q106 (P Loop) – T349 (α6)	4.13 (0.51)	3.00 (0.24)	−1.13
T107 (P Loop) – T226 (L9/Sw1)	6.62 (1.2)	4.27 (0.59)	−2.35
T107 (P Loop) – M228 (L9/Sw1)	7.68 (2.07)	3.85 (0.63)	−3.83
T107 (P Loop) – N229 (L9/Sw1)	7.8 (0.95)	4.9 (0.39)	−2.9
T107 (P Loop) – S233 (L9/Sw1)	4.2 (0.73)	6.21 (0.43)	2.01
G108 (P Loop) – M228 (L9/Sw1)	8.91 (2.3)	4.36 (0.46)	−4.55
G108 (P Loop) – N229 (L9/Sw1)	7.92 (0.69)	3.81 (0.27)	−4.11
T112 (P Loop) – Y231 (L9/Sw1)	8.31 (1.26)	5.39 (0.36)	−2.92
T112 (P Loop) – S233 (L9/Sw1)	6.08 (0.5)	3.29 (0.88)	−2.79
R119 (L5) – E123 (L5)	4.35 (1.48)	5.75 (1.4)	1.4
R119 (L5) – Y125 (L5)	3.85 (1.4)	5.03 (1.46)	1.18
R119 (L5) – E215 (α3)	8.23 (3.41)	4.52 (1.67)	−3.71
W127 (L5) – E215 (α3)	5.68 (1.44)	4.11 (0.55)	−1.57
E129 (L5) – R138 (α2b)	3.76 (1.09)	4.91 (0.86)	1.15
Y164 (β4) - K280 (L11/Sw2)	7.81 (1.27)	5.06 (0.89)	−2.75
N173 (L8a) – V178 (L8a)	5.32 (1.24)	3.82 (0.94)	−1.5

Residue position	WT distance (Å)	K146Q distance (Å)	K146Q − WT (Å)
D186 (β5b) – R312 (L12)	5.49 (1.33)	3.78 (1.14)	−1.71
R192 (L8b) – R318 (α5)	4.78 (0.47)	7.27 (1.61)	2.49
R192 (L8b) – L324 (L13)	5.2 (1.89)	8.1 (2.02)	2.9
R192 (L8b) – G325 (L13)	4.77 (1.34)	7.91 (1.76)	3.14
R192 (L8b) – G326 (L13)	4.3 (0.84)	6.58 (1.09)	2.28
G193 (L8b) – Q321 (α5)	5.14 (0.66)	3.9 (0.93)	−1.24
V194 (β5c) – Q321 (α5)	5.48 (0.83)	3.92 (0.48)	−1.56
A218 (α3) – T223 (L9)	6.49 (0.68)	4.97 (0.72)	−1.52
R221 (L9/Sw1) – S233 (L9/Sw1)	4.73 (0.39)	3.47 (0.22)	−1.26
T223 (L9/Sw1) – R281 (α4)	5.1 (1.8)	6.43 (1.02)	1.33
A224 (L9/Sw1) – E284 (α4)	4.46 (0.61)	5.68 (0.65)	1.22
A225 (L9/Sw1) – S275 (L11/Sw2)	5.41 (0.9)	8.28 (0.6)	2.87
A225 (L9/Sw1) – G276 (L11/Sw2)	5.82 (1.51)	8.99 (1.01)	3.17
T226 (L9/Sw1) – R274 (L11/Sw2)	4.09 (1.2)	8.81 (0.75)	4.72
T226 (L9/Sw1) – S275 (L11/Sw2)	4.38 (0.63)	5.88 (0.8)	1.5
T226 (L9/Sw1) – G276 (L11/Sw2)	5.18 (1.01)	8.08 (0.8)	2.9
T226 (L9/Sw1) – R281 (α4)	7.46 (1.06)	5.51 (1.25)	−1.95
L227 (L9/Sw1) – E270 (L11/Sw2)	6.77 (1.3)	4.56 (1.04)	−2.21
L227 (L9/Sw1) – G273 (L11/Sw2)	4.92 (1.7)	7.89 (0.71)	2.97
L227 (L9/Sw1) – G276 (L11/Sw2)	4.84 (1.07)	6.1 (0.75)	1.26
L227 (L9/Sw1) – A277 (L11/Sw2)	5.06 (1.06)	8.27 (1.31)	3.21
M228 (L9/Sw1) – E270 (L11/Sw2)	6.67 (1.45)	4.61 (0.96)	−2.06
M228 (L9/Sw1) – L341 (L14)	7.31 (2.4)	4.3 (1.01)	−3.01
M228 (L9/Sw1) – N342 (α6)	8.62 (2.79)	4.95 (1.19)	−3.67
N229 (L9/Sw1) – S233 (L9)	6.98 (1.15)	4.94 (0.23)	−2.04
S232 (L9/Sw1) – E270 (L11/Sw2)	4.7 (1.26)	6.16 (0.46)	1.46
S233 (L9/Sw1) – D265 (L11/Sw2)	5.95 (0.42)	3.55 (0.78)	−2.4
R234 (L9/Sw1) – G268 (L11/Sw2)	4.74 (1.03)	3.59 (1.13)	−1.15
R234 (L9/Sw1) – E270 (L11/Sw2)	3.31 (0.9)	5.05 (0.48)	1.74
R234 (L9/Sw1) – S275 (L11/Sw2)	5.35 (0.65)	8.44 (0.73)	3.09
H236 (β6) – N287 (α4)	6.84 (0.64)	5.12 (1.36)	−1.72
V256 (β7) – K371 (NL)	8.01 (2.08)	5.67 (1.97)	−2.34
T328 (β8) – P363 (NL)	5.55 (1.22)	4.3 (0.45)	−1.25

APPENDIX C

Additional Tables for Chapter IV

Table C.1: Average inter-residue distances from MD simulations of WT and apo states. All residue pairs listed, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference between ATP-bound wild type (WT) and nucleotide-free (apo) simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop.

Residue position	WT distance (Å)	apo distance (Å)	apo − WT (Å)
P4 (CS) – V331 (NL)	5.52 (2.53)	28.61 (5.73)	23.09
A5 (CS) – V331 (NL)	4.69 (2.47)	28.09 (4.81)	23.40
A5 (CS) – T330 (NL)	6.02 (3.36)	26.51 (3.43)	20.49
A5 (CS) – N295 (L13)	9.72 (3.19)	6.23 (3.58)	−3.49
E6 (CS) – V331 (NL)	4.80 (0.72)	25.76 (3.46)	20.96
E6 (CS) – N329 (NL)	4.60 (1.46)	21.85 (2.27)	17.25
E6 (CS) – K328 (NL)	4.31 (1.66)	19.00 (2.65)	14.69
E6 (CS) – G294 (L13)	9.09 (1.63)	4.98 (2.24)	−4.11
E6 (CS) – N295 (L13)	9.72 (1.80)	5.65 (2.36)	−4.07
C7 (CS) – V331 (NL)	4.01 (0.55)	24.83 (2.54)	20.82
C7 (CS) – N329 (NL)	2.89 (0.13)	20.74 (2.05)	17.85
C7 (CS) – K328 (NL)	3.45 (0.24)	17.57 (2.61)	14.12
C7 (CS) – I327 (NL)	4.52 (0.28)	15.24 (2.16)	10.72
C7 (CS) – G294 (L13)	7.68 (1.50)	4.30 (2.17)	−3.38
S8 (CS) – N329 (NL)	4.61 (0.48)	19.57 (3.71)	14.96
S8 (CS) – K328 (NL)	3.81 (0.49)	16.85 (4.78)	13.04

Residue position	WT distance (Å)	apo distance (Å)	apo − WT (Å)
S8 (CS) – I327 (NL)	3.50 (0.22)	14.14 (4.15)	10.64
S8 (CS) – G294 (L13)	9.74 (1.22)	5.07 (0.96)	−4.67
I9 (CS) – N329 (NL)	3.86 (0.46)	17.28 (2.90)	13.42
I9 (CS) – K328 (NL)	4.59 (0.42)	14.42 (4.51)	9.83
I9 (CS) – I327 (NL)	2.98 (0.17)	12.28 (3.42)	9.30
I9 (CS) – A270 (α4)	9.02 (0.67)	4.87 (0.85)	−4.15
I9 (CS) – S291 (α5)	4.62 (0.85)	8.54 (0.64)	3.92
I9 (CS) – K325 (α6)	3.88 (0.43)	7.25 (3.13)	3.37
I9 (CS) – L269 (α4)	6.44 (0.73)	4.07 (0.44)	−2.37
I9 (CS) – I266 (α4)	6.12 (0.82)	3.95 (0.41)	−2.17
V11 (β1) – K325 (α6)	3.75 (0.21)	7.35 (2.82)	3.60
R14 (β1) – P55 (β2d)	4.28 (1.30)	5.93 (0.47)	1.65
R16 (β1) – E22 (α0)	5.28 (2.58)	2.74 (0.08)	−2.54
R16 (β1) – S307 (L14)	6.39 (0.73)	4.51 (0.99)	−1.88
N19 (α0) – M198 (L9/S1)	3.91 (1.19)	9.93 (4.46)	6.02
A21 (α0) – N197 (L9/S1)	5.06 (2.25)	12.57 (4.74)	7.51
K34 (β2a) – N334 (NL)	35.54 (0.93)	5.71 (3.07)	−29.83
G44 (L2c) – V333 (NL)	40.35 (0.74)	6.61 (3.31)	−33.74
P46 (β2c) – S332 (NL)	31.10 (0.72)	5.24 (1.65)	−25.86
P46 (β2c) – V331 (NL)	26.91 (0.63)	5.17 (1.90)	−21.74
V48 (β2c) – I327 (NL)	11.33 (0.48)	5.29 (1.91)	−6.04
V48 (β2c) – T326 (NL)	8.65 (0.53)	4.44 (1.77)	−4.21
V48 (β2c) – K325 (α6)	3.20 (0.65)	6.45 (1.41)	3.25
F49 (L2d) – K325 (α6)	4.57 (0.51)	7.98 (1.83)	3.41
E76 (α1) – V333 (NL)	4.42 (0.92)	39.35 (2.24)	34.93
G77 (α1) – N334 (NL)	2.97 (0.21)	41.82 (2.98)	38.85
G77 (α1) – V333 (NL)	3.56 (0.47)	41.35 (1.76)	37.79
G77 (α1) – S332 (NL)	4.94 (0.57)	40.06 (2.35)	35.12
Y78 (β3) – V333 (NL)	3.63 (0.25)	37.53 (1.83)	33.90
Y78 (β3) – N334 (NL)	4.50 (0.47)	38.19 (2.73)	33.69
Y78 (β3) – S332 (NL)	3.62 (0.28)	36.20 (2.49)	32.58
Y78 (β3) – V331 (NL)	3.97 (0.40)	35.16 (2.30)	31.19
Y78 (β3) – N295 (L13)	5.67 (0.33)	3.75 (0.59)	−1.92
N79 (β3) – N334 (NL)	3.61 (1.21)	36.68 (2.63)	33.07

Residue position	WT distance (Å)	apo distance (Å)	apo − WT (Å)
N79 (β3) – S291 (α5)	7.35 (0.43)	3.76 (0.43)	−3.59
N79 (β3) – D290 (α5)	5.35 (0.62)	3.04 (0.43)	−2.31
G80 (β3) – S291 (α5)	5.56 (0.35)	3.56 (0.41)	−2.00
T81 (β3) – L292 (L13)	6.71 (0.38)	3.87 (0.35)	−2.84
T81 (β3) – C296 (β8)	3.75 (0.48)	5.73 (0.32)	1.98
Y85 (β3) – G235 (L11/S2)	7.13 (0.54)	4.08 (0.64)	−3.05
Y85 (β3) – L262 (α4)	4.25 (0.46)	7.05 (0.49)	2.80
Y85 (β3) – A234 (L11/S2)	6.14 (0.39)	3.74 (0.45)	−2.40
G86 (PL) – S236 (L11/S2)	3.61 (1.10)	6.31 (0.69)	2.70
Q87 (PL) – N310 (α6)	4.02 (1.47)	6.47 (0.53)	2.45
Q87 (PL) – K92 (PL)	3.54 (0.38)	5.23 (0.35)	1.69
T88 (PL) – E237 (L11/S2)	4.24 (0.76)	8.03 (1.07)	3.79
T88 (PL) – S236 (L11/S2)	3.83 (0.60)	7.53 (0.62)	3.70
T88 (PL) – G235 (L11/S2)	3.70 (0.26)	6.60 (0.67)	2.90
T88 (PL) – E313 (α6)	6.95 (1.41)	4.42 (0.96)	−2.53
S89 (PL) – T196 (L9/S1)	4.51 (0.96)	8.24 (2.07)	3.73
K92 (PL) – G235 (L11/S2)	3.67 (0.31)	5.71 (0.39)	2.04
T93 (PL) – S203 (L9/S1)	2.82 (0.13)	5.96 (0.96)	3.14
T93 (PL) – D232 (L11/S2)	2.72 (0.14)	5.35 (0.56)	2.63
T93 (PL) – R191 (L9/S1)	4.15 (0.41)	6.58 (0.77)	2.43
F117 (α2b) – Y121 (α2b)	3.16 (0.37)	5.65 (0.90)	2.49
D118 (α2b) – Y121 (α2b)	3.69 (0.49)	6.12 (0.43)	2.43
I120 (α2b) – M123 (α2b)	3.45 (0.43)	5.47 (0.33)	2.02
Y121 (α2b) – E128 (β4)	9.69 (0.97)	3.87 (1.31)	−5.82
Y121 (α2b) – L127 (β4)	8.42 (0.79)	4.40 (1.10)	−4.02
Y121 (α2b) – F129 (β4)	5.89 (0.94)	3.21 (0.29)	−2.68
Y121 (α2b) – D124 (α2b)	6.29 (0.53)	4.20 (0.76)	−2.09
M123 (α2b) – Q215 (β6)	5.18 (1.50)	8.54 (1.20)	3.36
K151 (L8a) – T170 (L8c)	5.93 (2.43)	3.61 (1.19)	−2.32
E158 (β5b) – R280 (L12)	6.64 (1.32)	3.44 (0.65)	−3.20
K188 (α3) – H201 (L9/S1)	4.17 (1.65)	6.38 (1.44)	2.21
H192 (L9/S1) – E251 (L11/S2)	6.49 (1.49)	3.94 (0.94)	−2.55
K214 (β6) – E336 (NL)	4.49 (1.68)	48.05 (2.94)	43.56
L224 (β7) – E336 (NL)	5.32 (1.59)	46.81 (3.07)	41.49

Residue position	WT distance (Å)	apo distance (Å)	apo − WT (Å)
L224 (β7) – N334 (NL)	4.40 (1.16)	45.88 (3.41)	41.48
S225 (β7) – E336 (NL)	3.24 (0.85)	45.90 (2.68)	42.66
S225 (β7) – N334 (NL)	2.98 (0.18)	44.78 (2.97)	41.80
G226 (β7) – N334 (NL)	4.06 (0.56)	42.61 (2.93)	38.55
K227 (β7) – N334 (NL)	5.01 (1.19)	39.08 (2.74)	34.07
K227 (β7) – D290 (α5)	7.30 (1.05)	2.91 (0.39)	−4.39
K227 (β7) – S291 (α5)	8.22 (0.55)	3.96 (0.98)	−4.26
Y229 (β7) – C296 (β8)	4.79 (0.88)	8.40 (0.49)	3.61
Y229 (β7) – L288 (α5)	6.05 (1.24)	3.66 (0.38)	−2.39
Y229 (β7) – D290 (α5)	6.28 (0.96)	4.04 (0.86)	−2.24
E237 (L11/S2) – T242 (L11/S2)	3.67 (1.11)	7.79 (2.77)	4.12
K238 (L11/S2) – E313 (α6)	4.02 (1.54)	6.22 (1.34)	2.20
K241 (L11/S2) – E313 (α6)	4.81 (2.33)	8.62 (2.62)	3.81
L259 (α4) – F320 (α6)	5.57 (0.53)	3.93 (0.40)	−1.64
G263 (α4) – R323 (α6)	5.41 (0.61)	3.47 (0.45)	−1.94
N264 (α4) – R323 (α6)	7.15 (0.67)	4.56 (0.95)	−2.59
I266 (α4) – I327 (NL)	3.78 (0.30)	13.78 (2.67)	10.00
I266 (α4) – F320 (α6)	3.74 (0.27)	5.39 (0.91)	1.65
L269 (α4) – I327 (NL)	3.73 (0.19)	14.18 (3.12)	10.45
L269 (α4) – T326 (NL)	5.11 (0.73)	11.35 (3.04)	6.24
L269 (α4) – K274 (L12)	4.55 (1.18)	7.57 (2.10)	3.02
L269 (α4) – G294 (L13)	6.85 (0.93)	4.24 (0.55)	−2.61
A270 (α4) – I327 (NL)	4.18 (0.47)	11.03 (2.65)	6.85
A270 (α4) – T326 (NL)	3.87 (0.65)	8.13 (2.60)	4.26
A270 (α4) – K325 (α6)	6.92 (0.61)	5.02 (2.14)	−1.90
K274 (L12) – V277 (L12)	4.00 (0.89)	5.88 (1.24)	1.88
Q289 (α5) – G294 (L13)	4.00 (0.46)	7.34 (0.62)	3.34
D290 (α5) – G294 (L13)	3.26 (0.21)	8.35 (0.55)	5.09
D290 (α5) – N295 (L13)	2.96 (0.18)	6.82 (0.76)	3.86
S291 (α5) – N329 (NL)	4.52 (1.61)	28.64 (2.74)	24.12
S291 (α5) – G294 (L13)	4.51 (0.48)	7.02 (0.55)	2.51
L292 (L13) – N329 (NL)	5.28 (1.36)	24.96 (2.82)	19.68
L292 (L13) – I327 (NL)	5.06 (0.77)	20.43 (2.93)	15.37
L292 (L13) – C296 (β8)	5.73 (0.40)	4.03 (0.62)	−1.70

Residue position	WT distance (Å)	apo distance (Å)	apo − WT (Å)
G293 (L13) – T330 (NL)	4.68 (1.17)	27.87 (3.34)	23.19
G293 (L13) – N329 (NL)	3.99 (1.54)	25.33 (3.26)	21.34
G293 (L13) – K328 (NL)	4.61 (0.95)	21.76 (5.16)	17.15
G293 (L13) – I327 (NL)	5.03 (0.95)	20.25 (3.50)	15.22
G293 (L13) – C296 (β8)	6.10 (0.40)	3.15 (0.41)	−2.95
G293 (L13) – T298 (β8)	6.13 (0.55)	3.39 (0.86)	−2.74
G294 (L13) – T330 (NL)	3.15 (0.37)	28.15 (3.47)	25.00
G294 (L13) – V331 (NL)	4.95 (1.15)	29.62 (1.77)	24.67
G294 (L13) – N329 (NL)	3.88 (1.34)	25.61 (3.43)	21.73
N295 (L13) – V333 (NL)	4.54 (1.00)	34.80 (1.64)	30.26
N295 (L13) – S332 (NL)	3.15 (0.50)	33.04 (2.48)	29.89
N295 (L13) – V331 (NL)	4.23 (1.38)	31.79 (1.73)	27.56
N295 (L13) – T330 (NL)	3.62 (0.44)	30.39 (3.25)	26.77
C296 (β8) – V331 (NL)	4.81 (1.74)	29.63 (1.66)	24.82
C296 (β8) – N329 (NL)	3.33 (0.83)	25.98 (2.97)	22.65
R297 (β8) – V331 (NL)	3.97 (0.39)	28.28 (2.24)	24.31
R297 (β8) – N329 (NL)	4.47 (0.64)	25.01 (2.38)	20.54
T298 (β8) – N329 (NL)	3.90 (0.89)	21.62 (2.70)	17.72
I300 (β8) – A324 (α6)	4.03 (0.40)	7.63 (2.61)	3.60
F320 (α6) – A324 (α6)	2.97 (0.14)	6.13 (1.73)	3.16
G321 (α6) – K325 (α6)	3.13 (0.20)	7.02 (1.68)	3.89
Q322 (α6) – K325 (α6)	3.39 (0.38)	5.37 (1.02)	1.98
R323 (α6) – I327 (NL)	4.15 (0.36)	9.64 (1.00)	5.49
R323 (α6) – T326 (NL)	3.03 (0.39)	6.24 (1.01)	3.21
A324 (α6) – I327 (NL)	3.31 (0.23)	6.66 (0.76)	3.35

Table C.2: Average inter-residue distances from MD simulations of WT and CNB+Latch mutant. All residue pairs listed, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference between wild type (WT) and mutant (CNB+Latch) simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop.

Residue position	WT dist. (Å)	CNB+Latch (Å)	CNB+Latch − WT (Å)
D3 (CS) – K328 (NL)	11.38 (2.88)	6.45 (4.02)	−4.93

Residue position	WT dist. (Å)	CNB+Latch (Å)	CNB+Latch − WT (Å)
D3 (CS) – C7 (CS)	8.35 (2.73)	5.43 (2.70)	−2.92
D3 (CS) – E6 (CS)	5.98 (1.79)	4.28 (1.74)	−1.70
P4 (CS) – V333 (NL)	10.00 (3.78)	4.58 (1.89)	−5.42
A5 (CS) – V333 (NL)	10.14 (3.21)	5.25 (1.19)	−4.89
A5 (CS) – N329 (NL)	5.69 (2.02)	3.83 (1.03)	−1.86
E6 (CS) – I9 (CS)	7.48 (0.74)	5.33 (2.14)	−2.15
C7 (CS) – L335 (NL)	16.03 (0.96)	7.77 (4.61)	−8.26
C7 (CS) – V333 (NL)	10.58 (0.94)	5.41 (2.07)	−5.17
C7 (CS) – N329 (NL)	2.89 (0.13)	4.77 (1.89)	1.88
C7 (CS) – K328 (NL)	3.45 (0.24)	4.99 (1.61)	1.54
S8 (CS) – V333 (NL)	14.19 (0.66)	7.27 (3.64)	−6.92
S8 (CS) – V331 (NL)	7.13 (0.67)	4.67 (1.10)	−2.46
S8 (CS) – I327 (NL)	3.50 (0.22)	5.26 (1.79)	1.76
S8 (CS) – K328 (NL)	3.81 (0.49)	5.29 (1.22)	1.48
R14 (β1) – ATP	3.69 (0.50)	5.56 (2.15)	1.87
R16 (β1) – E22 (α0)	5.28 (2.58)	2.72 (0.07)	−2.56
L18 (α0) – ATP	6.82 (0.60)	5.09 (1.27)	−1.73
N19 (α0) – ATP	5.90 (1.25)	3.80 (0.83)	−2.10
N19 (α0) – M198 (L9/S1)	3.91 (1.19)	5.29 (0.79)	1.38
A21 (α0) – N197 (L9/S1)	5.06 (2.25)	6.62 (1.08)	1.56
E22 (α0) – ATP	8.99 (1.95)	4.86 (0.42)	−4.13
E22 (α0) – S89 (PL)	7.49 (1.97)	4.40 (1.29)	−3.09
R25 (α0) – M198 (L9/S1)	7.38 (2.28)	3.92 (0.56)	−3.46
I41 (β2b) – K45 (β2c)	4.04 (0.98)	5.63 (0.67)	1.59
G42 (L2c) – K45 (β2c)	4.66 (1.40)	6.44 (0.68)	1.78
Q59 (α1) – R111 (α2b)	4.20 (0.83)	5.91 (0.42)	1.71
Q69 (α1) – E336 (NL)	18.86 (0.96)	8.62 (5.60)	−10.24
I70 (α1) – G80 (β3)	5.13 (1.25)	7.00 (0.52)	1.87
K72 (α1) – E336 (NL)	14.82 (0.99)	7.55 (5.01)	−7.27
D73 (α1) – S332 (NL)	8.03 (0.45)	5.48 (2.77)	−2.55
V74 (α1) – G80 (β3)	4.56 (0.62)	6.20 (0.54)	1.64
E76 (α1) – V333 (NL)	4.42 (0.92)	7.68 (2.08)	3.26
G77 (α1) – N334 (NL)	2.97 (0.21)	13.38 (1.03)	10.41
G77 (α1) – V333 (NL)	3.56 (0.47)	10.85 (0.76)	7.29

Residue position	WT dist. (Å)	CNB+Latch (Å)	CNB+Latch − WT (Å)
G77 (α1) – S332 (NL)	4.94 (0.57)	9.82 (2.29)	4.88
Y78 (β3) – N334 (NL)	4.50 (0.47)	8.64 (1.10)	4.14
Y78 (β3) – V331 (NL)	3.97 (0.40)	6.88 (0.81)	2.91
Y78 (β3) – V333 (NL)	3.63 (0.25)	6.21 (0.93)	2.58
N79 (β3) – N334 (NL)	3.61 (1.21)	13.19 (1.58)	9.58
N79 (β3) – S332 (NL)	5.29 (0.37)	8.61 (1.42)	3.32
N79 (β3) – G226 (β7)	3.63 (0.28)	5.34 (0.38)	1.71
N79 (β3) – Y229 (β7)	5.77 (0.73)	4.23 (0.85)	−1.54
Q87 (PL) – K238 (L11/S2)	7.18 (1.32)	4.84 (0.88)	−2.34
Q87 (PL) – M198 (L9/S1)	8.29 (1.51)	6.94 (3.24)	−1.35
Q104 (L5) – R111 (α2b)	4.10 (0.98)	6.26 (1.40)	2.16
K151 (L8a) – T170 (L8c)	5.93 (2.43)	3.46 (1.04)	−2.47
K151 (L8a) – G168 (L8c)	4.87 (2.00)	3.07 (0.46)	−1.80
A194 (L9/S1) – E251 (L11/S2)	6.43 (1.15)	4.53 (0.81)	−1.90
T196 (L9/S1) – E237 (L11/S2)	6.97 (1.38)	4.35 (0.64)	−2.62
T196 (L9/S1) – R204 (L9/S1)	5.99 (1.82)	4.03 (0.60)	−1.96
N197 (L9/S1) – F309 (L14)	9.36 (2.39)	4.53 (1.06)	−4.83
M198 (L9/S1) – F309 (L14)	8.53 (1.68)	5.13 (1.38)	−3.40
M198 (L9/S1) – N310 (α6)	7.64 (1.65)	6.53 (3.17)	−1.11
R204 (L9/S1) – N254 (L11/S2)	7.35 (0.67)	4.82 (1.88)	−2.53
R204 (L9/S1) – E237 (L11/S2)	5.34 (1.95)	3.02 (0.68)	−2.32
K214 (β6) – E336 (NL)	4.49 (1.68)	21.72 (3.38)	17.23
K223 (β7) – L335 (NL)	6.23 (2.81)	21.43 (3.36)	15.20
K223 (β7) – E336 (NL)	5.60 (1.94)	20.55 (3.09)	14.95
L224 (β7) – N334 (NL)	4.40 (1.16)	15.71 (1.48)	11.31
L224 (β7) – L335 (NL)	6.18 (2.26)	17.13 (3.20)	10.95
L224 (β7) – E336 (NL)	5.32 (1.59)	16.11 (2.94)	10.79
S225 (β7) – E336 (NL)	3.24 (0.85)	18.40 (3.01)	15.16
S225 (β7) – N334 (NL)	2.98 (0.18)	17.24 (0.99)	14.26
G226 (β7) – N334 (NL)	4.06 (0.56)	17.06 (0.96)	13.00
K227 (β7) – N334 (NL)	5.01 (1.19)	16.45 (1.36)	11.44
Y229 (β7) – C296 (β8)	4.79 (0.88)	6.44 (0.62)	1.65
E237 (L11/S2) – T242 (L11/S2)	3.67 (1.11)	6.35 (2.66)	2.68
E237 (L11/S2) – N256 (α4)	4.12 (1.18)	5.79 (0.85)	1.67

Residue position	WT dist. (Å)	CNB+Latch (Å)	CNB+Latch − WT (Å)
K238 (L11/S2) – N256 (α4)	5.21 (1.58)	8.03 (1.06)	2.82
G272 (L12) – H276 (L12)	7.05 (0.72)	4.90 (1.44)	−2.15
N295 (L13) – N334 (NL)	5.23 (1.11)	12.33 (0.96)	7.10
N295 (L13) – V333 (NL)	4.54 (1.00)	8.77 (1.28)	4.23
N295 (L13) – S332 (NL)	3.15 (0.50)	6.52 (0.66)	3.37
R297 (β8) – E336 (NL)	14.16 (1.35)	8.70 (5.04)	−5.46
R297 (β8) – V333 (NL)	7.56 (1.19)	4.84 (0.99)	−2.72

Table C.3: Average inter-residue distances from MD simulations of WT and Latch mutant. All residue pairs listed, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference between wild type (WT) and mutant (Latch, N334A) simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop.

Residue position	WT dist. (Å)	Latch (Å)	Latch − WT (Å)
D25 (L2a) – F309 (L14)	4.63 (1.32)	7.41 (1.32)	2.78
D25 (L2a) – S307 (L14)	4.12 (2.30)	6.34 (2.00)	2.22
G40 (L2c) – K45 (β2c)	4.66 (1.40)	6.52 (0.69)	1.86
D48 (L2d) – K325 (α6)	5.68 (1.03)	4.06 (1.55)	−1.62
E74 (α1) – V333 (NL)	4.42 (0.92)	10.38 (4.68)	5.96
G75 (α1) – N334 (NL)	2.97 (0.21)	8.81 (4.94)	5.84
G75 (α1) – V333 (NL)	3.56 (0.47)	9.15 (4.19)	5.59
G75 (α1) – S332 (NL)	4.94 (0.57)	9.16 (2.35)	4.22
Y76 (β3) – V331 (NL)	3.97 (0.40)	6.20 (1.87)	2.23
N77 (β3) – N334 (NL)	3.61 (1.21)	10.41 (3.70)	6.80
T86 (PL) – T242 (L11/S2)	8.03 (1.27)	5.57 (3.09)	−2.46
T86 (PL) – N199 (L9/S1)	6.27 (1.50)	4.12 (0.54)	−2.15
Y137 (β4) – E251 (L11/S2)	3.19 (0.97)	5.01 (1.39)	1.82
R170 (L8c) – E186 (α3)	3.59 (1.53)	5.33 (1.74)	1.74
T194 (L9/S1) – T242 (L11/S2)	9.41 (1.82)	4.92 (2.50)	−4.49
T194 (L9/S1) – E237 (L11/S2)	6.97 (1.38)	4.49 (0.83)	−2.48
N195 (L9/S1) – T242 (L11/S2)	9.47 (2.70)	3.70 (1.43)	−5.77
N195 (L9/S1) – G243 (L11/S2)	11.00 (3.86)	5.35 (2.57)	−5.65
M196 (L9/S1) – K241 (L11/S2)	11.86 (2.69)	6.03 (2.98)	−5.83

Residue position	WT dist. (Å)	Latch (Å)	Latch − WT (Å)
M196 (L9/S1) – T242 (L11/S2)	10.80 (1.79)	5.19 (2.49)	−5.61
M196 (L9/S1) – F309 (L14)	8.53 (1.68)	5.09 (1.45)	−3.44
R202 (L9/S1) – E251 (L11/S2)	3.42 (1.14)	6.06 (2.39)	2.64
R202 (L9/S1) – E237 (L11/S2)	5.34 (1.95)	3.16 (0.75)	−2.18
K212 (β6) – E336 (NL)	4.49 (1.68)	16.34 (6.66)	11.85
S223 (β7) – E336 (NL)	3.24 (0.85)	11.96 (6.28)	8.72
S223 (β7) – N334 (NL)	2.98 (0.18)	11.17 (4.91)	8.19
G224 (β7) – N334 (NL)	4.06 (0.56)	12.81 (4.77)	8.75
K225 (β7) – N334 (NL)	5.01 (1.19)	13.14 (3.74)	8.13
G244 (L11/S2) – D250 (L11/S2)	3.13 (0.31)	5.17 (2.11)	2.04

APPENDIX D

Additional Tables for Chapter V

Table D.1: Average inter-residue distances from MD simulations of WT and V8M mutant. All residue pairs listed, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference between wild type (WT) and mutant (V8M) simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop; S, switch.

Residue position	WT dist. (Å)	V8M(β1) (Å)	V8M − WT (Å)
A4 (CS) – R355 (NL)	5.84 (2.25)	3.49 (0.28)	−2.35
S5 (CS) – R324 (β8)	4.73 (1.46)	7.37 (1.07)	2.64
S5 (CS) – N357 (NL)	5.02 (1.64)	7.62 (0.87)	2.60
S5 (CS) – R355 (NL)	5.54 (1.26)	3.64 (0.55)	−1.90
V8 (β1) – A329 (β8)	5.72 (0.44)	3.98 (0.43)	−1.74
R13 (β1) – E19 (α0)	7.23 (1.32)	2.72 (0.07)	−4.51
R13 (β1) – M210 (L9/S1)	7.97 (1.49)	4.84 (1.70)	−3.13
P14 (β1) – E19 (α0)	6.94 (1.21)	4.50 (0.19)	−2.44
N16 (α0) – M210 (L9/S1)	5.91 (1.57)	4.02 (1.31)	−1.89
E19 (α0) – P333 (L14)	6.54 (0.89)	4.34 (0.39)	−2.20
Y89 (β3) – S242 (β7)	4.83 (1.13)	3.21 (0.47)	−1.62
F94 (β3) – N318 (α5)	7.19 (1.16)	3.98 (0.48)	−3.21
Y96 (β3) – S252 (L11/S2)	6.37 (1.57)	3.95 (0.70)	−2.42
G97 (PL) – S252 (L11/S2)	5.79 (1.08)	3.14 (0.38)	−2.65
Q98 (PL) – R254 (L11/S2)	4.00 (0.94)	7.54 (0.60)	3.54
Q98 (PL) – S332 (L14)	6.56 (0.66)	3.76 (1.02)	−2.80

Residue position	WT dist. (Å)	V8M(β1) (Å)	V8M − WT (Å)
Q98 (PL) – T341 (α6)	5.95 (0.74)	3.36 (0.50)	−2.59
Q98 (PL) – M210 (L9/S1)	5.11 (1.77)	7.51 (1.43)	2.40
Q98 (PL) – L331 (β8)	6.56 (0.28)	4.52 (1.22)	−2.04
T99 (PL) – E253 (L11/S2)	7.14 (0.87)	3.12 (0.58)	−4.02
T99 (PL) – R216 (L9/S1)	6.19 (1.54)	3.99 (0.91)	−2.20
T99 (PL) – S252 (L11/S2)	5.80 (0.49)	3.61 (0.53)	−2.19
T99 (PL) – N211 (L9/S1)	3.26 (0.43)	5.30 (1.11)	2.04
M108 (α2a) – Y189 (α3)	4.25 (0.89)	7.50 (1.25)	3.25
Q111 (L5) – D194 (α3)	4.07 (1.52)	7.31 (1.89)	3.24
P120 (α2b) – M196 (α3)	6.66 (0.78)	4.41 (0.97)	−2.25
C123 (α2b) – M196 (α3)	6.86 (0.83)	4.39 (0.96)	−2.47
N131 (α2b) – N135 (L6)	7.14 (1.24)	4.91 (1.95)	−2.23
V144 (β4) – L195 (α3)	7.21 (0.75)	4.57 (1.19)	−2.64
R155 (β5a) – N162 (L8a)	6.26 (0.38)	4.20 (1.18)	−2.06
N162 (L8a) – S182 (L8c)	3.80 (0.46)	6.31 (1.43)	2.51
N162 (L8a) – L181 (L8c)	4.06 (0.35)	6.15 (1.19)	2.09
N162 (L8a) – D180 (L8c)	2.95 (0.23)	5.02 (1.30)	2.07
K163 (L8a) – D180 (L8c)	4.28 (1.42)	7.58 (2.14)	3.30
K163 (L8a) – L166 (L8a)	3.92 (0.70)	6.02 (1.27)	2.10
L184 (L8c) – I192 (α3)	4.22 (0.56)	6.68 (1.24)	2.46
V186 (L8c) – L195 (α3)	7.49 (0.68)	4.78 (1.36)	−2.71
V186 (L8c) – D191 (α3)	6.03 (0.89)	3.85 (0.90)	−2.18
T187 (L8c) – D191 (α3)	5.95 (0.81)	3.47 (1.04)	−2.48
M210 (L9/S1) – R254 (L11/S2)	5.35 (2.01)	9.25 (2.00)	3.90
R216 (L9/S1) – N270 (L11/S2)	3.45 (0.89)	7.19 (0.89)	3.74
D231 (L10) – T234 (L10)	5.14 (1.38)	2.96 (0.36)	−2.18
A232 (L10) – N235 (L10)	3.79 (1.37)	6.53 (0.30)	2.74
A232 (L10) – I236 (L10)	4.03 (1.79)	6.71 (0.94)	2.68
G251 (L11/S2) – I271 (α4)	3.49 (0.26)	5.19 (1.30)	1.70
S252 (L11/S2) – T344 (α6)	8.74 (1.40)	4.87 (1.20)	−3.87
S252 (L11/S2) – N272 (α4)	3.61 (0.44)	5.70 (1.50)	2.09
E253 (L11/S2) – A269 (L11/S2)	4.60 (1.58)	7.81 (1.36)	3.21
E253 (L11/S2) – E340 (α6)	7.67 (1.18)	4.49 (1.21)	−3.18
E253 (L11/S2) – G268 (L11/S2)	4.13 (0.82)	6.75 (1.45)	2.62

Residue position	WT dist. (Å)	V8M(β1) (Å)	V8M − WT (Å)
E253 (L11/S2) – I271 (α4)	4.62 (1.08)	7.05 (1.33)	2.43
E253 (L11/S2) – L265 (L11/S2)	4.47 (0.52)	6.85 (1.45)	2.38
R254 (L11/S2) – D339 (α6)	6.29 (1.24)	3.78 (1.93)	−2.51
R254 (L11/S2) – S343 (α6)	6.31 (1.51)	4.20 (2.04)	−2.11
S257 (L11/S2) – K261 (L11/S2)	7.47 (0.72)	5.79 (2.95)	−1.68
G259 (L11/S2) – R264 (L11/S2)	7.55 (0.61)	4.88 (1.86)	−2.67
T263 (L11/S2) – E267 (L11/S2)	3.44 (0.85)	5.80 (2.40)	2.36
N270 (L11/S2) – S274 (α4)	4.57 (0.46)	2.99 (0.22)	−1.58
K273 (α4) – S309 (L12)	3.91 (0.51)	5.70 (0.55)	1.79
L285 (α4) – G321 (L13)	4.74 (1.16)	6.84 (0.42)	2.10
L285 (α4) – R355 (NL)	3.30 (0.58)	5.27 (0.40)	1.97
L285 (α4) – D289 (α4)	4.74 (1.03)	3.16 (0.35)	−1.58
A286 (α4) – R355 (NL)	4.52 (1.27)	7.09 (0.43)	2.57
A286 (α4) – R350 (α6)	5.98 (0.87)	4.17 (0.46)	−1.81
A286 (α4) – S290 (α4)	4.72 (1.49)	2.98 (0.37)	−1.74
M288 (α4) – G321 (L13)	5.85 (3.11)	9.05 (0.94)	3.20
M288 (α4) – N295 (L12)	7.08 (1.00)	4.79 (1.36)	−2.29
D289 (α4) – Q353 (NL)	6.69 (2.03)	3.37 (0.75)	−3.32
K299 (L12) – E317 (α5)	3.49 (1.21)	5.84 (2.59)	2.35
K300 (L12) – G320 (L13)	8.29 (1.17)	5.72 (2.97)	−2.57
R316 (α5) – G321 (L13)	6.31 (0.92)	4.10 (0.58)	−2.21
E317 (α5) – G321 (L13)	5.96 (1.69)	3.18 (0.22)	−2.78
L319 (L13) – M327 (β8)	7.68 (0.86)	5.00 (1.05)	−2.68
G320 (L13) – R355 (NL)	5.51 (1.70)	3.39 (0.39)	−2.12
G320 (L13) – N357 (NL)	5.38 (1.38)	3.40 (0.49)	−1.98
M327 (β8) – I354 (NL)	7.42 (0.82)	4.72 (0.98)	−2.70
M327 (β8) – Y347 (α6)	4.21 (0.47)	6.24 (1.01)	2.03
M327 (β8) – A348 (α6)	4.29 (0.73)	6.24 (0.67)	1.95

Table D.2: Average inter-residue distances from MD simulations of WT and Y89D mutant. All residue pairs listed, with their secondary structure location in parentheses, have a significant ($P < 10^{-5}$) difference between wild type (WT) and mutant (Y89D) simulations. Standard deviation values in parantheses. Abbreviations: CS, cover strand; NL, neck linker; L, loop; S, switch.

Residue position	WT dist. (Å)	Y89D($\alpha1/\beta3$) (Å)	Y89D − WT (Å)
G3 (CS) − R355 (NL)	6.89 (3.32)	3.84 (1.24)	−3.05
A4 (CS) − R355 (NL)	5.84 (2.25)	3.42 (0.21)	−2.42
A4 (CS) − R324 ($\beta8$)	5.02 (1.72)	6.98 (1.38)	1.96
S5 (CS) − A358 (NL)	6.34 (2.96)	9.62 (1.05)	3.28
S5 (CS) − K352 ($\alpha6$)	6.75 (1.61)	3.78 (0.72)	−2.97
S5 (CS) − R324 ($\beta8$)	4.73 (1.46)	7.29 (0.99)	2.56
S5 (CS) − S323 ($\beta8$)	5.90 (1.93)	8.41 (0.80)	2.51
S5 (CS) − N357 (NL)	5.02 (1.64)	7.50 (0.48)	2.48
S5 (CS) − R355 (NL)	5.54 (1.26)	3.84 (0.49)	−1.70
S5 (CS) − T325 ($\beta8$)	3.89 (0.95)	5.54 (0.61)	1.65
S5 (CS) − I354 (NL)	5.01 (1.16)	3.51 (0.34)	−1.50
R13 ($\beta1$) − M210 (L9/S1)	7.97 (1.49)	4.84 (1.16)	−3.13
R18 ($\alpha0$) − N209 (L9/S1)	5.60 (2.49)	3.58 (1.28)	−2.02
R22 ($\alpha0$) − N209 (L9/S1)	8.50 (2.63)	5.15 (1.94)	−3.35
R22 ($\alpha0$) − D335 (L14)	4.75 (1.84)	6.45 (0.76)	1.70
M30 ($\beta2a$) − N66 (L3)	3.89 (1.10)	8.80 (2.86)	4.91
S31 ($\beta2a$) − N66 (L3)	4.98 (1.13)	10.10 (2.95)	5.12
G32 (L2b) − N66 (L3)	4.94 (1.31)	10.07 (2.78)	5.13
D53 (L2d) − K352 ($\alpha6$)	6.04 (1.26)	4.38 (1.66)	−1.66
Y54 ($\beta2d$) − N66 (L3)	4.93 (1.22)	10.20 (3.24)	5.27
S55 ($\beta2d$) − N66 (L3)	5.09 (1.32)	9.16 (2.21)	4.07
W57 (L3) − D64 (L3)	3.90 (1.29)	7.26 (2.04)	3.36
W57 (L3) − N66 (L3)	3.39 (0.33)	4.98 (1.56)	1.59
W57 (L3) − A68 (L3)	3.07 (0.32)	4.49 (1.41)	1.42
P62 (L3) − Q115 (L5)	8.78 (1.47)	5.62 (2.11)	−3.16
H84 ($\alpha1$) − G88 ($\alpha1$)	3.28 (0.53)	4.88 (0.45)	1.60
E87 ($\alpha1$) − I360 (NL)	7.19 (5.06)	12.21 (3.23)	5.02
E87 ($\alpha1$) − R324 ($\beta8$)	5.86 (2.34)	4.37 (2.64)	−1.49
G88 ($\alpha1$) − I360 (NL)	6.62 (5.14)	12.95 (2.14)	6.33

Residue position	WT dist. (Å)	Y89D($\alpha1/\beta3$) (Å)	Y89D $-$ WT (Å)
G88 ($\alpha1$) $-$ V241 ($\beta7$)	4.92 (1.05)	3.26 (0.31)	-1.66
Y89 ($\beta3$) $-$ I360 (NL)	5.57 (3.49)	10.04 (1.64)	4.47
Y89 ($\beta3$) $-$ V359 (NL)	5.22 (2.30)	7.75 (1.43)	2.53
Y89 ($\beta3$) $-$ V241 ($\beta7$)	6.29 (1.04)	3.77 (0.37)	-2.52
Y89 ($\beta3$) $-$ A358 (NL)	4.66 (1.92)	6.73 (1.13)	2.07
Y89 ($\beta3$) $-$ K243 ($\beta7$)	4.41 (0.49)	3.05 (0.20)	-1.36
Y89 ($\beta3$) $-$ S242 ($\beta7$)	4.83 (1.13)	3.48 (0.25)	-1.35
Q98 (PL) $-$ R254 (L11/S2)	4.00 (0.94)	6.91 (0.95)	2.91
Q98 (PL) $-$ M210 (L9/S1)	5.11 (1.77)	6.59 (0.54)	1.48
G100 (PL) $-$ M210 (L9/S1)	6.17 (1.27)	4.50 (1.18)	-1.67
S104 (PL) $-$ S215 (L9/S1)	5.49 (0.35)	3.71 (1.17)	-1.78
N135 (L6) $-$ H230 ($\beta6$)	7.25 (1.57)	5.04 (1.63)	-2.21
S139 ($\beta4$) $-$ H230 ($\beta6$)	6.52 (1.09)	4.39 (1.47)	-2.13
Y150 (L7) $-$ E267 (L11/S2)	4.67 (1.94)	6.25 (1.29)	1.58
M210 (L9/S1) $-$ R254 (L11/S2)	5.35 (2.01)	10.25 (2.65)	4.90
S215 (L9/S1) $-$ MG2	4.19 (0.31)	2.85 (0.84)	-1.34
R216 (L9/S1) $-$ G268 (L11/S2)	4.04 (1.03)	6.16 (1.66)	2.12
R216 (L9/S1) $-$ N270 (L11/S2)	3.45 (0.89)	5.37 (1.22)	1.92
K228 ($\beta6$) $-$ N235 (L10)	8.85 (1.44)	6.14 (2.99)	-2.71
R229 ($\beta6$) $-$ T234 (L10)	7.94 (0.94)	5.08 (1.35)	-2.86
R229 ($\beta6$) $-$ T238 ($\beta7$)	2.91 (0.16)	4.25 (1.20)	1.34
H230 ($\beta6$) $-$ T237 ($\beta7$)	3.91 (0.53)	5.82 (1.94)	1.91
H230 ($\beta6$) $-$ I236 (L10)	3.77 (0.80)	5.44 (1.55)	1.67
D231 (L10) $-$ T238 ($\beta7$)	4.27 (1.13)	8.11 (2.16)	3.84
D231 (L10) $-$ T237 ($\beta7$)	4.48 (0.81)	6.79 (1.87)	2.31
D231 (L10) $-$ T234 (L10)	5.14 (1.38)	2.94 (0.35)	-2.20
D231 (L10) $-$ I236 (L10)	3.29 (0.92)	5.48 (1.75)	2.19
A232 (L10) $-$ I236 (L10)	4.03 (1.79)	7.28 (1.19)	3.25
A232 (L10) $-$ N235 (L10)	3.79 (1.37)	6.18 (0.97)	2.39
S252 (L11/S2) $-$ N272 ($\alpha4$)	3.61 (0.44)	5.59 (1.47)	1.98
E253 (L11/S2) $-$ T258 (L11/S2)	9.04 (1.99)	3.98 (1.23)	-5.06
E253 (L11/S2) $-$ S257 (L11/S2)	9.90 (1.75)	5.21 (1.70)	-4.69
E253 (L11/S2) $-$ L265 (L11/S2)	4.47 (0.52)	6.56 (0.92)	2.09
E253 (L11/S2) $-$ A269 (L11/S2)	4.60 (1.58)	6.38 (0.94)	1.78

Residue position	WT dist. (Å)	Y89D($\alpha1/\beta3$) (Å)	Y89D $-$ WT (Å)
R254 (L11/S2) – T258 (L11/S2)	7.06 (2.26)	4.33 (1.58)	-2.73
R254 (L11/S2) – S257 (L11/S2)	6.68 (1.93)	4.04 (1.24)	-2.64
R254 (L11/S2) – E340 ($\alpha6$)	3.31 (1.25)	5.12 (2.00)	1.81
R254 (L11/S2) – D339 ($\alpha6$)	6.29 (1.24)	4.66 (1.78)	-1.63
A255 (L11/S2) – T258 (L11/S2)	5.73 (1.35)	3.83 (0.94)	-1.90
G268 (L11/S2) – I271 ($\alpha4$)	3.47 (0.39)	5.32 (1.75)	1.85
L285 ($\alpha4$) – C356 (NL)	5.16 (1.19)	7.38 (1.02)	2.22
L285 ($\alpha4$) – R355 (NL)	3.30 (0.58)	4.71 (0.98)	1.41
A286 ($\alpha4$) – R355 (NL)	4.52 (1.27)	6.10 (0.81)	1.58
K299 (L12) – E317 ($\alpha5$)	3.49 (1.21)	5.86 (3.18)	2.37
T301 (L12) – G321 (L13)	3.89 (0.55)	6.05 (1.77)	2.16
D302 (L12) – R316 ($\alpha5$)	4.57 (1.21)	6.33 (1.37)	1.76
N322 (L13) – V359 (NL)	3.53 (0.94)	5.56 (2.43)	2.03
S323 ($\beta8$) – A358 (NL)	4.40 (1.06)	5.98 (1.10)	1.58
R324 ($\beta8$) – A358 (NL)	4.80 (1.79)	6.56 (1.34)	1.76

APPENDIX E

Additional Tables for Chapter VII

Table E.1: Experimental structures selected for PCA. PPI: protein-protein interactions present in a given structure.

PDB ID	Chains	Nucleotide	Ligands (α)	Ligand (β)	PPI	Reference
5W3J	A, B	GDP	GTP,MG	GDP,TA1		Howes et al. J. Cell Biol. (2017)
4U3J	A, B	GTP	GTP,MG	GTP,MG	TOG	Ayaz et al. Elife (2014)
4FFB	A, B	GTP	GTP,MG	GTP,MG	TOG	Ayaz et al. Science (2012)
5W3F	A, B	GDP	GTP,MG	GDP		Howes et al. J. Cell Biol. (2017)
5W3H	A, B	GDP	GTP,MG	EP,GDP		Howes et al. J. Cell Biol. (2017)
5MJS	H, A	GDP	GTP	GDP	EB1/3	von Loeffelholz et al. Nat Commun (2017)
5UCY	A, B	GDP	GTP,MG	GDP		Ichikawa et al. Nat Commun (2017)
5JCO	B, K	GMPCPP	GTP,MG	G2P,MG		Vemu et al. J.Biol.Chem. (2016)
5IJ9	A, B	GDP	GTP,MG	GDP		Ti et al. Dev.Cell (2016)
5IJ0	A, B	GDP	GTP,MG	GDP		Ti et al. Dev.Cell (2016)
5N5N	K, B	GMPCPP	GTP,MG	G2P,MG		Vemu et al. Mol. Biol. Cell (2017)
3E22	C, D	GDP	GTP,MG	GDP,LOC	Stathmin	Cormier et al. Embo Rep. (2008)
3DU7	C, D	GDP	MG	CN2,GDP,GTP,HOS	Stathmin	Cormier et al. Embo Rep. (2008)
5SYG	A, B	GDP	GTP,MG	GDP,ZPN		Kellogg et al. J. Mol. Biol. (2017)
5SYF	A, B	GDP	GTP,MG	GDP,TA1		Kellogg et al. J. Mol. Biol. (2017)
5SYE	A, B	GDP	GTP,MG	GDP,POU,TA1		Kellogg et al. J. Mol. Biol. (2017)
5SYC	A, B	GDP	GTP,MG	GDP,POU		Kellogg et al. J. Mol. Biol. (2017)
5FNV	C, D	GTP	CA,GTP,MG,X3H	GTP,MG	Stathmin, TTL	Yang et al. Nat.Commun. (2016)
5JQG	C, D	GTP	CA,GTP,MG	GTP,MG	Stathmin, TTL	Yang et al. Nat Commun (2016)
4ZOL	C, D	GDP	CA,GTP,MG	55Q,GDP,MG	Stathmin, TTL	Wang et al. Mol.Pharmacol. (2016)
4ZHQ	C, D	GDP	CA,GTP,MG	GDP,MG	Stathmin, TTL	Wang et al. Mol.Pharmacol. (2016)
4ZI7	C, D	GDP	CA,GTP,MG	GDP,MG	Stathmin, TTL	Wang et al. Mol.Pharmacol. (2016)

PDB ID	Chains	Nucleotide	Ligands (α)	Ligand (β)	PPI	Reference
5BMV	C, D	GDP	CA,GTP,MG,VLB	GDP,MG	Stathmin, TTL	Wang et al. Mol.Pharmacol. (2016)
5KMG	A, B	GDP	GTP,MG	GDP,POU	PRC1	Kellogg et al. Proc.Natl.Acad.Sci.USA (2016)
3JAK	K, B	GDP-Pi	GTP,MG	GSP	EB1/3	Zhang et al. Cell (2015)
3JAL	K, B	GMPCPP	GTP,MG	G2P	EB1/3	Zhang et al. Cell (2015)
3JAR	K, B	GDP	GTP,MG	GDP	EB1/3	Zhang et al. Cell (2015)
3JAW	A, B	GDP-Pi	GTP,MG	GSP		Zhang et al. Cell (2015)
6BJC	K, B	GMPCPP	GTP,MG	G2P,MG	TPX2	Zhang et al. Elife (2017)
3JAS	K, B	GDP	GTP,MG	GDP		Zhang et al. Cell (2015)
3JAT	K, B	GMPCPP	GTP,MG	G2P,MG		Zhang et al. Cell (2015)
1SA1	C, D	GDP	GTP,MG	GDP,POD	Stathmin	Ravelli et al. Nature (2004)
1SA0	C, D	GDP	GTP,MG	CN2,GDP	Stathmin	Ravelli et al. Nature (2004)
1JFF	A, B	GDP	GTP,MG,ZN	GDP,TA1		Lowe et al. J.Mol.Biol. (2001)
1TVK	A, B	GDP	GTP	EP,GDP		Nettles et al. Science (2004)
5ND4	A, B	GDP	GTP,MG	GDP,TA1	Kinesin	Atherton et al. Elife (2017)
5OGC	A, B	GDP	GTP,MG,ZN	GDP,TA1	Kinesin	Locke et al. Proc. Natl. Acad. Sci. U.S.A. (2017)
3J6E	A, B	GMPCPP	GTP,MG	G2P,MG		Alushin et al. Cell (2014)
3J6F	A, B	GDP	GTP,MG	GDP		Alushin et al. Cell (2014)
5M5C	A, B	GDP	GTP,MG	GDP,TA1	CAMSAP	Atherton et al. Nat. Struct. Mol. Biol. (2017)
4LNU	A, B	GDP	GTP,MG,SO4	GDP,MES,SO4	DARPin, Kinesin	Cao et al. Nat Commun (2014)
4EB6	C, D	GDP	GTP,MG,VLB	GDP,SO4	Stathmin	Ranaivoson et al. Acta Crystallogr.,Sect.D (2012)
3UT5	C, D	GDP	GTP,MG	GDP,LOC,SO4	Stathmin, Vinca	Ranaivoson et al. Acta Crystallogr.,Sect.D (2012)
3RYI	C, D	GDP	GTP,MG,SO4	GDP,SO4	Stathmin	Nawrotek et al. J.Mol.Biol. (2011)
4F6R	A, B	GDP	GTP,MG,SO4	GDP,MES	DARPin, Stathmin	Mignot et al. J.Biol.Chem. (2012)

PDB ID	Chains	Nucleotide	Ligands (α)	Ligand (β)	PPI	Reference
3RYH	C, D	GMPCPP	GTP,MG,SO4	G2P,MG,SO4	Stathmin	Nawrotek et al. J.Mol.Biol. (2011)
3RYC	C, D	GTP	GTP,MG,SO4	GTP,MG,SO4	Stathmin	Nawrotek et al. J.Mol.Biol. (2011)
3RYF	C, D	GTP	GTP,MG,SO4	GTP,MG,SO4	Stathmin	Nawrotek et al. J.Mol.Biol. (2011)
1Z2B	C, D	GDP	GTP,MG,VLB	CN2,GDP	Stathmin	Gigant et al. Nature (2005)
5KX5	C, D	GDP	CA,GTP,MG	6YK,GDP,MG	Stathmin, TTL	Leverett et al. ACS Med Chem Lett (2016)
3HKD	C, D	GDP	GTP,MG	GDP,N16	Stathmin	Dorleans et al. Proc.Natl.Acad.Sci.USA (2009)
3N2G	C, D	GDP	GTP,MG	G2N,GDP	Stathmin	Barbier et al. J.Biol.Chem. (2010)
3N2K	C, D	GDP	GTP,MG	GDP,K2N	Stathmin	Barbier et al. J.Biol.Chem. (2010)
3HKE	C, D	GDP	GTP,MG	GDP,T13	Stathmin	Dorleans et al. Proc.Natl.Acad.Sci.USA (2009)
3HKC	C, D	GDP	GTP,MG	E70,GDP	Stathmin	Dorleans et al. Proc.Natl.Acad.Sci.USA (2009)
3HKB	C, D	GDP	GTP,MG	GDP	Stathmin	Dorleans et al. Proc.Natl.Acad.Sci.USA (2009)
4X1I	C, D	GDP	GTP,MG	3WD,GDP,LOC	Stathmin	Maderna et al. J.Med.Chem. (2014)
4X1K	C, D	GDP	GTP,MG	3WZ,GDP,LOC	Stathmin	Maderna et al. J.Med.Chem. (2014)
4X1Y	C, D	GDP	GTP,MG	3WV,GDP,LOC	Stathmin	Maderna et al. J.Med.Chem. (2014)
4X20	C, D	GDP	GTP,MG	3WY,GDP,LOC	Stathmin	Maderna et al. J.Med.Chem. (2014)
5EIB	C, D	GTP	GTP,MG	GTP,MG	CENPJ, DARPin	Zheng et al. Nat Commun (2016)
4DRX	A, B	GTP	GTP,MG	GTP,MG	DARPin	Pecqueur et al. Proc.Natl.Acad.Sci.USA (2012)
3J8X	A, B	GDP	GTP	GDP	Kinesin	Shang et al. Elife (2014)
3J8Y	A, B	GDP	GTP	GDP	Kinesin	Shang et al. Elife (2014)
4HNA	A, B	GDP	GTP,MG	ALF,GDP,MG	DARPin, Kinesin	Gigant et al. Nat.Struct.Mol.Biol. (2013)
4F61	C, D	GDP	GTP,MG	GDP	Stathmin	Mignot et al. J.Biol.Chem. (2012)
5ITZ	A, B	GDP	GTP,MG	GDP,LOC	CENPJ, DARPin	Sharma et al. Dev.Cell (2016)
5JCB	C, D	GDP	CA,GTP,IMD,MG,NA	GDP,MG,NV4	Stathmin, TTL	Zhao et al. ACS Chem. Biol. (2017)

PDB ID	Chains	Nucleotide	Ligands (α)	Ligand (β)	PPI	Reference
5M7G	C, D	GDP	CA,GTP,MG	FB7,GDP,MG	Stathmin, TTL	Bohnacker et al. Nat Commun (2017)
5NQT	A, B	GDP	GTP,MG	GDP	DARPin	Weinert et al. Nat Commun (2017)
5EYP	A, B	GDP	GTP,MG,SO4	GDP,LOC,SO4	DARPin	Ahmad et al. Sci Rep (2016)
5MIO	A, B	GDP	GTP,MG	GDP,LOC	DARPin, Kinesin	Wang et al. Nat Commun (2017)
4I50	C, D	GDP	CA,GTP,MG	EP,GDP,MES,MG	Stathmin, TTL	Prota et al. Science (2013)
5NG1	C, D	GDP	CA,GTP,MG	8WB,GDP,MG,ZPN	Stathmin, TTL	Field et al. J. Nat. Prod. (2017)
5LXS	C, D	GDP	CA,GTP,MG	7AO,GDP,MG	Stathmin, TTL	Prota et al. Chembiochem (2017)
5EZY	C, D	GTP	CA,GTP,MG	GTP,MG,TAJ	Stathmin, TTL	Wang et al. Nat Commun (2017)
4O4L	C, D	GDP	GTP,MG	EP,GDP,MG,POU	Stathmin, TTL	Prota et al. Angew.Chem.Int.Ed.Engl. (2014)
4O4J	C, D	GDP	GTP,MG	GDP,MG,POU	Stathmin, TTL	Prota et al. Angew.Chem.Int.Ed.Engl. (2014)
4O4I	C, D	GDP	CA,GTP,MG	EP,GDP,MG	Stathmin, TTL	Prota et al. Angew.Chem.Int.Ed.Engl. (2014)
4I4T	C, D	GDP	CA,GTP,MG	GDP,MG,ZPN	Stathmin, TTL	Prota et al. Science (2013)
5OV7	C, D	GDP	GTP,MG	6FS,GDP,MG	Stathmin, TTL	Jost et al. Mol. Cell (2017)
5NJH	C, D	GDP	GTP,MG	GDP,MG	Stathmin, TTL	Saez-Calvo et al. Cell Chem Biol (2017)
5JH7	C, D	GDP	CA,GTP,MG	6K9,GDP	Stathmin, TTL	Doodhi et al. Curr.Biol. (2016)
5J2T	C, D	GDP	GTP,MG,VLB	GDP	Stathmin, TTL	Waight et al. Plos One (2016)
5JVD	C, D	GDP	CA,GTP,MG	6NL,GDP,MG	Stathmin	Canela et al. Oncotarget (2017)
4IIJ	C, D	GDP	CA,GTP,MG	GDP,MG	Stathmin, TTL	Prota et al. J.Cell Biol. (2013)
4I55	C, D	GDP	CA,GTP,MG	GDP,MG	Stathmin, TTL	Prota et al. Science (2013)
5NM5	A, B	GDP	GTP,MG	GDP,LOC	DARPin	Weinert et al. Nat Commun (2017)
5O7A	C, D	GDP	GTP	9N5,GDP,MG	Stathmin, TTL	Sharma et al. Int J Mol Sci (2017)
4O2A	C, D	GDP	CA,GTP	2RR,GDP,MG	Stathmin, TTL	Prota et al. J.Mol.Biol. (2014)
4O2B	C, D	GDP	CA,GTP,IMD,MG	GDP,LOC,MG	Stathmin, TTL	Prota et al. J.Mol.Biol. (2014)

PDB ID	Chains	Nucleotide	Ligands (α)	Ligand (β)	PPI	Reference
4TV9	C, D	GDP	CA,GTP,MG	3H4,GDP,MG	Stathmin, TTL	Prota et al. Proc.Natl.Acad.Sci.USA (2014)
4YJ2	C, D	GDP	CA,GTP,IMD,MG	4ED,GDP,MG	Stathmin, TTL	McNamara et al. Protein Sci. (2015)
4YJ3	C, D	GDP	CA,GTP,IMD,MG	4EE,GDP,MG	Stathmin, TTL	McNamara et al. Protein Sci. (2015)
5C8Y	C, D	GDP	CA,GTP,MG	GDP,PN6	Stathmin, TTL	Wang et al. Febs J. (2016)
5CA0	C, D	GDP	CA,GTP,MG	GDP,LXL	Stathmin, TTL	Wang et al. Febs J. (2016)
5CA1	C, D	GDP	CA,GTP,MG	GDP,MES,NZO	Stathmin, TTL	Wang et al. Febs J. (2016)
5CB4	C, D	GDP	CA,GTP,MG	GDP,TIV	Stathmin, TTL	Wang et al. Febs J. (2016)
5GON	C, D	GDP	CA,GTP,IMD,MG	GDP,MG	Stathmin, TTL	Zhou et al. J. Med. Chem. (2016)
5H7O	C, D	GTP	CA,GTP,MG	7Q7,GTP,MG	Stathmin, TTL	Arnst et al. Cancer Res. (2018)
5XAF	C, D	GDP	CA,GTP,IMD,MG	84F,GDP,MG	Stathmin, TTL	Zhou et al. Eur J Med Chem (2017)
5XHC	C, D	GTP	CA,GTP,MG	GTP,MG	Stathmin, TTL	Chu et al. To Be Published (null)
5XI5	C, D	GTP	CA,GTP,MG	GTP,MG,PN6	Stathmin, TTL	Chu et al. To Be Published (null)
5XI7	C, D	GTP	CA,GTP,MG	GTP,MG,PO7	Stathmin, TTL	Chu et al. To Be Published (null)
5XLT	C, D	GDP	CA,GTP,MG	89O,GDP,MG	Stathmin, TTL	Niu et al. Biochem. Biophys. Res. Commun. (2017)
5XLZ	C, D	GTP	CA,GTP,MG	89U,GTP,MG	Stathmin, TTL	Cheng et al. Biochem. Biophys. Res. Commun. (2018)
5XP3	C, D	GTP	CA,GTP,MG	GTP,MG	Stathmin, TTL	Wei et al. To Be Published (null)
5YL4	C, D	GTP	CA,GTP,MG	GTP,MG	Stathmin, TTL	Fu et al. To Be Published (null)
5OSK	C, D	GDP	CA,GTP,IMD,MG	GDP,MG	Stathmin, TTL	Dohle et al. J. Med. Chem. (2018)
5LOV	C, D	GDP	71E,GTP,MG	GDP	Stathmin, TTL	Wieczorek et al. Sci Transl Med (2016)
5H74	C, D	GTP	CA,GTP,MES,MG	7LG,GTP,MG	Stathmin, TTL	Wang et al. To Be Published (null)
5IYZ	C, D	GDP	CA,GTP,MG	4Q5,GDP,MG	Stathmin, TTL	Waight et al. Plos One (2016)
5NFZ	C, D	GDP	CA,GTP,MG	8WB,GDP,MG	Stathmin, TTL	Field et al. J. Nat. Prod. (2017)
5LYJ	C, D	GDP	CA,GTP,MG	7BA,GDP,MG	Stathmin, TTL	Gaspari et al. Chem (2017)

PDB ID	Chains	Nucleotide	Ligands (α)	Ligand (β)	PPI	Reference
5M7E	C, D	GDP	CA,GTP,MG	GDP,MG	Stathmin, TTL	Bohnacker et al. Nat Commun (2017)
5M8D	C, D	GDP	CA,GTP,MG	GDP,MG	Stathmin, TTL	Bohnacker et al. Nat Commun (2017)
5M8G	C, D	GDP	CA,GTP,MG	918,GDP,MES,MG	Stathmin, TTL	Bohnacker et al. Nat Commun (2017)
4WBN	C, D	GDP	CA,CL,GTP,MG	CA,GDP	Stathmin, TTL	Weinert et al. Nat.Methods (2015)
5LP6	C, D	GDP	CA,GTP,MG	GDP	Stathmin, TTL	Marangon et al. Biochem. Biophys. Res. Commun. (2016)
5MF4	C, D	GDP	CA,GTP,MG	7LZ,GDP,MES,MG	Stathmin, TTL	Trigili et al. Acs Omega (2017)
4IHJ	C, D	GDP	CA,GTP,MG	GDP,MG	Stathmin, TTL	Prota et al. J.Cell Biol. (2013)
5J2U	C, D	GDP	GTP,MG	6DO,GDP	Stathmin	Waight et al. Plos One (2016)
4O4H	C, D	GDP	GTP,MG	GDP,MG	Stathmin, TTL	Prota et al. Angew.Chem.Int.Ed.Engl. (2014)
5LXT	C, D	GDP	CA,GTP,MG	7AK,GDP,MG	Stathmin, TTL	Prota et al. Chembiochem (2017)
4TUY	C, D	GDP	CA,GTP,MG	36L,GDP,MG	Stathmin, TTL	Prota et al. Proc.Natl.Acad.Sci.USA (2014)
4TV8	C, D	GDP	CA,GTP,MG	3GT,GDP,MG	Stathmin, TTL	Prota et al. Proc.Natl.Acad.Sci.USA (2014)
5NQU	A, B	GDP	GTP,MG	GDP	DARPin	Weinert et al. Nat Commun (2017)

BIBLIOGRAPHY

BIBLIOGRAPHY

Abraham, M. J., T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindah (2015), Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX*, doi: 10.1016/j.softx.2015.06.001.

Alushin, G. M., G. C. Lander, E. H. Kellogg, R. Zhang, D. Baker, and E. Nogales (2014), High-Resolution microtubule structures reveal the structural transitions in $\alpha\beta$-tubulin upon GTP hydrolysis, *Cell*, *157*(5), 1117–1129, doi: 10.1016/j.cell.2014.03.053.

Andreu, J. M., J. Garcia de Ancos, D. Starling, J. L. Hodgkinson, and J. Bordas (1989), A synchrotron x-ray scattering characterization of purified tubulin and of its expansion induced by mild detergent binding, *Biochemistry*, *28*(9), 4036–4040, doi: 10.1021/bi00435a060.

Arfken, G. (1985), Determinants, Matrices, And Group Theory, in *Mathematical Methods for Physicists*, pp. 168–276, Elsevier, doi: 10.1016/b978-0-12-059820-5.50012-4.

Arrhenius, S. (1889), Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren, *Zeitschrift für Physikalische Chemie*, *4*(1), 226–248, doi: 10.1515/zpch-1889-0116.

Asenjo, A. B., Y. Weinberg, and H. Sosa (2006), Nucleotide binding and hydrolysis induces a disorder-order transition in the kinesin neck-linker region, *Nature Structural & Molecular Biology*, *13*(7), 648–654, doi: 10.1038/nsmb1109.

Atherton, J., I. Farabella, I.-M. Yu, S. S. Rosenfeld, A. Houdusse, M. Topf, and C. A. Moores (2014), Conserved mechanisms of microtubule-stimulated ADP release, ATP binding, and force generation in transport kinesins, *eLife*, *3*, doi: 10.7554/eLife.03680.

Atherton, J., et al. (2017), The divergent mitotic kinesin MKLP2 exhibits atypical structure and mechanochemistry, *eLife*, *6*, doi: 10.7554/eLife.27793.

Bachmann, A., and A. Straube (2015), Kinesins in cell migration, *Biochemical Society Transactions*, doi: 10.1042/bst20140280.

Bahi-Buisson, N., et al. (2014), The wide spectrum of tubulinopathies: what are the key features for the diagnosis?, *Brain*, *137*(6), 1676–1700, doi: 10.1093/brain/awu082.

Bakan, A., L. M. Meireles, and I. Bahar (2011), ProDy: Protein dynamics inferred from theory and experiments, *Bioinformatics*, doi: 10.1093/bioinformatics/btr168.

Barber, K. W., and J. Rinehart (2018), The ABCs of PTMs, *Nature Chemical Biology*, doi: 10.1038/nchembio.2572.

Barducci, A., G. Bussi, and M. Parrinello (2008), Well-tempered metadynamics: A smoothly converging and tunable free-energy method, *Physical Review Letters*, doi: 10.1103/PhysRevLett.100.020603.

Behnke-Parks, W. M., J. Vendome, B. Honig, Z. Maliga, C. Moores, and S. S. Rosenfeld (2011), Loop L5 acts as a conformational latch in the mitotic kinesin Eg5, *Journal of Biological Chemistry*, *286*(7), 5242–5253, doi: 10.1074/jbc.M110.192930.

Benjamini, Y., and Y. Hochberg (1995), Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300, doi: 10.1111/j.2517-6161.1995.tb02031.x.

Berman, H. M., et al. (2002), The protein data bank, *Acta Crystallographica Section D: Biological Crystallography*, doi: 10.1107/S0907444902003451.

Bickel, K. G., B. J. Mann, J. S. Waitzman, T. A. Poor, S. E. Rice, and P. Wadsworth (2017), Src family kinase phosphorylation of the motor domain of the human kinesin-5, Eg5, *Cytoskeleton*, doi: 10.1002/cm.21380.

Bieling, P., L. Laan, H. Schek, E. L. Munteanu, L. Sandblad, M. Dogterom, D. Brunner, and T. Surrey (2007), Reconstitution of a microtubule plus-end tracking system in vitro, *Nature*, *450*(7172), 1100–1105, doi: 10.1038/nature06386.

Blee, T. K. P., N. K. Gray, and M. Brook (2015), Modulation of the cytoplasmic functions of mammalian post-transcriptional regulatory proteins by methylation and acetylation: a key layer of regulation waiting to be uncovered?, *Biochemical Society Transactions*, *43*(6), 1285–1295, doi: 10.1042/BST20150172.

Boehr, D. D., R. Nussinov, and P. E. Wright (2009), The role of dynamic conformational ensembles in biomolecular recognition, doi: 10.1038/nchembio.232.

Bonomi, M., et al. (2009), PLUMED: A portable plugin for free-energy calculations with molecular dynamics, *Computer Physics Communications*, doi: 10.1016/j.cpc.2009.05.011.

Braun, A., K. Dang, F. Buslig, M. A. Baird, M. W. Davidson, C. M. Waterman, and K. A. Myers (2014), Rac1 and Aurora A regulate MCAK to polarize microtubule growth in migrating endothelial cells, *Journal of Cell Biology*, doi: 10.1083/jcb.201401063.

Brouhard, G. J., and L. M. Rice (2018), Microtubule dynamics: an interplay of biochemistry and mechanics, *Nature Reviews Molecular Cell Biology*, *19*(7), 451–463, doi: 10.1038/s41580-018-0009-y.

Brownhill, K., L. Wood, and V. Allan (2009), Molecular motors and the Golgi complex: Staying put and moving through, *Seminars in Cell & Developmental Biology*, *20*(7), 784–792, doi: 10.1016/j.semcdb.2009.03.019.

Budaitis, B. G., S. Jariwala, D. N. Reinemann, K. I. Schimert, G. Scarabelli, B. J. Grant, D. Sept, M. J. Lang, and K. J. Verhey (2019), Neck Linker Docking is Critical for Kinesin-1 Force Generation in Cells but at a Cost to Motor Speed and Processivity, *eLife*, doi: 10.7554/eLife.

Buss, F., and J. Kendrick-Jones (2008), How are the cellular functions of myosin VI regulated within the cell?, *Biochemical and Biophysical Research Communications*, doi: 10.1016/j.bbrc.2007.11.150.

Cao, L., W. Wang, Q. Jiang, C. Wang, M. Knossow, and B. Gigant (2014), The structure of apo-kinesin bound to tubulin links the nucleotide cycle to movement, *Nature Communications*, doi: 10.1038/ncomms6364.

Case, D., et al. (2014), *AMBER 14*, University of California, San Francisco.

Case, D. A., et al. (2012), *AMBER 12*, University of California, San Francisco.

Case, D. e. a. (2018), *AMBER 18*, University of California, San Francisco.

Case, R. B., S. Rice, C. L. Hart, B. Ly, and R. D. Vale (2000), Role of the kinesin neck linker and catalytic core in microtubule-based motility, *Current Biology*, *10*(3), 157–160, doi: 10.1016/S0960-9822(00)00316-X.

Castle, B. T., and D. J. Odde (2013), Brownian Dynamics of Subunit Addition-Loss Kinetics and Thermodynamics in Linear Polymer Self-Assembly, *Biophysical Journal*, *105*(11), 2528–2540, doi: 10.1016/j.bpj.2013.10.009.

Chaaban, S., S. Jariwala, C.-T. Hsu, S. Redemann, J. M. Kollman, T. Müller-Reichert, D. Sept, K. H. Bui, and G. J. Brouhard (2018), The Structure and Dynamics of C. elegans Tubulin Reveals the Mechanistic Basis of Microtubule Growth, *Developmental Cell*, pp. 1–14, doi: 10.1016/j.devcel.2018.08.023.

Chiba, K., C. Min, S. Arai, K. Hashimoto, R. J. McKenney, and S. Niwa (2019), Disease-associated mutations hyperactivate kif1a motility and anterograde axonal transport of synaptic vesicle precursors, *bioRxiv*, doi: 10.1101/597906.

Choudhary, C., C. Kumar, F. Gnad, M. L. Nielsen, M. Rehman, T. C. Walther, J. V. Olsen, and M. Mann (2009), Lysine acetylation targets protein complexes and co-regulates major cellular functions, *Science*, doi: 10.1126/science.1175371.

Clancy, B. E., W. M. Behnke-Parks, J. O. L. Andreasson, S. S. Rosenfeld, and S. M. Block (2011a), A universal pathway for kinesin stepping, *Nature Structural and Molecular Biology*, doi: 10.1038/nsmb.2104.

Clancy, B. E., W. M. Behnke-Parks, J. O. L. Andreasson, S. S. Rosenfeld, and S. M. Block (2011b), A universal pathway for kinesin stepping, *Nature Structural and Molecular Biology*, doi: 10.1038/nsmb.2104.

Cohen, T. J., J. L. Guo, D. E. Hurtado, L. K. Kwong, I. P. Mills, J. Q. Trojanowski, and V. M. Y. Lee (2011), The acetylation of tau inhibits its function and promotes pathological tau aggregation, *Nature Communications*, *2*(1), 252, doi: 10.1038/ncomms1255.

Davtyan, A., N. P. Schafer, W. Zheng, C. Clementi, P. G. Wolynes, and G. A. Papoian (2012), AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing, *The Journal of Physical Chemistry B*, *116*(29), 8494–8503, doi: 10.1021/jp212541y.

de Boor, S., et al. (2015), Small GTP-binding protein Ran is regulated by posttranslational lysine acetylation, *Proceedings of the National Academy of Sciences*, *112*(28), E3679–E3688, doi: 10.1073/pnas.1505995112.

De Vivo, M., M. Masetti, G. Bottegoni, and A. Cavalli (2016), Role of Molecular Dynamics and Related Methods in Drug Discovery, *Journal of Medicinal Chemistry*, *59*(9), 4035–4061, doi: 10.1021/acs.jmedchem.5b01684.

DeBerg, H. A., et al. (2013), Motor Domain Phosphorylation Modulates Kinesin-1 Transport, *Journal of Biological Chemistry*, *288*(45), 32,612–32,621, doi: 10.1074/jbc.M113.515510.

Debiec, K. T., A. M. Gronenborn, and L. T. Chong (2014), Evaluating the Strength of Salt Bridges: A Comparison of Current Biomolecular Force Fields, *The Journal of Physical Chemistry B*, *118*(24), 6561–6569, doi: 10.1021/jp500958r.

Dogan, M. Y., S. Can, F. B. Cleary, V. Purde, and A. Yildiz (2015), Kinesin's Front Head Is Gated by the Backward Orientation of Its Neck Linker, *Cell Reports*, *10*(12), 1967–1973, doi: 10.1016/j.celrep.2015.02.061.

Doshi, U., M. J. Holliday, E. Z. Eisenmesser, and D. Hamelberg (2016), Dynamical network of residue–residue contacts reveals coupled allosteric effects in recognition, catalysis, and mutation, *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1523573113.

Eddy, S. R. (2011), Accelerated profile HMM searches, *PLoS Computational Biology*, doi: 10.1371/journal.pcbi.1002195.

Edgar, R. C. (2004), MUSCLE: Multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research*, doi: 10.1093/nar/gkh340.

Elber, R., and M. Karplus (1987), Multiple conformational states of proteins: A molecular dynamics analysis of myoglobin, *Science*, doi: 10.1126/science.3798113.

Esmaeeli Nieh, S., et al. (2015), De novo mutations in KIF1A cause progressive encephalopathy and brain atrophy, *Annals of Clinical and Translational Neurology*, *2*(6), 623–635, doi: 10.1002/acn3.198.

Eyal, E., G. Lum, and I. Bahar (2015), The anisotropic network model web server at 2015 (ANM 2.0), *Bioinformatics*, doi: 10.1093/bioinformatics/btu847.

Falnikar, A., S. Tole, and P. W. Baas (2011), Kinesin-5, a mitotic microtubule-associated motor protein, modulates neuronal migration, *Molecular Biology of the Cell*, *22*(9), 1561–1574, doi: 10.1091/mbc.e10-11-0905.

Finn, R. D., et al. (2014), Pfam: The protein families database, doi: 10.1093/nar/gkt1223.

Fletcher, D. A., and R. D. Mullins (2010), Cell mechanics and the cytoskeleton, *Nature*, *463*(7280), 485–492, doi: 10.1038/nature08908.

Forsberg, E. C., and E. H. Bresnick (2001), Histone acetylation beyond promoters: long-range acetylation patterns in the chromatin world, *BioEssays*, *23*(9), 820–830, doi: 10.1002/bies.1117.

Freedman, H., T. Luchko, R. F. Luduena, and J. A. Tuszynski (2011), Molecular dynamics modeling of tubulin C-terminal tail interactions with the microtubule surface, *Proteins: Structure, Function and Bioinformatics*, *79*(10), 2968–2982, doi: 10.1002/prot.23155.

Fygenson, D. K., E. Braun, and A. Libchaber (1994), Phase diagram of microtubules, *Physical Review E*, *50*(2), 1579–1588, doi: 10.1103/PhysRevE.50.1579.

Garcia-Saez, I., D. Blot, R. Kahn, and F. Kozielski (2004), Crystallization and preliminary crystallographic analysis of the motor domain of human kinetochore-associated protein CENP-E using an automated crystallization procedure, *Acta Crystallographica Section D Biological Crystallography*, *60*(6), 1158–1160, doi: 10.1107/S0907444904009564.

Gardner, M. K., B. D. Charlebois, I. M. Jánosi, J. Howard, A. J. Hunt, and D. J. Odde (2011), Rapid Microtubule Self-Assembly Kinetics, *Cell*, *146*(4), 582–592, doi: 10.1016/j.cell.2011.06.053.

Gebremichael, Y., J.-W. Chu, and G. A. Voth (2008), Intrinsic Bending and Structural Rearrangement of Tubulin Dimer: Molecular Dynamics Simulations and Coarse-Grained Analysis, *Biophysical Journal*, *95*(5), 2487–2499, doi: 10.1529/biophysj.108.129072.

Geyer, E. A., A. Burns, B. A. Lalonde, X. Ye, F.-A. Piedra, T. C. Huffaker, and L. M. Rice (2015), A mutation uncouples the tubulin conformational and GTPase cycles, revealing allosteric control of microtubule dynamics, *eLife*, *4*, doi: 10.7554/eLife.10113.

Gigant, B., W. Wang, B. Dreier, Q. Jiang, L. Pecqueur, A. Plückthun, C. Wang, and M. Knossow (2013), Structure of a kinesin-tubulin complex and implications for kinesin motility, *Nature Structural and Molecular Biology*, doi: 10.1038/nsmb.2624.

Goldman, R. D., M. M. Cleland, S. P. Murthy, S. Mahammad, and E. R. Kuczmarski (2012), Inroads into the structure and function of intermediate filament networks, *Journal of Structural Biology*, *177*(1), 14–23, doi: 10.1016/j.jsb.2011.11.017.

Gorfe, A. A., B. J. Grant, and J. A. McCammon (2008), Mapping the Nucleotide and Isoform-Dependent Structural and Dynamical Features of Ras Proteins, *Structure*, *16*(6), 885–896, doi: 10.1016/j.str.2008.03.009.

Gorsky, M. K., S. Burnouf, J. Dols, E. Mandelkow, and L. Partridge (2016), Acetylation mimic of lysine 280 exacerbates human Tau neurotoxicity in vivo, *Scientific Reports*, *6*(1), 22,685, doi: 10.1038/srep22685.

Goulet, A., J. Major, Y. Jun, S. P. Gross, S. S. Rosenfeld, and C. A. Moores (2014), Comprehensive structural model of the mechanochemical cycle of a mitotic motor highlights molecular adaptations in the kinesin family, *Proceedings of the National Academy of Sciences*, *111*(5), 1837–1842, doi: 10.1073/pnas.1319848111.

Grafmüller, A., and G. A. Voth (2011), Intrinsic bending of microtubule protofilaments, *Structure*, *19*(3), 409–417, doi: 10.1016/j.str.2010.12.020.

Grafmüller, A., E. G. Noya, and G. A. Voth (2013), Nucleotide-dependent lateral and longitudinal interactions in microtubules, *Journal of Molecular Biology*, *425*(12), 2232–2246, doi: 10.1016/j.jmb.2013.03.029.

Grant, B. J., A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon, and L. S. D. Caves (2006), Bio3d: An R package for the comparative analysis of protein structures, *Bioinformatics*, doi: 10.1093/bioinformatics/btl461.

Grant, B. J., J. A. McCammon, and A. A. Gorfe (2010), Conformational Selection in G-Proteins: Lessons from Ras and Rho, *Biophysical Journal*, *99*(11), L87–L89, doi: 10.1016/j.bpj.2010.10.020.

Hahlen, K., B. Ebbing, J. Reinders, J. Mergler, A. Sickmann, and G. Woehlke (2006), Feedback of the Kinesin-1 Neck-linker Position on the Catalytic Site, *Journal of Biological Chemistry*, *281*(27), 18,868–18,877, doi: 10.1074/jbc.M508019200.

Hanson, J. A., J.-W. Chu, K. Duderstadt, J. Brokaw, H. Yang, S. Bhattacharyya, and L. P. Watkins (2007), Illuminating the mechanistic roles of enzyme conformational dynamics, *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.0708600104.

Hariharan, V., and W. O. Hancock (2009), Insights into the Mechanical Properties of the Kinesin Neck Linker Domain from Sequence Analysis and Molecular Dynamics Simulations, *Cellular and Molecular Bioengineering*, *2*(2), 177–189, doi: 10.1007/s12195-009-0059-5.

Hartl, F. U., A. Bracher, and M. Hayer-Hartl (2011), Molecular chaperones in protein folding and proteostasis, doi: 10.1038/nature10317.

Henzler-Wildman, K., and D. Kern (2007), Dynamic personalities of proteins, *Nature*, *450*(7172), 964–972, doi: 10.1038/nature06522.

Henzler-Wildman, K. A., et al. (2007), Intrinsic motions along an enzymatic reaction trajectory, *Nature*, doi: 10.1038/nature06410.

Hesse, W. R., M. Steiner, M. L. Wohlever, R. D. Kamm, W. Hwang, and M. J. Lang (2013), Modular Aspects of Kinesin Force Generation Machinery, *Biophysical Journal*, *104*(9), 1969–1978, doi: 10.1016/j.bpj.2013.03.051.

Hirokawa, N., and R. Takemura (2004), Kinesin superfamily proteins and their various functions and dynamics, doi: 10.1016/j.yexcr.2004.08.010.

Hirokawa, N., Y. Noda, Y. Tanaka, and S. Niwa (2009), Kinesin superfamily motor proteins and intracellular transport, *Nature Reviews Molecular Cell Biology*, *10*(10), 682–696, doi: 10.1038/nrm2774.

Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling (2006), Comparison of multiple amber force fields and development of improved protein backbone parameters, doi: 10.1002/prot.21123.

Hornbeck, P. V., B. Zhang, B. Murray, J. M. Kornhauser, V. Latham, and E. Skrzypek (2015), PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations, *Nucleic Acids Research*, doi: 10.1093/nar/gku1267.

Hotchkiss, L., et al. (2016), Novel De Novo Mutations in KIF1A as a Cause of Hereditary Spastic Paraplegia With Progressive Central Nervous System Involvement, *Journal of Child Neurology*, *31*(9), 1114–1119, doi: 10.1177/0883073816639718.

Howes, S. C., E. A. Geyer, B. LaFrance, R. Zhang, E. H. Kellogg, S. Westermann, L. M. Rice, and E. Nogales (2017), Structural differences between yeast and mammalian microtubules revealed by cryo-EM, *The Journal of Cell Biology*, p. jcb.201612195, doi: 10.1083/jcb.201612195.

Humphrey, W., A. Dalke, and K. Schulten (1996), VMD: Visual molecular dynamics, *Journal of Molecular Graphics*, *14*(1), 33–38, doi: 10.1016/0263-7855(96)00018-5.

Hwang, W., M. J. Lang, and M. Karplus (2008), Force Generation in Kinesin Hinges on Cover-Neck Bundle Formation, *Structure*, *16*(1), 62–71, doi: 10.1016/j.str.2007.11.008.

Hwang, W., M. J. Lang, and M. Karplus (2017), Kinesin motility is driven by sub-domain dynamics, *eLife*, *6*, 1–25, doi: 10.7554/eLife.28948.

Igaev, M., and H. Grubmüller (2018), Microtubule assembly governed by tubulin allosteric gain in flexibility and lattice induced fit, *eLife*, *7*, 1–21, doi: 10.7554/eLife.34353.

Iqbal, Z., S. L. Rydning, I. M. Wedding, J. Koht, L. Pihlstrøm, A. H. Rengmark, S. P. Henriksen, C. M. E. Tallaksen, and M. Toft (2017), Targeted high through-put sequencing in hereditary ataxia and spastic paraplegia, *PLOS ONE*, *12*(3), e0174,667, doi: 10.1371/journal.pone.0174667.

Isojima, H., R. Iino, Y. Niitani, H. Noji, and M. Tomishige (2016), Direct observation of intermediate states during the stepping motion of kinesin-1, *Nature Chemical Biology*, *12*(4), 290–297, doi: 10.1038/nchembio.2028.

Janke, C. (2014), The tubulin code: Molecular components, readout mechanisms, functions, *Journal of Cell Biology*, doi: 10.1083/jcb.201406055.

Jariwala, S., L. Skjærven, X.-Q. Yao, and B. J. Grant (2017), Investigating Protein Sequence-structure-dynamics Relationships with Bio3D-web, *Journal of Visualized Experiments*, doi: 10.3791/55640.

Jenuwein, T., and C. D. Allis (2001), Translating the histone code, doi: 10.1126/science.1063127.

Jolly, A. L., H. Kim, D. Srinivasan, M. Lakonishok, A. G. Larson, and V. I. Gelfand (2010), Kinesin-1 heavy chain mediates microtubule sliding to drive changes in cell shape, *Proceedings of the National Academy of Sciences*, *107*(27), 12,151–12,156, doi: 10.1073/pnas.1004736107.

Jon Kull, F., E. P. Sablin, R. Lau, R. J. Fletterick, and R. D. Vale (1996), Crystal structure of the kinesin motor domain reveals a structural similarity to myosin, *Nature*, *380*(6574), 550–555, doi: 10.1038/380550a0.

Kabsch, W. (1978), A discussion of the solution for the best rotation to relate two sets of vectors, *Acta Crystallographica Section A*, *34*(5), 827–828, doi: 10.1107/S0567739478001680.

Kabsch, W., and C. Sander (1983), Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers*, *22*(12), 2577–2637, doi: 10.1002/bip.360221211.

Kapitein, L. C., M. A. Schlager, W. A. van der Zwan, P. S. Wulf, N. Keijzer, and C. C. Hoogenraad (2010), Probing Intracellular Motor Protein Activity Using an Inducible Cargo Trafficking Assay, *Biophysical Journal*, *99*(7), 2143–2152, doi: 10.1016/j.bpj.2010.07.055.

Kerns, S. J., et al. (2015), The energy landscape of adenylate kinase during catalysis, *Nature Structural and Molecular Biology*, doi: 10.1038/nsmb.2941.

Khalil, A. S., D. C. Appleyard, A. K. Labno, A. Georges, M. Karplus, A. M. Belcher, W. Hwang, and M. J. Lang (2008), Kinesin's cover-neck bundle folds forward to generate force., *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19,247–19,252, doi: 10.1073/pnas.0805147105.

Klejnot, M., A. Falnikar, V. Ulaganathan, R. A. Cross, P. W. Baas, and F. Kozielski (2014), The crystal structure and biochemical characterization of Kif15: a bifunctional molecular motor involved in bipolar spindle formation and neuronal development, *Acta Crystallographica Section D Biological Crystallography*, *70*(1), 123–133, doi: 10.1107/S1399004713028721.

Kononova, O., Y. Kholodov, K. E. Theisen, K. A. Marx, R. I. Dima, F. I. Ataullakhanov, E. L. Grishchuk, and V. Barsegov (2014), Tubulin Bond Energies and Microtubule Biomechanics Determined from Nanoindentation in Silico, *Journal of the American Chemical Society*, *136*(49), 17,036–17,045, doi: 10.1021/ja506385p.

Kouzarides, T. (2000), Acetylation: a regulatory modification to rival phosphorylation?, *The EMBO Journal*, *19*(6), 1176–1179, doi: 10.1093/emboj/19.6.1176.

Kozielski, F., S. Sack, A. Marx, M. Thormählen, E. Schönbrunn, V. Biou, A. Thompson, E.-M. Mandelkow, and E. Mandelkow (1997), The Crystal Structure of Dimeric Kinesin and Implications for Microtubule-Dependent Motility, *Cell*, *91*(7), 985–994, doi: 10.1016/S0092-8674(00)80489-4.

Kruskal, W. H. (1952), A Nonparametric test for the Several Sample Problem, *The Annals of Mathematical Statistics*, *23*(4), 525–540, doi: 10.1214/aoms/1177729332.

Kruskal, W. H., and W. A. Wallis (1952), Use of Ranks in One-Criterion Variance Analysis, *Journal of the American Statistical Association*, doi: 10.1080/01621459.1952.10483441.

Kueh, H. Y., and T. J. Mitchison (2009), Structural Plasticity in Actin and Tubulin Polymer Dynamics, *Science*, *325*(5943), 960–963, doi: 10.1126/science.1168823.

Kull, F. J., and S. A. Endow (2002), Kinesin: switch I & II and the motor mechanism., *Journal of cell science*, *115*(Pt 1), 15–23.

Kunwar, A., M. Vershinin, J. Xu, and S. P. Gross (2008), Stepping, Strain Gating, and an Unexpected Force-Velocity Curve for Multiple-Motor-Based Transport, *Current Biology*, doi: 10.1016/j.cub.2008.07.027.

Kutner, M. H., C. J. Nachtsheim, J. Neter, W. Li, et al. (2005), *Applied Linear Statistical Models*, vol. 5, 744–745 pp., McGraw-Hill Irwin Boston.

Lang, M. J., and W. Hwang (2010), Motor Proteins: Kinesin Drives with an Underhead Cam, *Current Biology*, *20*(9), R408–R410, doi: 10.1016/j.cub.2010.03.015.

Lee, J.-R., et al. (2015), De Novo Mutations in the Motor Domain of KIF1A Cause Cognitive Impairment, Spastic Paraparesis, Axonal Neuropathy, and Cerebellar Atrophy, *Human Mutation*, *36*(1), 69–78, doi: 10.1002/humu.22709.

Li, H., A. D. Robertson, and J. H. Jensen (2005), Very fast empirical prediction and rationalization of protein pK a values, *Proteins: Structure, Function and Genetics*, doi: 10.1002/prot.20660.

Little, M., and T. Seehaus (1988), Comparative analysis of tubulin sequences, *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, *90*(4), 655–670, doi: 10.1016/0305-0491(88)90320-3.

Liu, D., X. Liu, Z. Shang, and C. V. Sindelar (2017), Structural basis of cooperativity in kinesin revealed by 3D reconstruction of a two-head-bound state on microtubules, *eLife*, *6*, doi: 10.7554/eLife.24490.

Liu, Z., et al. (2014), CPLM: A database of protein lysine modifications, *Nucleic Acids Research*, doi: 10.1093/nar/gkt1093.

Löwe, J., H. Li, K. H. Downing, and E. Nogales (2001), Refined structure of $\alpha\beta$-tubulin at 3.5 Å resolution, *Journal of Molecular Biology*, doi: 10.1006/jmbi.2001.5077.

Luchko, T., J. Torin Huzil, M. Stepanova, and J. Tuszynski (2008), Conformational Analysis of the Carboxy-Terminal Tails of Human β-Tubulin Isotypes, *Biophysical Journal*, *94*(6), 1971–1982, doi: 10.1529/biophysj.107.115113.

Ludueña, R. F., and A. Banerjee (2008), *The Isotypes of Tubulin*, pp. 123–175, Humana Press, Totowa, NJ, doi: 10.1007/978-1-59745-336-3_6.

Lue, R., A. Viel, and J. Liebler (), The inner life series.

Mallik, R., D. Petrov, S. A. Lex, S. J. King, and S. P. Gross (2005), Building complexity: An in vitro study of cytoplasmic dynein with in vivo implications, *Current Biology*, doi: 10.1016/j.cub.2005.10.039.

Marmorstein, R., and M. M. Zhou (2014), Writers and readers of histone acetylation: Structure, mechanism, and inhibition, *Cold Spring Harbor Perspectives in Biology*, doi: 10.1101/cshperspect.a018762.

McKenney, R. J., M. Vershinin, A. Kunwar, R. B. Vallee, and S. P. Gross (2010), LIS1 and NudE Induce a Persistent Dynein Force-Producing State, *Cell*, *141*(2), 304–314, doi: 10.1016/j.cell.2010.02.035.

Meagher, K. L., L. T. Redman, and H. A. Carlson (2003), Development of polyphosphate parameters for use with the AMBER force field, *Journal of Computational Chemistry*, *24*(9), 1016–1025, doi: 10.1002/jcc.10262.

Menzies, K. J., H. Zhang, E. Katsyuba, and J. Auwerx (2016), Protein acetylation in metabolism — metabolites and cofactors, *Nature Reviews Endocrinology*, *12*(1), 43–60, doi: 10.1038/nrendo.2015.181.

Minor, D. L., and P. S. Kim (1994), Measurement of the β-sheet-forming propensities of amino acids, *Nature*, *367*(6464), 660–663, doi: 10.1038/367660a0.

Mitchison, T., and M. Kirschner (1984), Dynamic instability of microtubule growth, *Nature*, *312*(5991), 237–242, doi: 10.1038/312237a0.

Monticelli, L., S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink (2008), The MARTINI coarse grained force field: extension to proteins, *J. Chem. Theory Comput.*, *4*(5), 819–834, doi: 10.1021/ct700324x.

Muretta, J. M., Y. Jun, S. P. Gross, J. Major, D. D. Thomas, and S. S. Rosenfeld (2015), The structural kinetics of switch-1 and the neck linker explain the functions of kinesin-1 and Eg5, *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1512305112.

Muretta, J. M., et al. (2018), A posttranslational modification of the mitotic kinesin Eg5 that enhances its mechanochemical coupling and alters its mitotic function, *Proceedings of the National Academy of Sciences*, *115*(8), E1779–E1788, doi: 10.1073/pnas.1718290115.

Muthukrishnan, G., Y. Zhang, S. Shastry, and W. O. Hancock (2009), The Processivity of Kinesin-2 Motors Suggests Diminished Front-Head Gating, *Current Biology*, *19*(5), 442–447, doi: 10.1016/j.cub.2009.01.058.

Nalawansha, D. A., I. D. Gomes, M. K. Wambua, and M. K. H. Pflum (2017), HDAC Inhibitor-Induced Mitotic Arrest Is Mediated by Eg5/KIF11 Acetylation, *Cell Chemical Biology*, *24*(4), 481–492.e5, doi: 10.1016/j.chembiol.2017.03.008.

Nick Pace, C., and J. Martin Scholtz (1998), A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins, *Biophysical Journal*, *75*(1), 422–427, doi: 10.1016/S0006-3495(98)77529-0.

Nitta, R., Y. Okada, and N. Hirokawa (2008), Structural model for strain-dependent microtubule activation of Mg-ADP release from kinesin, *Nature Structural & Molecular Biology*, *15*(10), 1067–1075, doi: 10.1038/nsmb.1487.

Nogales, E. (2000), Structural Insights into Microtubule Function, *Annual Review of Biochemistry*, *69*(1), 277–302, doi: 10.1146/annurev.biochem.69.1.277.

Nogales, E., S. G. Wolf, and K. H. Downing (1998), Structure of the $\alpha\beta$ tubulin dimer by electron crystallography, *Nature*, *391*(6663), 199–203, doi: 10.1038/34465.

Ohba, C., et al. (2015), De novo KIF1A mutations cause intellectual deficit, cerebellar atrophy, lower limb spasticity and visual disturbance, *Journal of Human Genetics*, *60*(12), 739–742, doi: 10.1038/jhg.2015.108.

Okamoto, N., F. Miya, T. Tsunoda, K. Yanagihara, M. Kato, S. Saitoh, M. Yamasaki, Y. Kanemura, and K. Kosaki (2014), KIF1A mutation in a patient with progressive neurodegeneration, *Journal of Human Genetics*, *59*(11), 639–641, doi: 10.1038/jhg. 2014.80.

Okazaki, K.-i., and S. Takada (2011), Structural Comparison of F1-ATPase: Interplay among Enzyme Structures, Catalysis, and Rotations, *Structure*, *19*(4), 588–598, doi: 10.1016/j.str.2011.01.013.

Olsen, J. V., et al. (2010), Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis, *Science Signaling*, *3*(104), ra3– ra3, doi: 10.1126/scisignal.2000475.

Onufriev, A., D. Bashford, and D. A. Case (2004), Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model, *Proteins: Structure, Function and Genetics*, doi: 10.1002/prot.20033.

Oosawa, F., and S. Asakura (1975), *Thermodynamics of the Polymerization of Protein*, Academic Press.

Parke, C. L., E. J. Wojcik, S. Kim, and D. K. Worthylake (2010a), ATP hydrolysis in Eg5 kinesin involves a catalytic two-water mechanism, *Journal of Biological Chemistry*, doi: 10.1074/jbc.M109.071233.

Parke, C. L., E. J. Wojcik, S. Kim, and D. K. Worthylake (2010b), ATP hydrolysis in Eg5 kinesin involves a catalytic two-water mechanism, *Journal of Biological Chemistry*, doi: 10.1074/jbc.M109.071233.

Peng, L. X., M. T. Hsu, M. Bonomi, D. A. Agard, and M. P. Jacobson (2014), The Free Energy Profile of Tubulin Straight-Bent Conformational Changes, with Implications for Microtubule Assembly and Drug Discovery, *PLoS Computational Biology*, *10*(2), e1003,464, doi: 10.1371/journal.pcbi.1003464.

Perez-Melero, C. (2014), KSP Inhibitors as Antimitotic Agents, *Current Topics in Medicinal Chemistry*, *14*(20), 2286–2311, doi: 10.2174/ 1568026614666141130095532.

Pollard, T. D., W. C. Earnshaw, J. Lippincott-Schwartz, and G. T. Johnson (2017), Cytoskeleton and cellular motility, in *Cell Biology (Third Edition)*, third edition ed., pp. 571 – 573, Elsevier, doi: https://doi.org/10.1016/B978-0-323-34126-4.00061-X.

Prota, A. E., K. Bargsten, D. Zurwerra, J. J. Field, J. F. Díaz, K.-H. Altmann, and M. O. Steinmetz (2013), Molecular Mechanism of Action of Microtubule-Stabilizing Anticancer Agents, *Science*, *339*(6119), 587–590, doi: 10.1126/science.1230582.

Raman, S. (2018), Systems Approaches to Understanding and Designing Allosteric Proteins, *Biochemistry*, *57*(4), 376–382, doi: 10.1021/acs.biochem.7b01094.

Reddy, B. J. N., M. Mattson, C. L. Wynne, O. Vadpey, A. Durra, D. Chapman, R. B. Vallee, and S. P. Gross (2016), Load-induced enhancement of Dynein force production by LIS1–NudE in vivo and in vitro, *Nature Communications*, *7*(1), 12,259, doi: 10.1038/ncomms12259.

Reiner, O., R. Carrozzo, Y. Shen, M. Wehnert, F. Faustinella, W. B. Dobyns, C. T. Caskey, and D. H. Ledbetter (1993), Isolation of a Miller–Dicker lissencephaly gene containing G protein β-subunit-like repeats, *Nature*, *364*(6439), 717–721, doi: 10.1038/364717a0.

Ren, J., Y. Zhang, S. Wang, L. Huo, J. Lou, and W. Feng (2018), Structural Delineation of the Neck Linker of Kinesin-3 for Processive Movement, *Journal of Molecular Biology*, *430*(14), 2030–2041, doi: 10.1016/j.jmb.2018.05.010.

Rice, L. M., E. A. Montabana, and D. A. Agard (2008), The lattice as allosteric effector: Structural studies of - and -tubulin clarify the role of GTP in microtubule assembly, *Proceedings of the National Academy of Sciences*, *105*(14), 5378–5383, doi: 10.1073/pnas.0801155105.

Rice, S., et al. (1999a), A structural change in the kinesin motor protein that drives motility, *Nature*, *402*(6763), 778–784, doi: 10.1038/45483.

Rice, S., et al. (1999b), A structural change in the kinesin motor protein that drives motility, *Nature*, doi: 10.1038/45483.

Richard, J., E. D. Kim, H. Nguyen, C. D. Kim, and S. Kim (2016), Allostery wiring map for kinesin energy transduction and its evolution, *Journal of Biological Chemistry*, *291*(40), 20,932–20,945, doi: 10.1074/jbc.M116.733675.

Rigbolt, K. T. G., et al. (2011), System-Wide Temporal Characterization of the Proteome and Phosphoproteome of Human Embryonic Stem Cell Differentiation, *Science Signaling*, *4*(164), rs3–rs3, doi: 10.1126/scisignal.2001570.

Robert, X., and P. Gouet (2014), Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Research*, *42*(W1), W320–W324, doi: 10.1093/nar/gku316.

Rosenfeld, S. S., G. M. Jefferson, and P. H. King (2001), ATP Reorients the Neck Linker of Kinesin in Two Sequential Steps, *Journal of Biological Chemistry*, *276*(43), 40,167–40,174, doi: 10.1074/jbc.M103899200.

Rosenfeld, S. S., P. M. Fordyce, G. M. Jefferson, P. H. King, and S. M. Block (2003), Stepping and Stretching, *Journal of Biological Chemistry*, *278*(20), 18,550–18,556, doi: 10.1074/jbc.M300849200.

Rosenfeld, S. S., M. van Duffelen, W. M. Behnke-Parks, C. Beadle, J. Corrreia, and J. Xing (2009), The ATPase Cycle of the Mitotic Motor CENP-E, *Journal of Biological Chemistry*, *284*(47), 32,858–32,868, doi: 10.1074/jbc.M109.041210.

RStudio (2013), Shiny: A Web Application Framework for R., *http://shiny.rstudio.com.*

Sack, S., J. Müller, A. Marx, M. Thormählen, E.-M. Mandelkow, S. T. Brady, and E. Mandelkow (1997), X-ray Structure of Motor and Neck Domains from Rat Brain Kinesin † , ‡, *Biochemistry, 36*(51), 16,155–16,165, doi: 10.1021/bi9722498.

Šali, A., and T. L. Blundell (1993), Comparative Protein Modelling by Satisfaction of Spatial Restraints, *Journal of Molecular Biology, 234*(3), 779–815, doi: 10.1006/jmbi.1993.1626.

Saunders, A. M., J. Powers, S. Strome, and W. M. Saxton (2007), Kinesin-5 acts as a brake in anaphase spindle elongation, *Current Biology, 17*(12), R453–R454, doi: 10.1016/j.cub.2007.05.001.

Scarabelli, G., and B. J. Grant (2013), Mapping the Structural and Dynamical Features of Kinesin Motor Domains, *PLoS Computational Biology*, doi: 10.1371/journal.pcbi.1003329.

Scarabelli, G., V. Soppina, X.-Q. Yao, J. Atherton, C. A. Moores, K. J. Verhey, and B. J. Grant (2015), Mapping the Processivity Determinants of the Kinesin-3 Motor Domain, *Biophysical Journal, 109*(8), 1537–1540, doi: 10.1016/j.bpj.2015.08.027.

Shang, Z., K. Zhou, C. Xu, R. Csencsits, J. C. Cochran, and C. V. Sindelar (2014), High-resolution structures of kinesin on microtubules provide a basis for nucleotide-gated force-generation, *eLife, 3*, doi: 10.7554/eLife.04686.

Shastry, S., and W. O. Hancock (2010), Neck Linker Length Determines the Degree of Processivity in Kinesin-1 and Kinesin-2 Motors, *Current Biology, 20*(10), 939–943, doi: 10.1016/j.cub.2010.03.065.

Shastry, S., and W. O. Hancock (2011), Interhead tension determines processivity across diverse N-terminal kinesins, *Proceedings of the National Academy of Sciences, 108*(39), 16,253–16,258, doi: 10.1073/pnas.1102628108.

Shen, M.-y., and A. Sali (2006), Statistical potential for assessment and prediction of protein structures, *Protein Science, 15*(11), 2507–2524, doi: 10.1110/ps.062416606.

Shilatifard, A. (2006), Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression, *Annual Review of Biochemistry, 75*(1), 243–269, doi: 10.1146/annurev.biochem.75.103004.142422.

Shojania Feizabadi, M., B. R. Janakaloti Narayanareddy, O. Vadpey, Y. Jun, D. Chapman, S. Rosenfeld, and S. P. Gross (2015), Microtubule C-Terminal Tails Can Change Characteristics of Motor Force Production, *Traffic*, doi: 10.1111/tra.12307.

Sindelar, C. V. (2011), A seesaw model for intermolecular gating in the kinesin motor protein, doi: 10.1007/s12551-011-0049-4.

Sindelar, C. V., and K. H. Downing (2007), The beginning of kinesin's force-generating cycle visualized at 9-Å resolution, *Journal of Cell Biology*, doi: 10.1083/jcb.200612090.

Sindelar, C. V., and K. H. Downing (2010), An atomic-level mechanism for activation of the kinesin molecular motors, *Proceedings of the National Academy of Sciences*, *107*(9), 4111–4116, doi: 10.1073/pnas.0911208107.

Sindelar, C. V., M. J. Budny, S. Rice, N. Naber, R. Fletterick, and R. Cooke (2002), Two conformations in the human kinesin power stroke defined by X-ray crystallography and EPR spectroscopy, *Nature Structural Biology*, doi: 10.1038/nsb852.

Sirajuddin, M., L. M. Rice, and R. D. Vale (2014), Regulation of microtubule motors by tubulin isotypes and post-translational modifications, *Nature Cell Biology*, doi: 10.1038/ncb2920.

Skiniotis, G. (2003), Nucleotide-induced conformations in the neck region of dimeric kinesin, *The EMBO Journal*, *22*(7), 1518–1528, doi: 10.1093/emboj/cdg164.

Skjærven, L., X. Q. Yao, G. Scarabelli, and B. J. Grant (2014), Integrating protein structural dynamics and evolutionary analysis with Bio3D, *BMC Bioinformatics*, *15*(1), 1–11, doi: 10.1186/s12859-014-0399-6.

Skjærven, L., J. Cuellar, A. Martinez, and J. M. Valpuesta (2015), Dynamics, flexibility, and allostery in molecular chaperonins, doi: 10.1016/j.febslet.2015.06.019.

Skjærven, L., S. Jariwala, X.-Q. Yao, and B. J. Grant (2016), Online interactive analysis of protein structure ensembles with Bio3D-web, *Bioinformatics*, (July), btw482, doi: 10.1093/bioinformatics/btw482.

Skoumpla, K., A. T. Coulton, W. Lehman, M. A. Geeves, and D. P. Mulvihill (2007), Acetylation regulates tropomyosin function in the fission yeast Schizosaccharomyces pombe, *Journal of Cell Science*, *120*(9), 1635–1645, doi: 10.1242/jcs.001115.

Sol, E. M., S. A. Wagner, B. T. Weinert, A. Kumar, H.-S. Kim, C.-X. Deng, and C. Choudhary (2012), Proteomic Investigations of Lysine Acetylation Identify Diverse Substrates of Mitochondrial Deacetylase Sirt3, *PLoS ONE*, *7*(12), e50,545, doi: 10.1371/journal.pone.0050545.

Soppina, V., S. R. Norris, A. S. Dizaji, M. Kortus, S. Veatch, M. Peckham, and K. J. Verhey (2014), Dimerization of mammalian kinesin-3 motors results in superprocessive motion, *Proceedings of the National Academy of Sciences*, *111*(15), 5562–5567, doi: 10.1073/pnas.1400759111.

Suhre, K., and Y. H. Sanejouand (2004), ElNémo: A normal mode web server for protein movement analysis and the generation of templates for molecular replacement, *Nucleic Acids Research*, doi: 10.1093/nar/gkh368.

Svoboda, K., and S. M. Block (1994), Force and velocity measured for single kinesin molecules, *Cell*, *77*(5), 773–784, doi: 10.1016/0092-8674(94)90060-4.

Terman, J. R., and A. Kashina (2013), Post-translational modification and regulation of actin, doi: 10.1016/j.ceb.2012.10.009.

Ti, S.-C., et al. (2016), Mutations in Human Tubulin Proximal to the Kinesin-Binding Site Alter Dynamic Instability at Microtubule Plus- and Minus-Ends, *Developmental Cell*, *37*(1), 72–84, doi: 10.1016/j.devcel.2016.03.003.

Tiwari, S. P., E. Fuglebakk, S. M. Hollup, L. Skjærven, T. Cragnolini, S. H. Grindhaug, K. M. Tekle, and N. Reuter (2014), WEBnmat v2.0: Web server and services for comparing protein flexibility, *BMC Bioinformatics*, doi: 10.1186/s12859-014-0427-6.

Tomishige, M., and R. D. Vale (2000), Controlling Kinesin by Reversible Disulfide Cross-Linking, *The Journal of Cell Biology*, *151*(5), 1081–1092, doi: 10.1083/jcb.151.5.1081.

Turner, J., R. Anderson, J. Guo, C. Beraud, R. Fletterick, and R. Sakowicz (2001), Crystal Structure of the Mitotic Spindle Kinesin Eg5 Reveals a Novel Conformation of the Neck-linker, *Journal of Biological Chemistry*, *276*(27), 25,496–25,502, doi: 10.1074/jbc.M100395200.

Uchimura, S., Y. Oguchi, Y. Hachikubo, S. Ishiwata, and E. Muto (2010), Key residues on microtubule responsible for activation of kinesin ATPase., *The EMBO journal*, *29*(7), 1167–1175, doi: 10.1038/emboj.2010.25.

Valdar, W. S. (2002), Scoring residue conservation, *Proteins: Structure, Function and Genetics*, doi: 10.1002/prot.10146.

Vale, R. D., and R. J. Fletterick (1997), The Design Plan of Kinesin Motors, *Annual Review of Cell and Developmental Biology*, *13*(1), 745–777, doi: 10.1146/annurev.cellbio.13.1.745.

Valentine, M. T., P. M. Fordyce, T. C. Krzysiak, S. P. Gilbert, and S. M. Block (2006), Individual dimers of the mitotic kinesin motor Eg5 step processively and support substantial loads in vitro, doi: 10.1038/ncb1394.

van Aalten, D., D. Conn, B. de Groot, H. Berendsen, J. Findlay, and A. Amadei (1997), Protein dynamics derived from clusters of crystal structures, *Biophysical Journal*, *73*(6), 2891–2896, doi: 10.1016/S0006-3495(97)78317-6.

VanBuren, V., D. J. Odde, and L. Cassimeris (2002), Estimates of lateral and longitudinal bond energies within the microtubule lattice, *Proceedings of the National Academy of Sciences*, *99*(9), 6035–6040, doi: 10.1073/pnas.092504999.

VanBuren, V., L. Cassimeris, and D. J. Odde (2005), Mechanochemical Model of Microtubule Structure and Self-Assembly Kinetics, *Biophysical Journal*, *89*(5), 2911–2926, doi: 10.1529/biophysj.105.060913.

Vemu, A., J. Atherton, J. O. Spector, A. Szyk, C. A. Moores, and A. Roll-Mecak (2016), Structure and Dynamics of Single-isoform Recombinant Neuronal Human Tubulin, *Journal of Biological Chemistry*, *291*(25), 12,907–12,915, doi: 10.1074/jbc.C116.731133.

Vemu, A., J. Atherton, J. O. Spector, C. A. Moores, and A. Roll-Mecak (2017), Tubulin isoform composition tunes microtubule dynamics, *Molecular Biology of the Cell*, *28*(25), 3564–3572, doi: 10.1091/mbc.e17-02-0124.

Venere, M., et al. (2015), The mitotic kinesin KIF11 is a driver of invasion, proliferation, and self-renewal in glioblastoma, *Science Translational Medicine*, *7*(304), 304ra143–304ra143, doi: 10.1126/scitranslmed.aac6762.

Verdone, L., M. Caserta, and E. D. Mauro (2005), Role of histone acetylation in the control of gene expression, *Biochemistry and Cell Biology*, *83*(3), 344–353, doi: 10.1139/o05-041.

Verhey, K. J., and J. Gaertig (2007), The Tubulin Code, *Cell Cycle*, *6*(17), 2152–2160, doi: 10.4161/cc.6.17.4633.

Vicente, J. J., and L. Wordeman (2015), Mitosis, microtubule dynamics and the evolution of kinesins, *Experimental Cell Research*, doi: 10.1016/j.yexcr.2015.02.010.

von Loeffelholz, O., N. A. Venables, D. R. Drummond, M. Katsuki, R. Cross, and C. A. Moores (2017), Nucleotide– and Mal3-dependent changes in fission yeast microtubules suggest a structural plasticity view of dynamics, *Nature Communications*, *8*(1), 2110, doi: 10.1038/s41467-017-02241-5.

Wagner, J. R., C. T. Lee, J. D. Durrant, R. D. Malmstrom, V. A. Feher, and R. E. Amaro (2016), Emerging Computational Methods for the Rational Discovery of Allosteric Drugs, *Chemical Reviews*, p. acs.chemrev.5b00631, doi: 10.1021/acs.chemrev.5b00631.

Walczak, C. E., S. Gayek, and R. Ohi (2013), Microtubule-Depolymerizing Kinesins, *Annual Review of Cell and Developmental Biology*, *29*(1), 417–441, doi: 10.1146/annurev-cellbio-101512-122345.

Wells, D. B., and A. Aksimentiev (2010), Mechanical Properties of a Complete Microtubule Revealed through Molecular Dynamics Simulation, *Biophysical Journal*, *99*(2), 629–637, doi: 10.1016/j.bpj.2010.04.038.

Widlund, P. O., M. Podolski, S. Reber, J. Alper, M. Storch, A. A. Hyman, J. Howard, and D. N. Drechsel (2012), One-step purification of assembly-competent tubulin from diverse eukaryotic sources, *Molecular Biology of the Cell*, *23*(22), 4393–4401, doi: 10.1091/mbc.e12-06-0444.

Wieczorek, M., S. Bechstedt, S. Chaaban, and G. J. Brouhard (2015), Microtubule-associated proteins control the kinetics of microtubule nucleation, *Nature Cell Biology*, *17*(7), 907–916, doi: 10.1038/ncb3188.

Williams, R. W., A. Chang, D. Juretić, and S. Loughran (1987), Secondary structure predictions and medium range interactions, *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, *916*(2), 200–204, doi: 10.1016/0167-4838(87)90109-9.

Xie, Y., R. Kang, X. Sun, M. Zhong, J. Huang, D. J. Klionsky, and D. Tang (2015), Posttranslational modification of autophagy-related proteins in macroautophagy, *Autophagy*, *11*(1), 28–45, doi: 10.4161/15548627.2014.984267.

Yildiz, A., M. Tomishige, A. Gennerich, and R. D. Vale (2008), Intramolecular Strain Coordinates Kinesin Stepping Behavior along Microtubules, *Cell*, doi: 10.1016/j.cell.2008.07.018.

Yu, I., C. P. Garnham, and A. Roll-Mecak (2015), Writing and reading the tubulin code, *Journal of Biological Chemistry*, doi: 10.1074/jbc.R115.637447.

Zhang, R., and E. Nogales (2015), A new protocol to accurately determine microtubule lattice seam location, *Journal of Structural Biology*, *192*(2), 245–254, doi: 10.1016/j.jsb.2015.09.015.

Zhang, R., G. M. Alushin, A. Brown, and E. Nogales (2015), Mechanistic origin of microtubule dynamic instability and its modulation by EB proteins, *Cell*, *162*(4), 849–859, doi: 10.1016/j.cell.2015.07.012.

Zimmermann, M. T., A. Kloczkowski, and R. L. Jernigan (2011), MAVENs: Motion analysis and visualization of elastic networks and structural ensembles, *BMC Bioinformatics*, doi: 10.1186/1471-2105-12-264.