THE UNIVERSITY OF WARWICK

Library Declaration and Deposit Agreement

1. STUDENT DETAILS

Please complete the following:

Full name: Neil Venables

University ID number: 1150112

2. THESIS DEPOSIT

2.1 Under your registration at the University, you are required to deposit your thesis with the University in BOTH hard copy and in digital format. The digital copy should normally be saved as a single pdf file.

2.2 The hard copy will be housed in the University Library. The digital copy will be deposited in the University's Institutional Repository (WRAP). Unless otherwise indicated (see 2.6 below), this will be made immediately openly accessible on the Internet and will be supplied to the British Library to be made available online via its Electronic Theses Online Service (EThOS) service.
[At present, theses submitted for a Master's degree by Research (MA, MSc, LLM, MS or MMedSci) are not being deposited in WRAP and not being made available via EthOS. This may change in future.]

2.3 In exceptional circumstances, the Chair of the Board of Graduate Studies may grant permission for an embargo to be placed on public access to the thesis **in excess of two years**. This must be applied for when submitting the thesis for examination (further information is available in the *Guide to Examinations for Higher Degrees by Research*.)

2.4 If you are depositing a thesis for a Master's degree by Research, the options below only relate to the hard copy thesis.

2.5 If your thesis contains material protected by third party copyright, you should consult with your department, and if appropriate, deposit an abridged hard and/or digital copy thesis.

2.6 Please tick one of the following options for the availability of your thesis (guidance is available in the *Guide to Examinations for Higher Degrees by Research)***:**

- [ ] Both the hard and digital copy thesis can be made publicly available immediately
- [x] The hard copy thesis can be made publicly available immediately and the digital copy thesis can be made publicly available after a period of two years (*should you subsequently wish to reduce the embargo period please inform the Library*)
- [ ] Both the hard and digital copy thesis can be made publicly available after a period of two years (*should you subsequently wish to reduce the embargo period please inform the Library*)
- [ ] Both the hard copy and digital copy thesis can be made publicly available after _____ (insert time period in excess of two years). **This option requires the prior approval of the Chair of the Board of Graduate Studies (see 2.3 above)**

The University encourages users of the Library to utilise theses as much as possible, and unless indicated below users will be able to photocopy your thesis.

- [ ] I **do not** wish for my thesis to be photocopied

3. GRANTING OF NON-EXCLUSIVE RIGHTS

Whether I deposit my Work personally or through an assistant or other agent, I agree to the following:

Rights granted to the University of Warwick and the British Library and the user of the thesis through this agreement are non-exclusive. I retain all rights in the thesis in its present version or future versions. I agree that the institutional repository administrators and the British Library or their agents may, without changing content, digitise and migrate the thesis to any medium or format for the purpose of future preservation and accessibility.

4. DECLARATIONS

I DECLARE THAT:

I am the author and owner of the copyright in the thesis and/or I have the authority of the authors and owners of the copyright in the thesis to make this agreement. Reproduction of any part of this thesis for teaching or in academic or other forms of publication is subject to the normal limitations on the use of copyrighted materials and to the proper and full acknowledgement of its source.

The digital version of the thesis I am supplying is either the same version as the final, hard-bound copy submitted in completion of my degree once any minor corrections have been completed, or is an abridged version (see 2.5 above).

I have exercised reasonable care to ensure that the thesis is original, and does not to the best of my knowledge break any UK law or other Intellectual Property Right, or contain any confidential material.

I understand that, through the medium of the Internet, files will be available to automated agents, and may be searched and copied by, for example, text mining and plagiarism detection software.

At such time that my thesis will be made publically available digitally (see 2.6 above), I grant the University of Warwick and the British Library a licence to make available on the Internet the thesis in digitised format through the Institutional Repository and through the British Library via the EThOS service.

If my thesis does include any substantial subsidiary material owned by third-party copyright holders, I have sought and obtained permission to include it in any version of my thesis available in digital format and that this permission encompasses the rights that I have granted to the University of Warwick and to the British Library.

5. LEGAL INFRINGEMENTS

I understand that neither the University of Warwick nor the British Library have any obligation to take legal action on behalf of myself, or other rights holders, in the event of infringement of intellectual property rights, breach of contract or of any other right, in the thesis.

Please sign this agreement and ensure it is bound into the final hard bound copy of your thesis, which should be submitted to Student Reception, Senate House.

Student's signature: N. Venables …… Date: 11/01/2016

THE ELECTROSTATIC ANALYSIS OF TOG-DOMAINS FROM XMAP215/DIS1 FAMILY MEMBERS

Neil Anthony Venables

September, 2015

University of Warwick

Systems Biology Doctoral Training Centre

THE ELECTROSTATIC ANALYSIS OF TOG-DOMAINS FROM XMAP215/DIS1 FAMILY MEMBERS

Neil Anthony Venables

1. Reviewer David Sept
Department of Biomedical Engineering
University of Michigan

2. Reviewer Stephen Royle
Division of Biomedical Cell Biology
University of Warwick

Supervisors Robert Cross and Till Bretschneider

September, 2015

Neil Anthony Venables

THE ELECTROSTATIC ANALYSIS OF TOG-DOMAINS FROM XMAP215/DIS1 FAMILY MEMBERS

Reviewers: David Sept and Stephen Royle

Supervisors: Robert Cross and Till Bretschneider

University of Warwick

Systems Biology Doctoral Training Centre

Senate House, Gibbet Hill Road

Coventry, CV4 7AL

ABSTRACT

TOG domain MT polymerases are catalysts of MT dynamics that track MT tips and chaperone tubulin exchange by mechanisms that are not yet understood. In this work, we use computational simulation to probe the detailed mechanisms by which TOGs capture and manipulate GTP-tubulin.

Natural TOGs display a ridge of basic surface loops that forms the core of the TOG-tubulin interface. Computational mutagenesis shows that these basic loops play a dominant role in setting the overall electrostatic field on the TOG domain, and that natural TOGs fall into subclasses depending on the detailed structure of these fields. Normal mode analysis reveals that natural TOG domains show characteristic patterns of flexibility that define their position within the TOG array. Nonetheless all TOGs are sufficiently stiff that the range of positions explored by the domains' common secondary structure is heavily restricted. Brownian dynamics simulations establish that diffusion-to-capture of tubulin by TOG domains is very strongly electrostatically steered. In all trajectories examined, TOGs were initially captured and oriented by tubulin to a degree that reflects their simulated association rates.

To be effective, TOG domain MT polymerases need to capture GTP-tubulin rapidly and specifically from solution and configure it so that it incorporates readily at the growing MT tip. Our data show that TOGs do this **(1)** by optimising long range, electrostatically-steered diffusion-to-capture, which is important for creating a tethered complex at the MT tip, and **(2)** by using conformational selection within the tethered complex to drive GTP-tubulin into conformation(s) that favours assembly and dissociation of the complex upon lattice incorporation.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisors Prof. Robert Cross and Dr. Till Bretschneider. I have learnt a lot from both of you over the last few years. I am especially thankful to Rob for diligently reading over this thesis and providing invaluable comments on drafts.

I am grateful to all members of the Cross and Bretschneider labs past and present. Colleagues in both groups have contributed immensely to my understanding of the field. I want to thank in particular Dr. Miho Katsuki and Dr. Douglas Drummond who were instrumental in providing help and advice on experimental techniques.

Next, and most importantly, I would like to thank my parents for their continual support and for enduring my many years of study. Last, but certainly not least, I would like to thank Faye. You made all this worthwhile.

Contents

1 INTRODUCTION **1**

1.1 Cytoskeleton 1

1.2 Microtubules 1

1.2.1 Microtubule Structure and Organisation 2

1.2.2 Microtubule Dynamics 8

1.3 XMAP215/Dis1 Family 15

1.3.1 Structure 15

1.3.2 Tubulin Binding 17

1.3.3 Catalysis 18

1.3.4 Differential TOG-activity 19

1.3.5 TOG domain Structures 21

1.3.6 CLASP 22

1.4 Theory 24

1.4.1 Continuum Electrostatics 24

1.4.2 Solving Poisson-Boltzmann Equation 26

1.4.3 Normal Mode Analysis 26

1.4.4 Brownian Dynamics 29

1.5 Motivation 33

2 STRUCTURAL & ELECTROSTATIC ANALYSIS **34**

2.1 Introduction 34

2.2 Aims 36

2.3 Results & Discussion 37

2.3.1 Cumulative electrostatic similarity distribution for TOG homologues 37

2.3.2 Perturbation maps of TOG electrostatic fields 38

2.3.3 Electrostatic clustering analysis 40

2.3.4 NMA of TOG-homologues 42

2.3.5 DCCM Analysis of Individual TOGs 45

2.3.6 Principal Component Analysis of TOG-Structures. 48

2.3.7 Electrostatic free energies of association. 49

2.4 Methods 56

2.4.1 Structure Preparation 56
2.4.2 Electrostatic Free Energies of Association 58
2.4.3 Normal Mode Analysis 58
2.4.4 Principal Component Analysis 58
2.5 Summary 59

3 BROWNIAN DYNAMICS OF TOG DOMAIN ASSOCIATION 60
3.1 Introduction 60
3.2 Aims 63
3.3 Results & Discussion 64
3.3.1 Density Maps of TOG Binding Site 64
3.3.2 Mixed Tubulin Association Rates 68
3.3.3 Effect of Mutation on Association Rate 71
3.3.4 BD Rotational & Translational Steering 74
3.3.5 TransComp Association Rates 86
3.3.6 Protein-protein Docking 90
3.4 Methods 93
3.4.1 Brownian Dynamics 93
3.4.2 Molecular Dynamics Refinement 93
3.4.3 Binding Maps 93
3.4.4 Occupancy Maps 94
3.4.5 Transient Complex 94
3.4.6 Tubulin Homology Modeling 95
3.5 Summary 95

4 THE DEVELOPMENT OF A MINIMAL +TIP SYSTEM 96
4.1 Introduction 96
4.2 Aims 97
4.3 Results & Discussion 97
4.3.1 Mal3 MT Plus-End Tracking *In Vitro* 97
4.3.2 Purification of Alp14-Avi 102
4.4 Materials & Methods 108
4.4.1 Tubulin Expression & Purification 108
4.4.2 Mal3 Expression & Purification 109
4.4.3 Alp14 Expression & Purification 110
4.4.4 Protein Detection Methods 112
4.4.5 Flow Cells 112
4.4.6 Microtubule Dynamics Assays 113
4.4.7 Microscopy 113

4.5 Summary 114

5 DISCUSSION & FUTURE WORK **115**

Bibliography **123**

List of Figures

1.1 Distribution of cytoskeletal filaments within a cell. 2
1.2 3D ribbon diagram of bovine αβ-tubulin. 4
1.3 Schematic representation of A- and B-lattice MTs 5
1.4 Inter-protofilament lateral contacts between GMPCPP and GDP tubulin. . . . 6
1.5 Deformation of MTs under load. 7
1.6 1D growth model of MT dynamic instability. 10
1.7 2D growth model of MT dynamic instability. 11
1.8 Composition of the MT tip structure. 13
1.9 Compaction and conformational change of α-tubulin 14
1.10 Domain organisation of XMAP215/Dis1 family members. 16
1.11 The structural decomposition of TOG domains. 17
1.12 Structure of TOG:αβ-tubulin complex 18
1.13 Catalysis of MT polymerisation by TOG proteins 19
1.14 TOG domain architecture and binding modes to MTs 20
1.15 Surface representation of Stu2 TOG1 and CLASP TOG2. 23
1.16 Division of the BD reaction scheme. 31

2.1 The electrostatic potential of a TOG:αβ-tubulin complex. 35
2.2 Conservation of electrostatic potential for TOG-domains homologues. 37
2.3 Mutation-based perturbation maps for TOG homologs. 39
2.4 Clustering of TOG domains by electrostatic potential 41
2.5 Atomic fluctuation profiles of TOG domain crystal structures 44
2.6 TOG domain correlation maps 47
2.7 Principal component analysis of TOG domains. 48
2.8 Electrostatic free energies of Tubulin-TOG association for cognate tubulin. . 52
2.9 Electrostatic free energies of Tubulin-TOG association for porcine tubulin. . 55

3.1 The spectrum of association rates. 61
3.2 Two-step binding model for a bimolecular reaction. 62
3.3 Density maps of TOG domain diffusion pathways. 67
3.4 TOG domain family association rates for porcine and cognate tubulin. 70
3.5 Cognate association rates for mutations of each TOG domain 71

3.6 Multiple sequence alignment of 12 TOG-domains. 73
3.7 Translational occupancy landscape. 79
3.8 Rotational occupancy landscape. 85
3.9 Location of the transient complex within the interaction energy landscape. . . 86
3.10 Predicted association rates for TOG family members. 88
3.11 Predicted association rates for selected alanine mutations in Alp14 TOG1. . 88
3.12 Protein docking results from PyDock. 92

4.1 Microtubule dynamics reconstituted *in vitro* 98
4.2 Titration of Mal3 from 300 nM to 0.5 nM 100
4.3 Effect of ionic strength on Mal3 binding 101
4.4 Western blot of Alp14-Avi pellets from insect cell expression 102
4.5 Elution profile of anion-exchange chromatograph column 103
4.6 Elution profile from avidin column under normal elution conditions. 104
4.7 Second elution profile of anion-exchange chromatography 105
4.8 Elution profile from avidin column under enhanced elution conditions. . . . 107

List of Tables

2.1 Homology modelled structures and associated templates 56

3.1 Association rates for porcine tubulin. 70
3.2 Association rates for cognate tubulin. 70
3.3 Association rates of TOG domains mutants to cognate tubulin. 72
3.4 List of charged residues mutated for each TOG domain. 72
3.5 TransComp predicted association rate for cognant tubulin 88
3.6 TransComp predicted association rates for select alanine mutations. 89
3.7 Correlations between PyDock energy score and iRMSD. 90

ABBREVIATIONS

AFM ..... Atomic force microscopy

alp14 ..... Altered polarity protein 14

BSA ..... Bovine serum albumin

C. elegans ..... *Caenorhabditis elegans*

ch-TOG ..... Colonic and hepatic tumour-overexpressed gene

CTT ..... Carboxyl-terminal tail

cyro-EM ..... Cyro-electron microscopy

CLASP ..... Cytoplasmic linker-associated protein

Da ..... Dalton (g mol^{-1})

Dis1 ..... Defective in sister chromatid separation

D. melanogaster ..... *Drosophila melanogaster*

DTT ..... Dithiothreitol

DDS ..... Dodecyl sulfoxide

DMSO ..... Dimethyl sulfoxide

EGFP ..... Enhanced green fluorescence protein

EGTA ..... Ethylene glycol tetracetic acid

EB ..... End binding protein

GDP ..... Guanosine-5'-diphosphate

GMPCPP ..... Guanosine-(α,β)-methylene-triphosphate

GTP ..... Guanosine triphosphate

GTPγS ..... Guanosine 5'-[γ-thio]-triphosphate

GTPase ..... Family of enzymes that bind to and hydrolyse GTP

HR ..... Heat repeat

HEAT ..... Huntingdon elongation factory 3, protein phosphatase 2A and yeast kinase TOR1

IPTG ..... Isopropyl β-D-1-triogalaclopyranoside

K-fibres ..... Kinetochore fibres

LN_2 ..... Liquid nitrogen

MAP ..... Microtubule associated protein

MSPS ..... Mini-spindles of *D. melanogaster*

MOR1 ..... *A. thaliana* homologue of XMAP215

MT ..... Microtubule

NMA ..... Normal mode analysis

PBS ..... Phosphate-buffered saline

PEM ..... PIPES-EGTA-$MgCl_2$

PFs ..... Protofilaments

PIPES ..... 1,4-Piperazinediethanesulfonic acid

S. cerevasisae ..... *Saccharomyces cerevisiae*

S. pombe ..... *Saccharomyces pombe*

SDS ..... Sodium dodecyl sulfate

Stu2 ..... Suppressor of tubulin 2 protein

Sf9 ..... *Spodoptera frugiperda* 9 cells

TOG ..... Tumour overexpression gene

X. laevis ..... *Xenopus laevis*

XMAP215 ..... *Xenopus laevis* MT associated protein at 215 kDa

Zyg9 ..... Zygote defective embryonic lethal protein 9

DECLARATION

I undertake that all the material presented here for examination is my own work and has not been written for me, in whole or part, by any other person. I also undertake that any quotation or paraphrase from the published or unpublished work of another person has been duly acknowledged in the work which I present for examination.

Coventry, September, 2015

N. Venables

Neil Anthony Venables

1 INTRODUCTION

1.1 Cytoskeleton

The cytoskeleton is a dynamic, three-dimensional scaffold that spans the cytoplasm of all eukaryotic cells. This supramolecular structure is composed of a system of interacting filaments: actin, microtubules, and intermediate filaments (Fig 1.1). Each of these filaments is assembled from the association of many repeated, non-covalently bonded protein subunits. These protein polymers are highly dynamic structures that reassemble at various locations in the cell. Together, these filaments form structures that enable cells to re-organise their internal components and interact mechanically with their environment. More importantly, the cytoskeleton collectively supports many essential cellular functions such as cell migration, intracellular transport and cell division.

This thesis explores the mechanisms of tubulin exchange by tumor overexpressed gene (TOG) domain polymerases at the microtubule tip. To understand how these proteins modulate tubulin binding and release, we first need to understand the structural and functional properties of single microtubules.

1.2 Microtubules

Microtubules (MTs) are dynamic polymers, assembled by the self-association of $\alpha\beta$-tubulin heterodimers. Tubulin subunits couple head-to-tail to form structurally polarised, hollow tubes with a diameter of 25 nm. Typically, each MT is formed by the lateral association of 13 protofilaments (Tilney et al., 1973) (Mandelkow et al., 1986). MTs switch stochastically between phases of growth, pause, catastrophe, shrinkage and rescue in a process termed dynamic instability (Mitchison and Kirschner, 1984). Tubulin is a GTPase and the polymerisation of tubulin into MTs is coupled to GTP turnover. Dynamic instability is driven principally by the dynamics of a variable-sized cap of GTP-tubulin that develops at the growing tips. Dynamic instability is the mechanism by which MTs explore intracellular space. It allows MTs to search space and capture cellular components, as exemplified by the capture

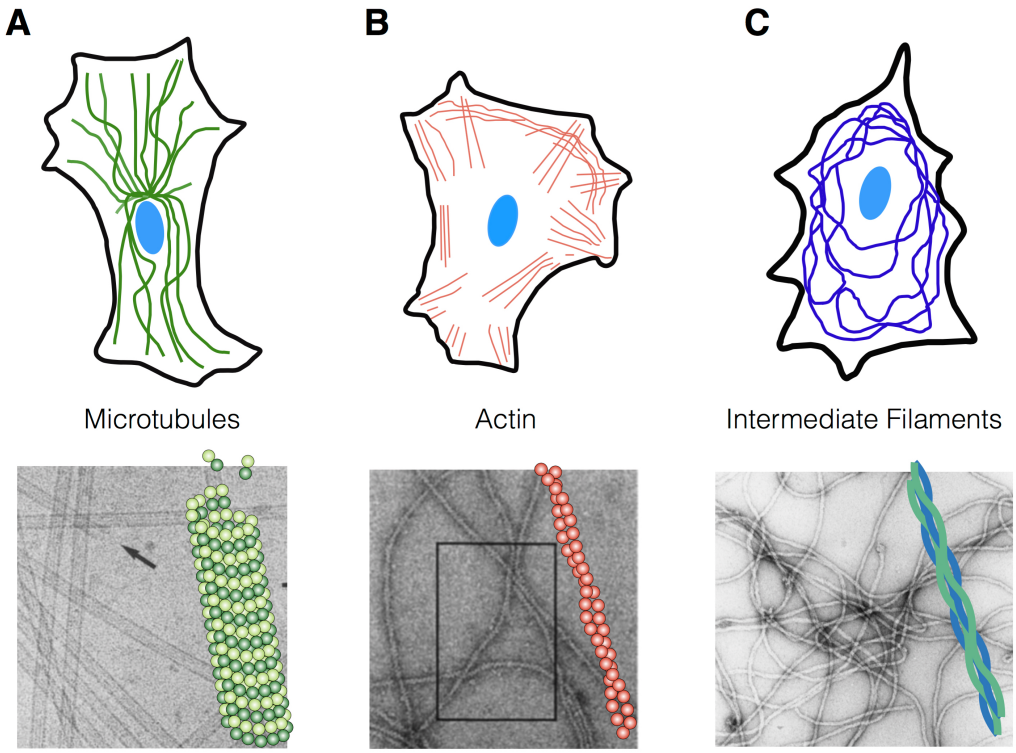


Figure 1.1: The distribution of cytoskeletal filaments within a cell. (**A**) Microtubules are long, hollow cylindrical filaments with a diameter of 25 nm and are composed of the protein tubulin. Microtubules typically radiate outwards from a microtubule-organising center. (**B**) Filamentous actin (F-actin) is a double-stranded polymer assembled from the globular monomer G-actin, and was discovered in skeletal muscle alongside the motor protein, myosin. Actin filaments measure ~7 nm in diameter with a helix repeating every 37 nm. In many cell types, actin filaments form a network that resides beneath the cell cortex. (**C**) Intermediate filaments have an average diameter of 10 nm and are highly deformable structures due to their hierarchical composition. The principle function of intermediate filaments is to provide mechanical support to the nuclear membrane and plasma membrane. Unlike the other filaments, they do not participate in cell motility. Figure adapted from (Alberts et al., 2014), p968. EM images were sourced in order from (Mandelkow et al., 1991) (Bremer and Aebi, 1992) (Mücke et al., 2004).

and alignment of a cell's chromosomes by the K-fibres of the mitotic spindle (Maddox et al., 2002).

1.2.1 Microtubule Structure and Organisation

$\alpha\beta$-tubulin

The subunit of MTs is tubulin, a heterodimer composed of two monomers: α- and β-tubulin (Krauhs et al., 1981). Tubulin was first purified using its affinity for colchicine (Weisenberg et al., 1968), a natural toxin obtained from the plant, *Colchicum autumnale*. α- and β-tubulin have an approximately equal mass of 55 kDa and show 40% sequence similarity at the

amino acid level (Burns and Surridge, 1990). In most eukaryotes, the tubulin gene family encodes multiple tubulin isotypes (Hammond et al., 2008). These isoforms derive from the expression of multiple genes and differences in sequence are mostly located at negatively charged residues in the C-terminal tail. Both α- and β-tubulin possess carboxyl-terminal tail (CTT) regions that project outwards from the microtubule surface and are highly mobile. CTTs are subject to numerous forms of post-translational modification, such as phosphorylation, acetylation, detyrosination and polyglutamylation. CTTs are thought to play critical roles in regulating microtubule structure and function (Aiken et al., 2014). The core fold of each tubulin monomer is composed of two β-sheets surrounded by α-helices (Fig. 1.2). α- and β-tubulin monomers associate non-covalently to form an extremely stable complex, whose dissociation is very thermodynamically unfavourable in the absence of protein co-factors (Caplow and Fee, 2002). Both α- and β-tubulin monomers bind GTP (MacNeal and Purich, 1978). Biochemical and structural studies show that key differences in conformation between the monomers occur during MT assembly. In the dimer, α-bound GTP is buried at the intra-dimer interface (N-site) and is non-exchangeable, whereas β-bound GTP at the E-site is hydrolysed to GDP during polymerisation. Tubulin subunits that become exposed at the MT tip by loss of the overlying tubulin, or that dissociate from the MT, can then exchange their GDP from the exposed surface site (E-site). The coupling of polymerisation to GTPase activity arises because the active site of β-tubulin only becomes catalytically competent following contribution of a glutamic acid residue by the overlying α-subunit (Nogales et al., 1998b) (Dai et al., 1994).

Protofilament Structure

Within a MT, GTP-tubulin subunits associate head-to-tail to form linear protofilaments. Protofilaments (PFs) can form two distinct geometrical arrangements, A- and B-lattices, which are mediated by lateral contacts (Fig. 1.3). This lateral interface occurs between the H1-S2 and H2-S3 loops of the nucleotide binding domain in one protofilament and the M-loop of the intermediate domain on the other (Fig. 1.4). The A- and B- lattice structures were first identified by cryo-EM and 3D reconstruction of flagella MTs (Amos and Klug, 1974). In an A-type lattice, lateral bonds between protofilaments arise from interactions between α- and β-monomers on adjacent protofilaments. In the B-type lattice, α-monomers laterally bind to α-monomers, and β-monomers with β-monomers, on adjacent protofilaments. In a 13-PF B-lattice MT, this pattern repeats itself except at the seam, where there is a dislocation equivalent to the length of three monomers between the neighbouring protofilaments, forming a single A-lattice seam at this junction. Structural evidence indicates that bonds between dimers occur at both the longitudinal and lateral interfaces (Chrétien and Fuller,

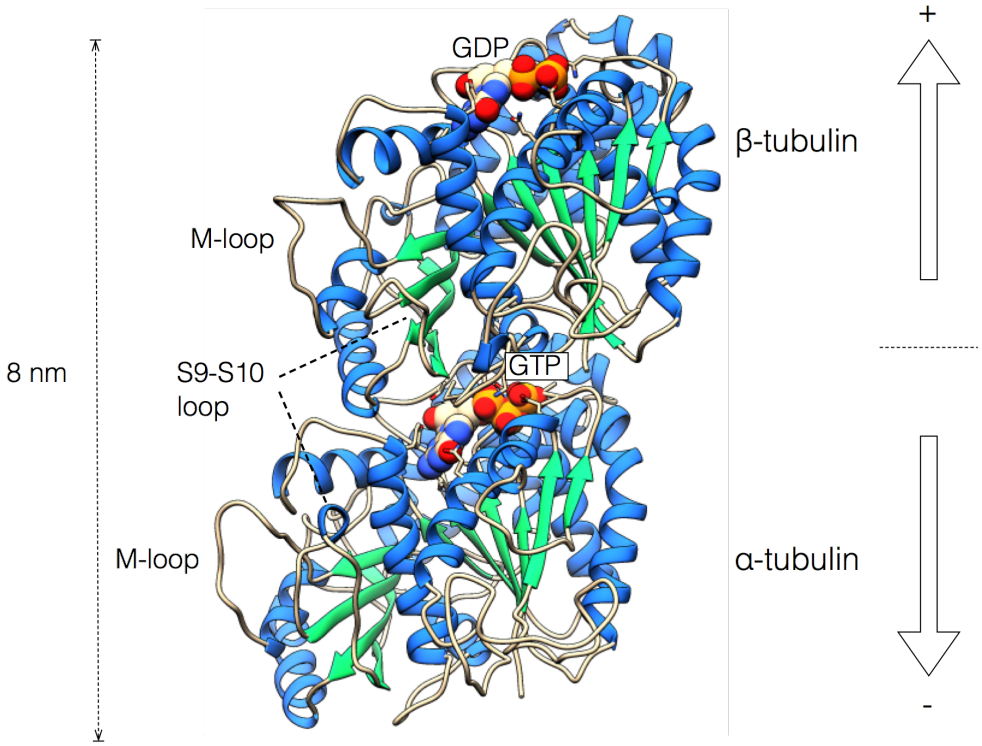


Figure 1.2: 3D ribbon diagram of bovine αβ-tubulin (PDB ID: 1JFF) (Lowe et al., 2001). The dimer consists of a core structure of two β-sheets (green) and 12 α-helices (blue). The nucleotide binding sites are located at the inter- and intra-dimer interfaces. GDP is bound to β-tubulin and α-tubulin contains GTP. Typically the cation Mg^{2+} is in complex with GTP. β-tubulin is closest to the plus-end and α-tubulin to the minus-end. The M-loop is required for establishing lateral tubulin contacts within the MT lattice and maintaining interactions with neighbouring subunits. The M-loop is stabilised in both subunits by the S9-S10 loops.

2000) (Chrétien et al., 1995). Longitudinal bonds are calculated to be stronger than lateral bonds (Sept et al., 2003). As lateral interactions at the seams are weaker, the seam is implicated as a site of modulation for controlling dynamic instability. A recent study has shown that the introduction of extra A-lattice seams into dynamic MTs promotes MT catastrophe by weakening the GTP cap (Katsuki et al., 2014). Computational modeling of lateral protofilament interactions indicates that the B-lattice is more energetically favourable than the A-lattice (Alushin et al., 2014). Overall, 13 protofilament B-lattice MTs are thought to be the most common type in cells (McIntosh et al., 2009).

Microtubule Polarity

MTs possess an intrinsic polarity due to the head-to-tail association of tubulin dimers. This polarisation of the lattice results in structural and kinetic differences at its ends. The faster growing and more dynamic end of a MT is termed the plus-end, and the slower and less

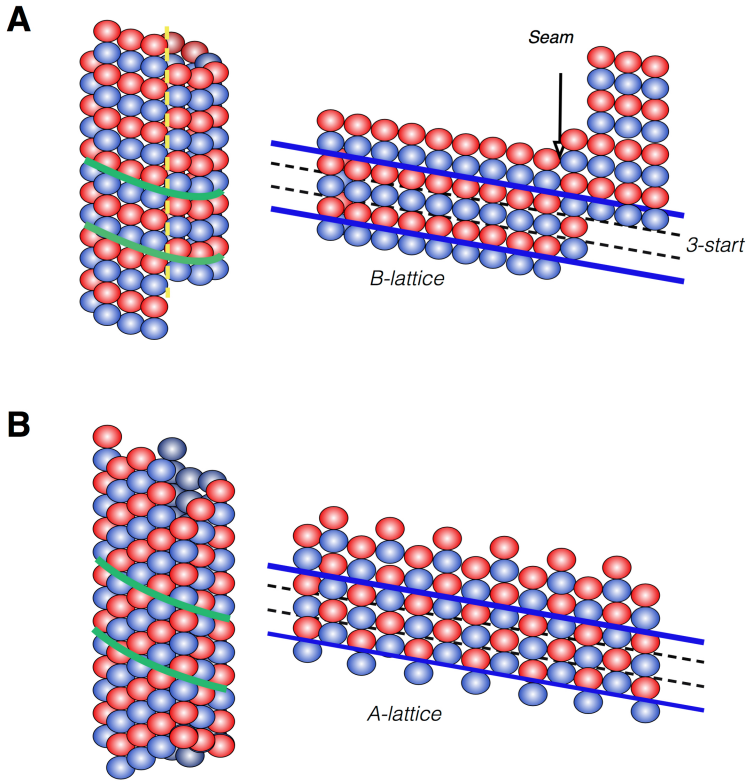


Figure 1.3: Schematic representation of A- and B-lattice MTs (**A,B**). Both lattice types are represented in 3D and as open-sheets. The solid-blue lines indicate the direction of the 3 start family of helices, where three tubulin monomers span each turn of the helix, common to each lattice type. The red and blue spheres are respectively α- and β-tubulin. Figure adapted from (Georges et al., 2008)

dynamic one is named the minus-end. Typically in cells, the MT growth rate is up to three-times faster at the plus-end than the minus-end (Kinoshita et al., 2001) (Zanic et al., 2013). Within a protofilament, tubulin dimers are orientated with their β-monomer pointing towards the plus-end and their α-monomer facing the minus-end of the MT. Direct evidence for this bias comes from biochemical experiments showing that GTP-coated fluorescent beads bind preferentially to MT plus ends, and beads coated with an antibody specific to α-tubulin bind to the minus-end (Mitchison, 1993) (Fan et al., 1996). MT polarity exists not only at the two ends of the structure, but all along the length of the lattice. MT polarity is central to the interactions of a variety of specialised motor and non-motor proteins (Baas, 1999).

Mechanical Properties

Understanding the mechanical properties of MTs is important because they have direct implications for MT assembly kinetics and the mechanism of force generation in cells. Early experimental studies, limited by spatial resolution, focused on measuring the flexural rigid-

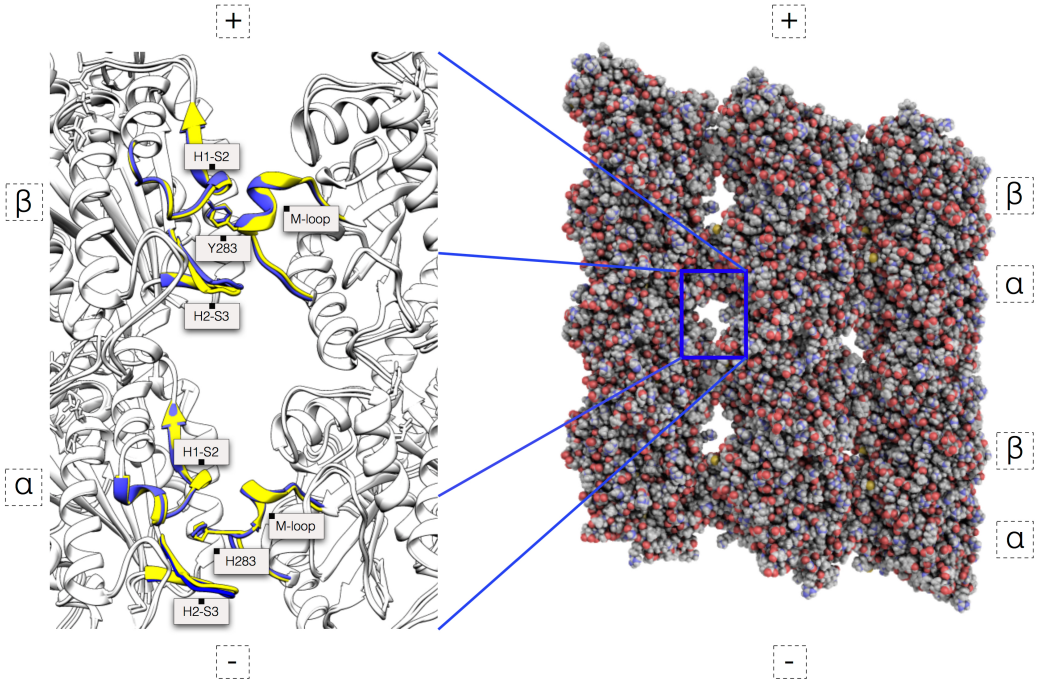


Figure 1.4: Inter-protofilament lateral contacts between GMPCPP (PDB code: 3J6E) and GDP (PDB code:3J6G) (Alushin et al., 2014) protofilaments. The GMPCPP structure is displayed in yellow and the GDP structures in blue. The two structures are aligned to show that lateral contacts are not affected by the tubulin nucleotide state. Key residues mediating lateral contacts are displayed in stick representation.

ity of MTs by optical microscopy (Gebremichael et al., 2008) (Asbury et al., 2011) (Venier et al., 1994). These mechanical studies arrived indirectly at a range of values for the Youngs modulus of a MT, determined in part by the choice of elastic model for each calculation (Cassimeris et al., 2001) (Hill, 1985) (Mickey and Howard, 1995). Unfortunately these experiments reached no consensus value and measurements differed by two orders of magnitude. Later, improvements in atomic force microscopy (AFM) led to direct measurements of microtubule deformation (Schaap et al., 2006) (Pablo et al., 2003). AFM improved on the earlier techniques because it allowed different modes of deformation to be applied locally to the structure. These studies revealed that single MTs under lateral deformation at low force experience an approximately linear response to the force applied. Longitudinal bonds between protofilaments are more resistant to mechanical force than lateral contacts. This evidence alongside prior thermodynamic calculations adds weight to the view that lateral bonds are weaker than longitudinal bonds (VanBuren et al., 2002). MTs experience irreversible failures in their structure at compressive forces beyond 0.3 nN and their resultant collapse follows a multi-step mechanism (Fig. 1.5). Initially, the MT lattice resists deformation and its response to force is linear. As the force increases, the developing mechanical strain breaks both longitudinal and lateral bonds. The lateral bonds dissociate first followed by the stronger longitudinal contacts. The MT structure collapses only when the longitudinal

bonds are lost. Interestingly, MTs can sustain a number of breaks in lateral contacts which can deform, unlike longitudinal bonds. More recently, these results were confirmed using computational modeling of MT structure under load *in silico* (Kononova et al., 2014). Simulations show a close match between simulated and experimental force-deformation spectra. This determined the free energies of dissociation for lateral (6.9 ± 0.4 kcal mol^{-1}) and longitudinal (14.9 ± 1.5 kcal mol^{-1}) breakage of inter-tubulin bonds and the flexural rigidity (18,000 - 26,000 pN nm^2) of tubulin protofilaments. Similarly, thermodynamic calculations estimate that ~1.7 kcal mol^{-1} of elastic strain energy is stored in the lattice by tubulin-GDP, based on measured angle differences from electron micrographs of curled, assembling protofilaments (Jánosi et al., 1998). These measurements support the conclusion that MTs are capable of generating or transmitting large mechanical forces in the cell. For instance, the motions of chromosomes are thought to be associated with depolymerising K-fibre MTs that are anchored to the poles. Depletion of all MT-dependent motors does not influence the speed of this movement in *S. pombe* cells (Grishchuk and McIntosh, 2006) (Tanaka et al., 2007). Direct *in vitro* assays have demonstrated force generation by depolymerising MTs. In these experiments, MTs were attached to beads coated with streptavidin and assembled from biotinylated tubulin (Grishchuk et al., 2005). The beads were trapped with laser tweezers and MTs were forced into depolymerisation by ablating the GTP-cap by photobleaching. The movement of the beads during shrinkage allowed a direct measurement of force.

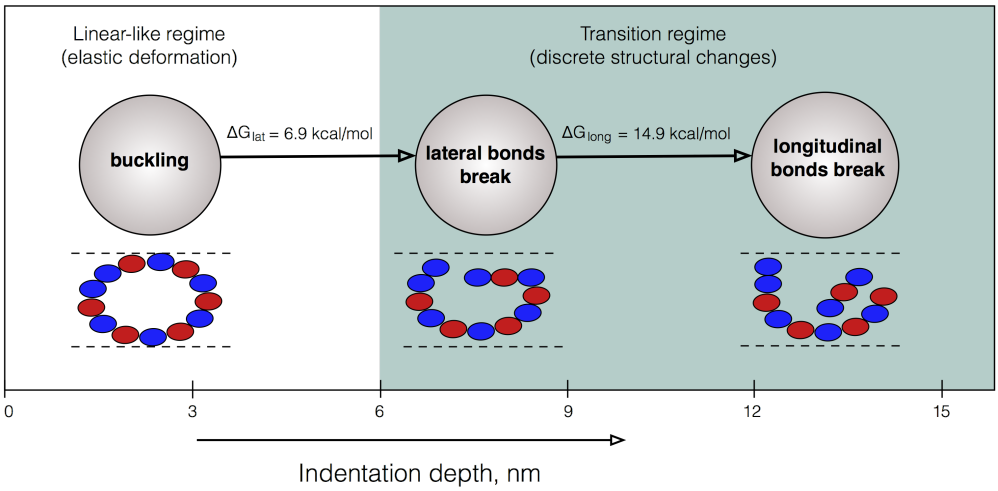


Figure 1.5: Microtubule deformation mechanics. The diagram illustrates the sequence of transitions during applied load to a MT. The white area indicates the linear regime for elastic deformations. The green area corresponds to the transition regime, characterised by rupture of lateral and longitudinal contacts. Figure adapted from (Kononova et al., 2014).

1.2.2 Microtubule Dynamics

Microtubule dynamic instability was first proposed based on the analysis of the length distributions of a fixed population of MTs (Mitchison and Kirschner, 1984). This distribution implied that, although these MTs exhibit a steady-state length, individual MTs interconvert between periods of growth and rapid shrinking. Dynamic instability is caused by a brief delay in hydrolysis of GTP-tubulin after polymerisation, which results in the formation of a GTP-rich region called the GTP-cap (Carlier and Pantaloni, 1981). This structure stabilises the tip due to strong lateral interactions between GTP-subunits, and predisposes the MT to growth because of its increased affinity for free GTP-tubulin subunits. When the cap is lost, the MT undergoes rapid GDP-tubulin dissociation, a process known as catastrophe. Lateral bonds between the exposed GDP-tubulin subunits break as they adopt their naturally curved conformation. When the GTP-cap of the depolymerising MT is restored, this triggers the transition back to growth - a process known as rescue. Although this model offers an attractive conceptual explanation for understanding dynamic instability, the model fails to account for the increasing complexity seen in MT structure in recent years. In the sections that follow, the evolution of our understanding of dynamic instability, with emphasis on the regulatory role of the GTP-cap, will be discussed.

GTP-cap model

The canonical structural explanation of dynamic instability is based upon high-resolution EM studies showing the structural details of growing and shrinking MTs (Chrétien et al., 1995) (Mandelkow et al., 1991). EM data reveals that assembly of MTs occurs by the addition of subunits to the ends of protofilaments. These protofilaments associate laterally to form a 2D sheet and the longitudinal edges of the sheet meet and anneal to form a tubular structure. GTP bound to β-tubulin provides the energy necessary for dynamic instability. The GTP at the α-monomer is neither hydrolysed nor exchanged due to its location within the dimer, whereas the GTP bound to the β-tubulin can be exchanged in free solution (Lowe et al., 2001) (Nogales et al., 1998a) (Spiegelman et al., 1977). Once incorporated into the MT, GTP-tubulin is hydrolysed to GDP-tubulin. The effect of GTP hydrolysis is to compact the intra-dimer interface within tubulin. Structural studies reveal that GDP protofilaments tend to curve outwards from the MT wall (Krebs et al., 2005) (Müller-Reichert et al., 1998). The minimum-energy conformation of GTP protofilaments, on the other hand, is almost entirely straight (Peng et al., 2014). Lateral interactions between GTP-tubulin subunits in a protofilament provide a constraining force that resists the natural tendency for GDP-tubulin to curve outwards. Typically for rapidly growing MTs, the constraining GTP-tubulin is found at the

tips of each protofilament. Loss of this stabilising GTP-cap renders the MT prone to depolymerisation. Experimentally, MTs polymerised with GMPCCP, a non-hydrolysable form of GTP, show continual growth and no shrinkage (Hyman et al., 1992). The interplay between a stabilising end structure and GTP hydrolysis is necessary for MTs to exhibit the stochastic switching behaviour underlying dynamic instability. Experiments using UV microbeam radiation and micro-needle severing found that newly exposed MT plus-ends rapidly depolymerise, confirming the need for a capping structure to protect the labile GDP-tubulin core from depolymerisation (Walker et al., 1989) (Tran et al., 1997). The discrete nature of the stabilizing GTP-cap was determined by dilution experiments which showed that when the free tubulin is removed, the time taken to switch from growth to shrinkage is the same regardless of the initial growth rate of the MT (Carlier et al., 1984). Furthermore, bulk biochemical assays found little lag between hydrolysis and tubulin incorporation (O'Brien et al., 1987) (Melki et al., 1996) (Vandecandelaere et al., 1999). Studies of MTs composed of GMPCPP and GDP-tubulin suggest it is possible that a single complete layer of GTP-tubulin is necessary and sufficient to stabilise the lattice (Caplow and Shanks, 1996).

Microtubule Growth

Recent experimental and computational studies have subverted the standard description of dynamic instability. MTs assemble due to frequent collisions between GTP-tubulin heterodimers and MT protofilament ends. The rate of MT growth increases linearly as a function of tubulin concentration, driven by more frequent and successful collisions. Based on these observations, researchers first applied Oosawa's equations for polymer growth (Oosawa and Kasai, 1962) to measurements of MT growth rates, taken from groundbreaking microscopy studies *in vitro* (Walker et al., 1988) (Hotani and Horio, 1988). These equations describe the theory of helical aggregation of macromolecules and relate the aggregation rate to the free monomer concentration. This linear or 1D model provided a simple description of MT growth and shrinkage, expressed in terms of measured rate constants. The 1D model treats the MT as a single protofilament end. The model states that MTs grow according to a second-order rate constant, proportional to the tubulin concentration, and shrinkage occurs with a first-order rate constant (Fig. 1.6). This model has held traction with MT researchers for many decades due to its simplicity and ability to fit well to experimental data.

However, early theoretical analyses began to uncover problems in the 1D model. These ideas were first put forward in two important biophysical papers (Odde et al., 1996) (Odde et al., 1995). First, a study was undertaken of the direct observation of individual MTs *in*

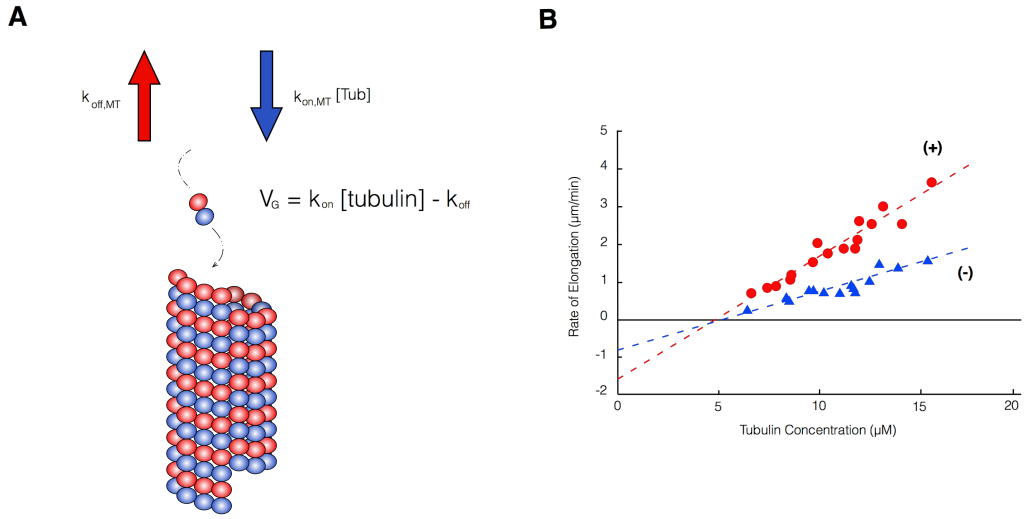


Figure 1.6: 1D model of MT growth. (**A**) During the growth phase, the rate of MT elongation is the result of a bimolecular association rate constant ($k_{on}[T]$) and a unimolecular dissociation rate constant k_{off}). (**B**) The rate of elongation as a function of free tubulin concentration. Mean rate values for the plus- and minus-ends are red circles and blue triangles. The values for the association and dissociation rate constants were determined from the slope of the linear regression lines through the plus- and minus-end data sets. The graph was sourced from (Walker et al., 1988).

vitro and combined with probabilistic analysis of the distribution of their growth times. This study determined that the faster growing plus-ends did not follow first-order catastrophe kinetics. Here the growing MT plus-ends have an effective frequency of catastrophe that depends on how long the MTs have been growing. This frequency is low initially but rises asymptotically to a limiting value, suggesting multiple sub-states within a growth phase. The second study sought to develop a model-independent method of characterising changes in MT growth and shrinkage rates, which is robust to experimental noise. The power spectrum was computed for individual MT life histories. This spectral analysis found that the power spectrum of individual MTs were broad, decaying by a roughly inverse square law, and is reasonably well approximated by a 1D model of growth for low frequencies. At higher frequencies, the 1D model was found to under-predict the level of oscillations observed during experiments, even when experimental error is accounted for. This implies that these oscillations occur even in periods of linear growth.

Recent high-resolution experiments of MT dynamics, observed with optical tweezers and sub-pixel resolution microscopy, reveal length fluctuations during the growth of individual MTs with rapid shortening and re-growth events distinct from large scale catastrophe and shrinkage (Schek et al., 2007) (Gardner et al., 2008). These observations are at odds with the predictions of the 1D model and measurements taken by light microscopy. To account for these kinetic differences at the tip, a 2D model was proposed which treats the MT end as having multiple tubulin binding sites. This model removes the assumption that the MT

tip structure is energetically constant and takes into consideration the structure of the MT tip as a whole. Tubulin binding events are then dependent upon the instantaneous state of the MT tip. Incoming tubulin dimers can form bonds that vary depending on the environment of each protofilament end. This 2D model was found to produce MTs that form tapers, where some protofilaments extend beyond others creating differences in the number of lateral bonds. This behaviour creates a dichotomy in terms of growth and shrinkage. Tubulin associates to the protofilaments more rapidly but shortening events increase because tapered tips are more prone to dissociate and break off due to a lack of lateral contacts. In kinetic terms, the subunit dissociation rate from the tip increases as a function of tubulin concentration and the second-order on-rate constant is considerably higher than in the 1D model (Gardner et al., 2011).

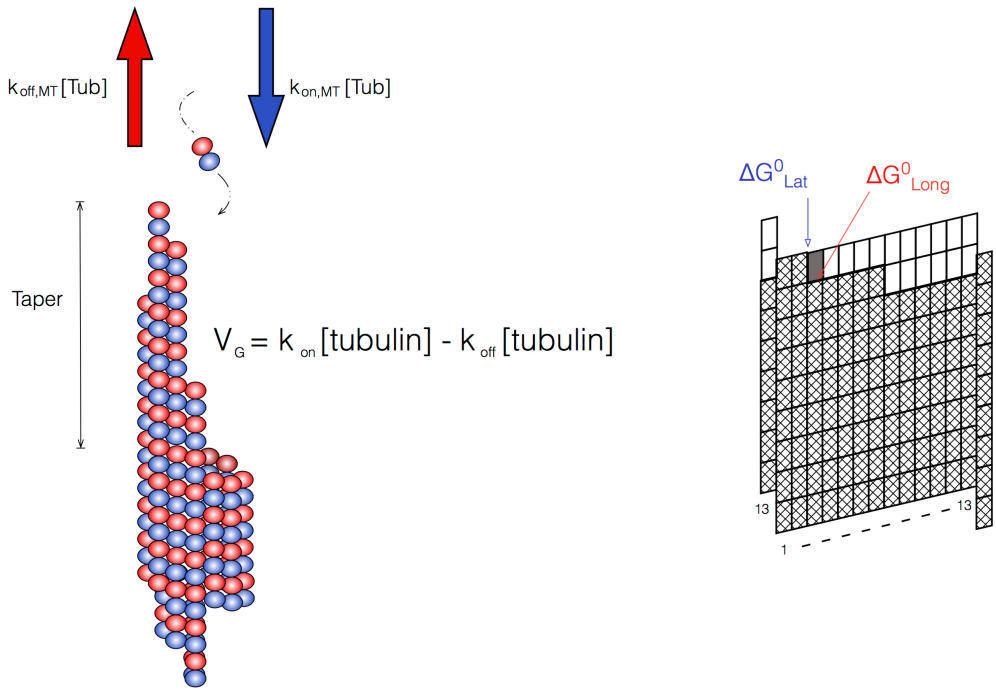


Figure 1.7: 2D model of MT growth. MTs grow and shorten by the addition and loss of $\alpha\beta$-tubulin subunits. The model predicts that tubulin dissociation rates increase at higher tubulin concentrations. Thus, the net assembly rate is driven by a small difference between both its large on-rate and off-rate. Tubulin dissociation varies due to the number of its lateral neighbours.

Recent computer simulations that reproduce MT growth rates and observed MT dynamics, also exhibit nanoscale behaviour at the tip (VanBuren et al., 2005) (Margolin et al., 2012). In these models, simulated MTs have tip structures that resemble closed tubes with dynamic multi-protofilament extensions of varied lengths. These simulations predict that cracks can appear between PFs during growth, and detachment of GTP-tubulin occurs with high frequency. As these MT extensions are more likely to undergo catastrophe than blunt sheets, due to fewer lateral contacts, this makes PFs less resistant to the curling propagated by

GDP-tubulin dimers found deeper in the lattice. This results in a frayed MT tip during shortening that has fewer GTP-tubulin dimers and deeper cracks between protofilament that each terminate in a GTP-tubulin-rich region.

Microtubule Catastrophe

Catastrophe of MTs built from pure tubulin is controlled by GTP hydrolysis and this modulates the size and distribution of the cap. In recent years, the standard description of this structure has been the subject of extensive study and has undergone many revisions. The cap is now considered as a distributed region in which the GTP concentration is higher at the tip and tapers off towards the body of the MT (Gardner et al., 2008) (Chrétien et al., 1995). The exponential-like shape results from GTP hydrolysis occurring as a first-order process on non-terminal subunits. This results in some GTP-tubulin being found deep in the lattice, providing a potential source for MT rescue events. Experimental evidence for this description has come from EB1, a +TIP protein that autonomously tracks the growing plus-end of MTs. This protein binds with a comet-like profile at the tip due to its high affinity for GTP-tubulin and or GDP.Pi tubulin (Bieling et al., 2007) (Maurer et al., 2012) . As EB1 is thought to recognise the nucleotide state of tubulin, based upon evidence of its ability of bind GTP analogues with high affinity, the EB1 signal is considered to be a readout for the GTP-cap. The comet-like shape of this signal reveals that EB1 decays exponentially over hundred of nanometers, suggesting the GTP cap is not small (Maurer et al., 2012). However, the EB1 signal does not reach the very tip of the MT. Super-resolution images of fixed MTs show that ch-TOG, a potent MT polymerase, binds ahead of the EB1 signal (Nakamura et al., 2012). This binding mode indicates that the very tip of the GTP-cap may differ from its laterally-bound neighbours (Fig. 1.8). Tapered structures seen at the very tip of the MT are likely to curve outwards and may offer a structural explanation for a distinct end structure. Shrinkage coincides with EB1s peak signal dropping to below 20%. This observation indicates that loss of the GTP-cap occurs when GTP-density drops below a threshold, indicative of a continuous rather than discrete structure.

It was assumed that GTP hydrolysis weakens lateral bonds between protofilaments and this causes instability in the lattice. Without GTP at the tip to protect GDP-tubulin below it, MTs switch to a disassembling state. GTP-tubulin was believed to form enhanced lateral contacts based on the observed structure of GTPϒS MTs (Maurer et al., 2011). GTPϒS is another slowly hydrolysable GTP-analog and GTPϒS MTs are thought to more closely mimic GTP caps. In contrast, no differences between lateral contacts were found between the structures of GMPCCP and GDP MTs. The primary difference between the structures

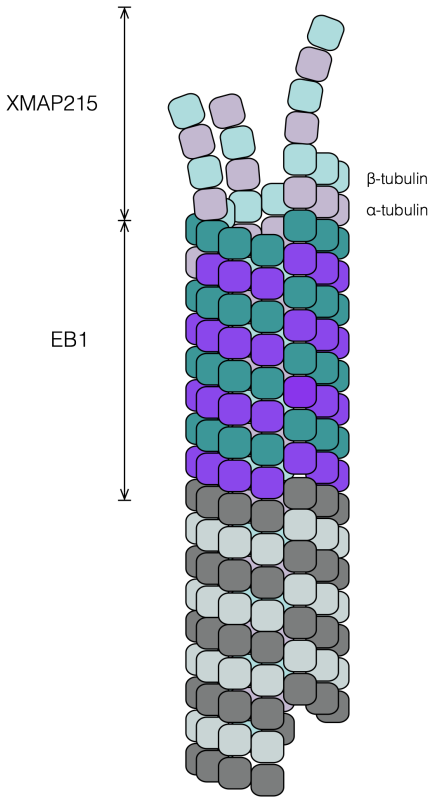


Figure 1.8: Composition of the MT tip structure. The distal end of the MT is thought to consist of curved protofilaments composed of GTP-tubulin with few lateral neighbours. This region is the preferred binding site of XMAP215-like proteins. A short distance away from the tip, the EB1 binding site (purple) extends over hundreds of nanometers. It is also composed of GTP-tubulin but in a straighter conformation that is locked into the MT lattice. The main body of the MT is composed of GDP-tubulin (grey).

is compaction of α-tubulin in GDP-tubulin subunits (Fig. 1.9), which disrupts longitudinal bonds between dimers (Alushin et al., 2014). These new cryoEM structures raise doubts about the role of lateral contacts during GTP hydrolysis and suggest that longitudinal bonds may be making more important contributions to stabilisation of the GTP-cap. This evidence points towards a GTP-cap whose detailed structure varies dynamically due to fluctuating growth and GTP hydrolysis in neighbouring protofilaments.

Simple GTP-cap models inadequately explain the frequency of catastrophe observed by experimental measurements. The canonical view of MT catastrophe describes loss of the GTP-cap as a single step random event with the probability of shrinkage being equal, regardless of how long the MT has been growing. In such a scheme, the distribution of MT lifetimes and lengths follows a decaying exponential distribution, a pattern not seen in direct experimental measurements. More curiously, the EB1 signal has been shown to increase as MTs grow faster, indicative of a larger GTP-cap (Maurer et al., 2012). Larger

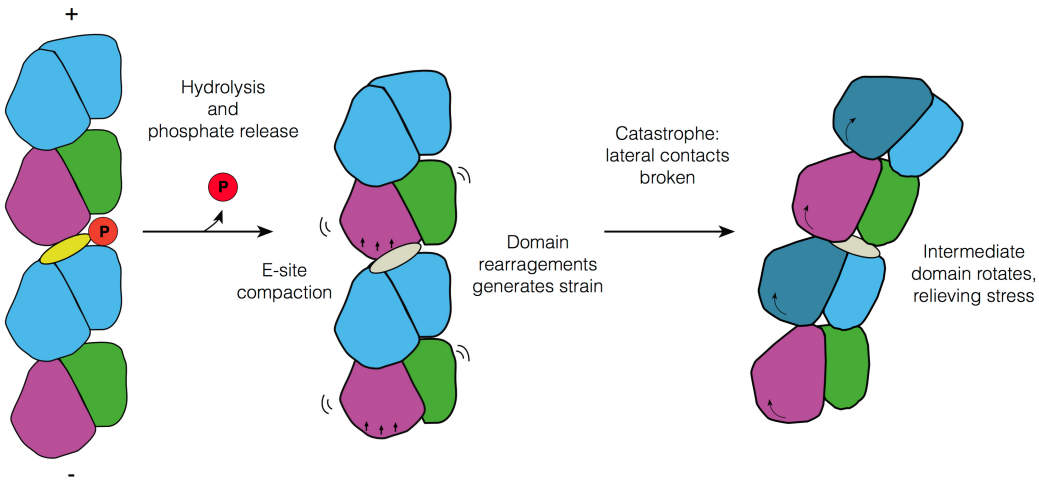


Figure 1.9: Compaction and conformational change of α-tubulin. Moving from left to right, hydrolysis and phosphate release leads to compaction of the E-site by movement of α-tubulins intermediate domain. Figure sourced from (Alushin et al., 2014).

GTP-cap sizes should endow a MT with more resistance to growth fluctuations and therefore lower its frequency of catastrophe. However, high-resolution experiments of growing MTs found that the catastrophe frequency remains relatively constant and is insensitive to the MT growth rate for young MTs (Gardner et al., 2013). Yet the distribution of MT lengths does still appear to decay exponentially at longer time scales. This implies that younger MTs have a lower probability of undergoing catastrophe than older MTs, and their catastrophe frequency increases with age. Young MTs are protected due to a lack of accumulated features that promote depolymerisation. These phenomena have recently been encapsulated into a model that describes MT catastrophe as a multi-step process (Bowne-Anderson et al., 2013). In this model, MT catastrophe occurs after three or more destabilising events on individual protofilaments, where terminal subunits exist in a GDP state. But what molecular mechanism explains age-dependent MT catastrophe? It has been proposed that evolution of tapered tip structures seen during MT growth could be the natural explanation for catastrophe (Coombes et al., 2013). Young MTs would begin to grow with a blunt tip, which does not predispose a MT to catastrophe, and thus young MTs will have a lower catastrophe frequency. However, as the MT grows over time, the tip structure becomes more tapered, which ultimately acts to destabilise the MT and increase the likelihood of a catastrophe event. The tapered tip could lead to a gradient in the tubulin subunit off-rate and this increases the probability of GTP-cap loss. GTP-cap loss from a critical number of unstable protofilaments could thus predispose the MT with a highly tapered structure to catastrophe. Here, the role of the GTP-cap and tip structure can be related as follows. At moderate tubulin concentrations, the increasingly tapered and unstable tip structure can compete with the countervailing effect of a large GTP-cap. However, the tip structure effect may eventually

saturate at high tubulin concentrations, and the stabilising effect of the GTP cap would then dominate.

Summary

The near equality of the on- and off-rates of the 2D model, implies an infrequent suppression of the off-rate should lead to a dramatic shift in the assembly for single MTs. Proteins that bind to the MT ends to regulate the parameters of dynamic instability could do so through only modest changes in the tubulin subunit exchange kinetics.

1.3 XMAP215/Dis1 Family

MT dynamics reconstituted *in vitro* exhibit a range of dynamic properties that are distinct from those of MTs in cells. MTs *in vivo* grow at least three times faster than those assembled *in vitro* from purified tubulin and show more dynamic behaviour (Cassimeris et al., 1988) (Charrasse et al., 1998) (Pryer et al., 1992) (Vasquez et al., 1994). The functional differences between MTs assembled in cells and those *in vitro* are due to the actions of numerous MT-associated proteins (MAPs) that use MTs as an important binding platform. Members of the highly evolutionarily-conserved XMAP215/Dis1 family help to regulate MT dynamics by promoting fast MT growth. These proteins use multiple tumour over-expressed gene (TOG) domains (Charrasse et al., 1998) to selectively bind both tubulin dimers and growing MT ends, so as to promote tubulin exchange. This protein family has characterised members in all of the major eukaryotes: fungi, plants and animals (Gard et al., 2004). The first member of this protein family to be characterised, XMAP215 from *Xenopus laevis*, was discovered through the biochemical fractionation of the Xenopus egg extracts (Vasquez et al., 1994). In vivo, XMAP215 family members localise to kinetochores and MT organising centres. Both cells immunodepleted of XMAP215 and cells carrying loss-of-function mutants exhibit short and slowly-growing interphase MTs and aberrant spindles during mitosis. At the kinetochores, family members are required for the regulation of kinetochore-MT attachment (Garcia et al., 2002). In general, XMAP215/Dis1 members are essential for the correct functioning of MT dynamics during cell division.

1.3.1 Structure

All XMAP215/Dis1 protein family members share a common structural organisation. All contain N-terminal TOG domains that form conserved linear arrays necessary for MT poly-

merisation (Fig. 1.10), and a C-terminal MT binding domain. The MT binding domain of yeasts differ from those of higher eukaryotes in that they contains a basic SK-rich region but no other sequence similarity, whereas the mammalian proteins have a common C-terminal motif. Lower eukaryotic family members contain two TOG domains and dimerise via a C-terminal coiled-coil region, yielding a tetrameric complex. The higher eukaryotic members are monomeric and contain trimeric or pentameric TOG arrays. All TOG domains are connected by flexible linkers of basic residues that promote lattice association.

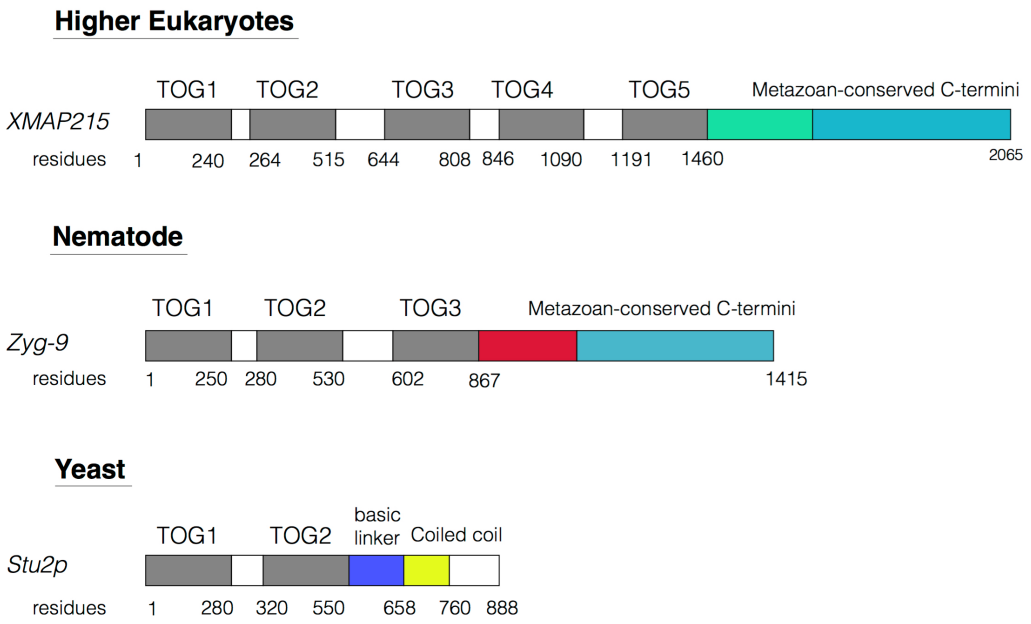


Figure 1.10: Domain organisation of XMAP215/Dis1 proteins. The higher eukaryotic family members are monomeric, have five TOG domains, and contains a unique C-termini. The nematode orthologs, with three TOG domains, contains a variable regions connected to a conserved C-terminal region. Yeast XMAP215 member including Stu2p have a conserved coiled-coil domains (yellow) that mediates homodimerisation. Figure adapted from (Al-Bassam et al., 2007)

The structure of a TOG-domain consists of a series of six (A-F) HEAT-repeats that form a 60 Å long paddle-like structure (Fig. 1.11). The HEAT-repeat motif was initially found in a diverse family of proteins, including the four from which it derives its name: Huntingtin, Elongation factor 3, the PR65/A subunit of protein phosphatase 2A, and the kinase TOR. HEAT-repeats (HRs) are composed of tandemly-repeated sequences of two anti-parallel helices, HA and HB, connected by a turn, and they occur in a wide range of proteins. These repeats provide a scaffold for facilitating protein-protein interactions. HRs share a common phylogenetic origin in the β-importins. HRs are variable in length and in amino acid structure, which renders their identification by sequence comparison difficult (Kippert and Gerloff, 2009). Hydrophobic residues within the helices of each repeat are conserved and form a hydrophobic core that provides a tight helical packing (Kappel et al., 2010). The highest degree of structural conservation maps to the face, delineated by the intra-HR loops,

primarily the loop regions between repeats A-C. Furthermore, conserved solvent-exposed residues are found within, or adjacent to, the turns connecting helix A to helix B of each repeat.

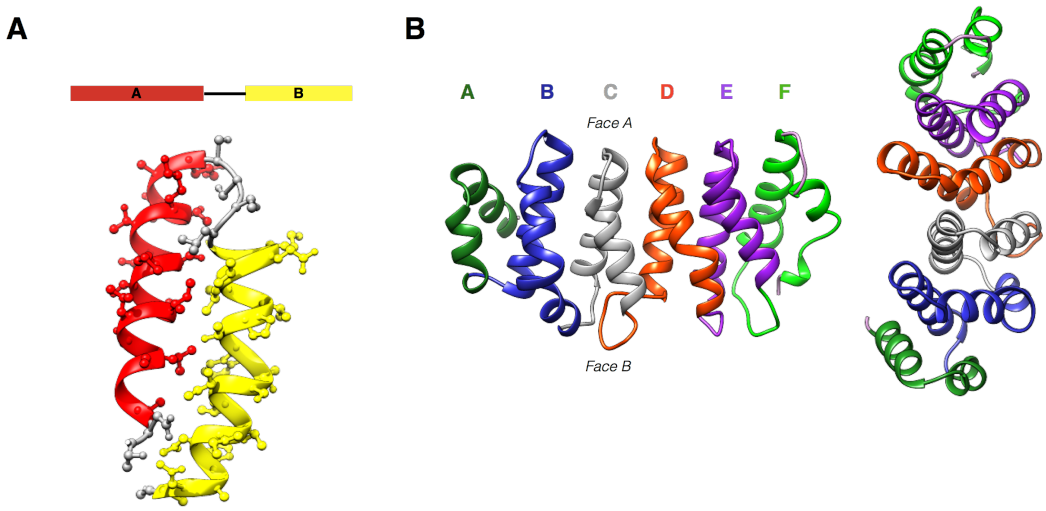


Figure 1.11: The composition of TOG-domains. (**A**) Typical structure of a HEAT-repeat - two anti-parallel α-helices joined by a turn. (**B**) TOG-domains consist of six HRs A-F. The loops of face A form the tubulin interaction surface.

1.3.2 Tubulin Binding

XMAP215 family members are able to act as MT polymerases due to the ability of their TOG domains to bind tubulin dimers and release them once incorporated into the MT lattice. This binding is mediated by the intra-HEAT-repeat loops that are positioned along the face of the domain that associates with a single tubulin heterodimer. Recent crystal structures of Stu2 TOG1 and TOG2 in complex with αβ-tubulin confirmed that the intra-HR loops are the main components that form the tubulin interaction interface. In this structure, the first four HRs engage β-tubulin, whereas the final two HRs engage α-tubulin (Fig. 1.12). TOG domains bind to GDP- or GTP-tubulin with equal affinity, interacting with a curved β-tubulin conformation (Ayaz et al., 2012) (Ayaz et al., 2012). The asymmetry of TOG-tubulin binding excludes the binding of a second TOG-domain. When conserved residues in these loop regions are mutated, TOG-tubulin interactions are abrogated and MT polymerase activity is severely compromised. Mutating the tubulin-binding activity of TOG-domains within the pentameric TOG array of XMAP215 has been shown to have differential effects on MT polymerase activity (Widlund et al., 2011). Mutations in TOG1 or TOG2 compromise XMAP215-mediated polymerase activity more dramatically than homologous mutations in TOG3 or TOG4. Mutating the tubulin-binding loops of TOG5 has no significant effect on the rate of MT polymerisation.

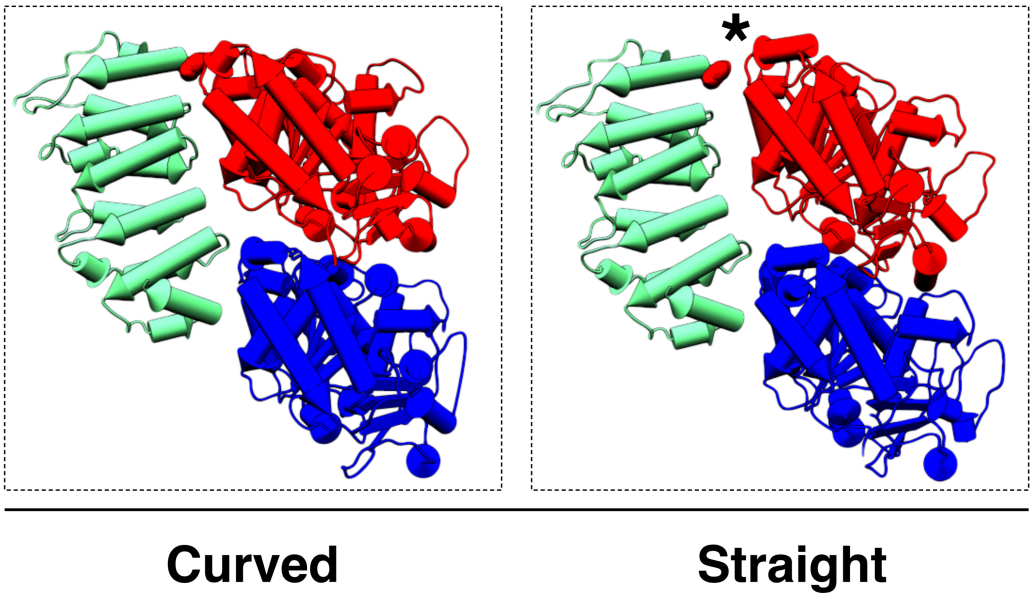


Figure 1.12: TOG:αβ-tubulin complex (PDB ID: 4U3J). The structure of a TOG domain docked with curved vs straight αβ-tubulin (PDB ID: 1JFF), illustrates how TOG-containing epitopes (red spheres) make contact with α- and β-tubulin in each of the two conformations. The TOG domain (green) for curved tubulin makes contact with both α- (blue) and β-tubulin (red) monomers. TOG domains bound to straight tubulin fail to make important residue contacts on the β-mononer

1.3.3 Catalysis

Studies investigating the mechanism of XMAP215 polymerase activity show that XMAP215 family members catalyse the incorporation of GTP-tubulin into a MT through multiple rounds of addition (Brouhard et al., 2008). XMAP215 binds a MT plus end tip and incorporates a tubulin dimer into the lattice before diffusing to the next reaction site, tracking the growing tip. This processive function is dependent on the coordinated action of multiple TOG domains together with a MT binding domain for MT association and polymerase activity (Fig. 1.13). Typically, tubulin dimers associate to the MT ends in a diffusion-limited reaction and form a short-lived complex where the dimer often fails to be incorporated. XMAP215 accelerates growth by stabilising this intermediate collision complex so that there is a higher chance of incorporation into the lattice. XMAP215 thus acts as a classical catalyst and is not consumed in the reaction. As a catalyst, XMAP215 can also accelerate MT depolymerisation at low tubulin concentrations. Kinetically, XMAP215 proteins increase both the net association and dissociation rate with no change in critical concentration. Both monomeric and dimeric family members show equivalent catalytic activity on MTs. However, how each TOG domain position in the array contributes to its differential activity is currently being addressed in structural studies.

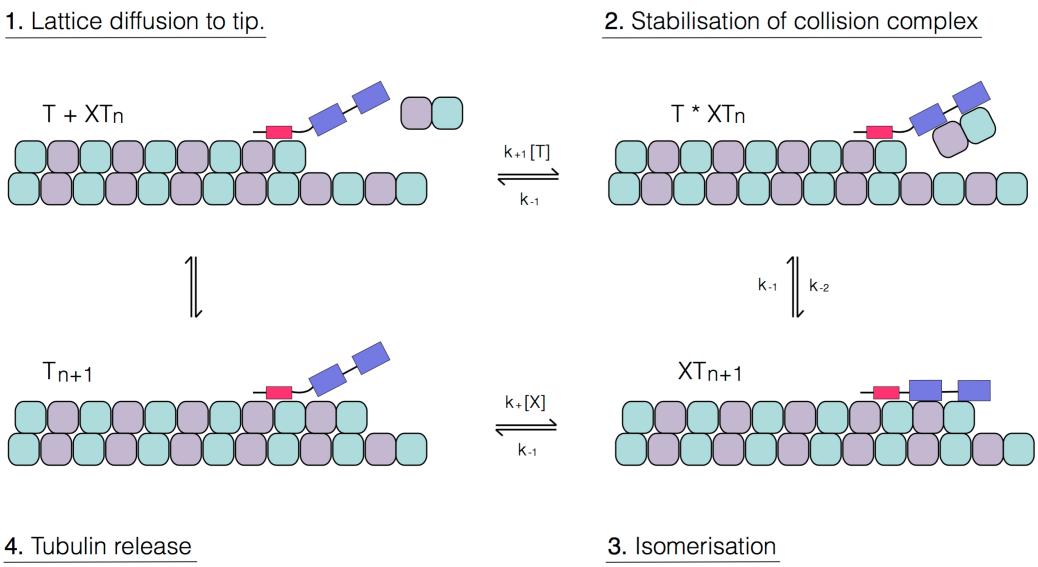


Figure 1.13: The catalytic cycle of TOG domain polymerisation. (**1**) Diffusion to the MT tip is facilitated by the polymerases MT-binding domain. (**2**) TOG1-2 stabilises the incoming dimer by forming a tethered complex. (3) TOG1-2 domains mediate the collision complex between the dimer and the MT tip. (4) The TOG domain releases the dimer on incorporation into the MT lattice. The polymerase is free to diffuse to the next available binding site at the tip. Figure adapted from (Widlund et al., 2011)

1.3.4 Differential TOG-activity

To understand the TOG array mechanism it is important to characterise the structural and functional properties of each domain within the array. Lower eukaryotes, like Stu2 from *S. cerevisiae* have only two TOG domains but drive polymerisation as a homodimer. Studies to date suggest that TOG1 and TOG2 bind and tether tubulin heterodimers proximal to the MT plus ends to incorporate tubulin into the lattice (Ayaz et al., 2014). Plus-end localisation of the polymerase may occur via a TOG domain binding to curved tubulin at the tip, with further TOG-domain(s) binding free β-tubulin dimer and physically tethering them to the MT end. Tethering then results in a higher frequency of association to the MT tip. In contrast, members with pentameric TOG arrays act as monomers and, although their TOG1-3 domains bear a structural similarity to fungal TOG1 and TOG2, their spatial arrangement on the MT is likely to differ (Fig. 1.14). However, it is still likely that a pentameric array interacts with a MT with the same polarity observed in the Stu2 TOG-tubulin crystal structures. The TOGs N- and C-termini will be directed towards the MT plus end and minus end, respectively. The full array may therefore bind to sequential tubulin dimers along a single protofilament, arranged from the plus to the minus end.

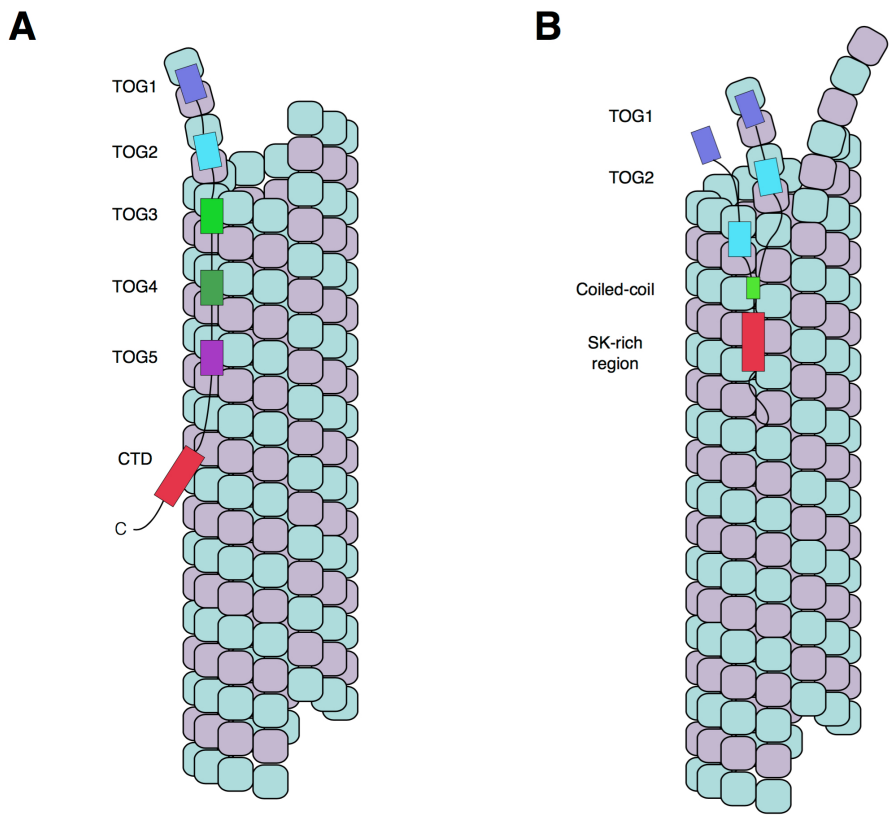


Figure 1.14: TOG domain architecture and possible binding modes of (**A**) higher and (**B**) lower-eukaryotic species. Fungal polymerases with only TOG1-2 domains bind and promote tubulin incorporation at the MT tip, consistent with their near equivalent GTP-tubulin binding affinity. Metazoan polymerases consist of the same high affinity TOG1-2 domains but also contain further TOG3-5 domains, which transiently interact with tubulin and have limited polymerase activity. The binding and mode of action for these polymerases remains elusive.

The polarised arrangement of the pentameric TOG array and its orientational bias suggests an individual function for each domain within the array. Some domains may selectively bind and deliver free tubulin or recognise different tubulin states along the MT lattice. This is consistent with observations of TOG-domains binding to tubulin heterodimers with differential affinity. For instance, studies show that Msps, XMAP215 and Stu2 TOG1-2 constructs co-elute with tubulin dimers over gel filtration (Fox et al., 2014) (Ayaz et al., 2012) (Ayaz et al., 2014) (Widlund et al., 2011) (Brouhard et al., 2008). In contrast, a Msps TOG3-4 construct did not show robust tubulin binding over gel filtration, indicating that this construct forms more transient interactions with tubulin (Fox et al., 2014). Additionally, a full length XMAP215 construct containing deleterious mutations in all TOG-domains except TOG1 and 2 retained a significant proportion of its tubulin binding activity that was not evident when all five TOG-domains were mutated (Widlund et al., 2011). This would suggest the TOGs 1 & 2 within the array are the most conserved for tubulin binding. Further evidence suggests that remaining TOG-domains may interact with a straight conformation of tubulin found in the

lattice. The Msps TOG3-4 construct, when over-expressed in cell culture, localises along the length of MTs. Similarly, full-length XMAP215 showed some MT lattice-binding in vitro that was not observed with the XMAP215 TOG1-2 construct. Collectively, these observations suggest that TOG-domains have differential tubulin binding affinities and differential capacities to recognise free versus lattice-incorporated tubulin.

1.3.5 TOG domain Structures

To date, TOG domain structure have been solved for Stu2 TOG1 & TOG2, Zyg9 TOG3, Msps TOG2 & TOG3 & TOG4 together with a high-resolution structure of the TOG-tubulin complex (Al-Bassam et al., 2007) (Fox et al., 2014) (Slep and Vale, 2007) (Ayaz et al., 2012) (Ayaz et al., 2014) (Howard et al., 2015). TOG domain crystals have resolved a number of structurally distinct features that contribute to each TOG domains position-dependent activity.

TOG domains 1-2 are similar within and across species and, based on crystal structures, bind to a curved tubulin conformation, predicted to be close to the minimum-energy structural state of free tubulin. The Stu2 and Msps TOG2 are formed by six HEAT-like repeats A-F. Only HEAT-repeats C & D of Msps TOG2 and C & F of Stu2 form canonical HEAT-repeats. The HEAT-repeat E, particularly of Stu2, is characterised by a segmented N-terminal α-helix.

Zyg9 TOG3 from *C. elegans* has seven HEAT-repeats but its core structure contains HRA-HRF (Al-Bassam et al., 2007). The additional N-terminal HR is thought to extend the domains lateral reach, and may facilitate binding across between protofilaments during polymerisation. Its unique, extra N-terminal HR0 is positioned orthogonal to HRs A-F and lies outside of the principal stack. Sequence comparisons reveal that the HR0 sequence is present only in XMAP215 TOG5, as well the fifth TOG domains of MOR1, Msps and other homologs from higher eukaryotes. Interestingly, HR6 packs onto HR5 with a 45° right-hand twist. These structural features make zyg9 TOG3 amongst the most structurally diverse of TOG domains. Sequence comparisons show that Zyg9 TOG3 is more closely related to TOG5 than to pentameric TOG3 domains from other species.

Msps TOG3 is another canonical TOG domain composed of six HRs, but with a unique stabilising C-terminal extension (Howard et al., 2015). Msps TOG3 is highly conserved across pentameric XMAP215 family members and its predicted intra-HEAT repeat loops are conserved with Stu2 TOG1 and TOG2, suggesting that Msps TOG3 also binds to curved tubulin. Msps TOG3 HRs are grouped into two characteristic triads that have slight twists: HR A-C and D-F. Specifically, the HR A-C triad has a right-hand twist, whereas the HR

D-F triad has a right-hand twist between HR D-R and a left-handed twist between HR E-F. Msps TOG3 has conserved solvent-accessible residues including a conserved tryptophan in the intra-HEAT loop A and a lysine in intra-HEAT loop C that are positioned to interact with β-tubulin. Msps TOG3s unique C-terminal extension binds alongside HR C-F and formed multiple conserved contacts. The tail promotes domain stability. Msps TOG3 has a truncated helix which is replaced by an extended intra-HR B loop. This extended loop consists of highly conserved residues that support its position and maintain optimal position to interact with tubulin.

Msps TOG4 is the most structurally distinct TOG domain . Msps TOG4 and ch-TOG TOG4 structurally align well via their HR A-F, sharing conserved loop and flanking residues, with minor differences in their terminal regions of HR D (Fox et al., 2014). The first HR triad in MSPS TOG4 is conserved and parallels the arrangement used by Stu2 TOGs in their interaction with tubulin. However, MSPS TOG4 HR D-F triad is dramatically re-orientated ~45° relative to Stu2 TOG2 HR D-F triad, effectively positioning HR F 19 Å away from its stiff TOG2 counterpart. This shift in the triad creates a distinct binding mode that is different from the geometry established in the Stu2 TOG1/2:αβ-tubulin crystal structures. The 18 Å shift in the relative positioning of HR F suggests this region of TOG4 may bind to a distinct site on α-tubulin. In short, the TOG domain structures of TOG1, TOG2 and TOG3 have similar architectures that differ significantly from TOG4, whilst TOG3 may engage β-tubulin in a mode that parallels that of TOG1 and TOG2. The structural divergence of TOG4 contrast with the lower RMSD values obtain when comparing Msps TOG3 to Stu2 TOG1-TOG2 and Msps TOG2.

In summary, structural evidence alongside biochemical assays would indicate that the TOG domain array is structurally polarised, with position-dependent features conserved across species.

1.3.6 CLASP

The CLASP family of proteins also contain TOG-domains. Although this thesis is concerned with only growth promoting TOG-domains of the XMAP215 family, CLASP is included here for reasons of completeness. CLASP TOG-domains share high-levels of sequence homology to XMAP215 TOG-domains but function to regulate the frequencies of rescue and catastrophe. In vitro reconstitution experiments show that CLASP promotes rescue, and loss of function mutants affect MT stability and cause the onset of spindle defects (Leano et al., 2013); interestingly, with no change in MT growth rate. Until recently, the function of CLASP TOG domains was perplexing since sequence comparisons to XMAP215 TOGs

found no significant sequence divergences to explain the difference in function between the protein classes. The recently published structure of a CLASP family TOG domain provided a possible explanation for the structural origin of this domains function. The binding surface of XMAP215 family TOGs interact with tubulin via a relatively flat interaction. CLASP, on the other hand, has a structure where some of its HRs are staggered which breaks the geometrical match with curved β-tubulin (Fig. 1.15). This suggest that CLASP TOG domains might bind to a curved conformation of β-tubulin that is reminiscent of curved PFs seen in depolymerising MTs.

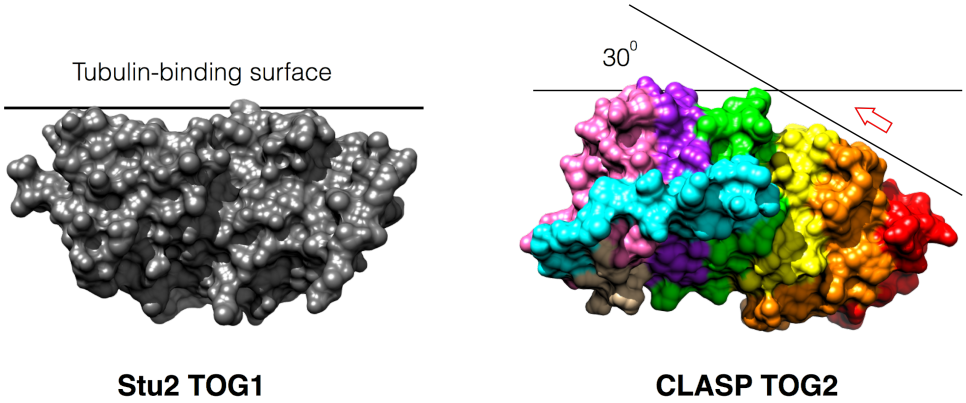


Figure 1.15: Surface representation of Stu2 TOG1 (PDB ID: 1JFF) (Ayaz et al., 2012) and CLASP TOG2 (PDB ID: 4K92) (Leano et al., 2013), showing the flat Stu2 TOG1 binding surface versus the bent CLASP TOG2 surface. The intra-HRs A-C and D-F form two surfaces and is angled 30° relative to other. This suggests that CLASP TOGs engage tubulin in a binding mode that is distinct from XMAP215 family members.

1.4 Theory

This section presents a brief introduction of the theoretical background for the main biophysical techniques applied in this thesis: Continuum Electrostatics, Normal Mode Analysis and Brownian Dynamics.

1.4.1 Continuum Electrostatics

In a continuum electrostatic model, the interior of a protein is represented as a homogeneous low dielectric medium, surrounded by a continuum of high dielectric solvent. This simplified treatment of a macromolecule offers computational efficiency at the cost of reduced accuracy. The electrostatic potential of a protein using this description can be best described by solving the Poisson-Boltzmann equation.

Electrostatic Potential

The electric potential is the work done to carry a test charge against an electric field (E) of the point charge (Q) along a path (s) connecting a point with zero potential to r.

$$\phi(r) = \frac{U(r)}{q_e} = \frac{1}{4\pi\epsilon_0}\frac{Q}{r} = \int_{\infty}^{r} E \cdot ds \tag{1.1}$$

As the electric field is associated with a set of stationary charges, the curl of the electrostatic field is zero, and is determined by the gradient of the potential:

$$E = -\nabla\phi \tag{1.2}$$

Poisson Equation

By the application of Gauss's law, the potential satisfies Poisson's equation which describes the potential energy field caused by a given charge density distribution:

$$\nabla^2\phi(r) = -\frac{\rho(r)}{\epsilon_0} \tag{1.3}$$

For a condensed medium in which dielectric effects can be represented by a dielectric constant then Eq. 1.3 becomes:

$$\nabla^2\phi(r) = -\frac{\rho(r)}{\epsilon(r)\epsilon_0} \tag{1.4}$$

$$\nabla[\epsilon(r)\nabla\phi(r)] = -\frac{\rho(r)}{\epsilon_0} \tag{1.5}$$

Poisson's equation captures a simplistic view of the electrostatic potential surrounding a protein by reducing the complexity of the protein-solvent system.

Poisson-Boltzmann Equation

In a real system, proteins interact with a range of molecular species. The cytoplasm of a cell is enriched with dissolved electrolytes. These mobile charge carriers interact with the protein surface, generating their own electrostatic field. This same field reciprocally influences the electrostatic field of the protein itself. The charge distribution of counterions can be modeled in a mean field approach by the Boltzmann distribution, giving rise to a mobile charge distribution:

$$\rho_e(r) = q_i \cdot \rho_{bulk} \cdot e^{\frac{\rho(r)\cdot q_i}{kT}} \tag{1.6}$$

This expansion of the Poisson equation to account for the mobile charge distribution leads to the Poisson-Boltzmann equation:

$$-\epsilon_0\nabla\cdot[\epsilon(r)\nabla\phi(r)] = \rho(r) + \sum_{i=1}^{n} q_i\rho_i^{bulk}e^{\frac{-\phi(r)\cdot q_i}{kT}} \tag{1.7}$$

This equation simplifies for binary electrolytes at a 1:1 concentration by reducing a two-term exponential series to a hyperbolic sine:

$$-\epsilon_0\nabla\cdot[\epsilon(r)\nabla\phi(r)] - \rho_{bulk}sinh\Big(\frac{\phi(r)}{kT}\Big) = \rho(r) \tag{1.8}$$

The linear form of the PBE is often used in protein applications to reduce computational overhead by assuming sinh(ϕ(r)) ~ ϕ(r) which results in:

$$-\epsilon_0 \nabla \cdot [\epsilon(r) \nabla \phi(r)] - \rho_{bulk} \left(\frac{\phi(r)}{kT} \right) = \rho(r) \tag{1.9}$$

The linear approximation can lead to significant errors for proteins with high charge densities.

1.4.2 Solving Poisson-Boltzmann Equation

To solve the PBE equation for various conditions found within a real protein environment, such as changes in atomic charge and dielectric boundaries, one must resort to using numerical methods. Exact, analytical solutions to the PBE exist only for simple geometries - sphere, plane or cylinder. Adapting these closed form solutions to more complicated geometrical shapes like those of large macromolecules has proved difficult. For arbitrarily shaped proteins, the PB equation can be solved by standard approaches such as finite-difference and finite-element methods (Baker et al., 2001). Each of these methods faces challenges; in particular, the difficulty of mapping an irregular molecule surface to a volumetric mesh, representing the source distribution as a set of discrete point charges and convergence issues associated with dielectric discontinuities. More recently, an alternative approach to solving the Poisson-Boltzmann equation directly is to use a boundary-integral formulation and the boundary-element method to determine the charge distribution on the molecule surface. The BEM formulation is often more efficient in terms of computational resources, as discretisation is performed over the surface of a model, rather than the volume occupied by the molecule and solvent (Yokota et al., 2011). Furthermore, later studies have extended the BEM method to include solvent-filled cavities and stern layers which are naturally defined by volumetric methods (Cooper et al., 2013).

1.4.3 Normal Mode Analysis

Normal mode analysis (NMA) is an important biophysical technique for studying large-scale protein flexibility. A detailed knowledge of the dynamics of macromolecules enables a more complete understanding of biological processes. The biological functions of macromolecules originate with thermal fluctuations and energy-dependent conformational rearrangements. These protein dynamics cover a broad timescale. Thermal fluctuations of bond lengths and torsion angles occur on relatively small scales but very rapidly. Large-scale rearrangements, on the other hand, occur over much longer time scales. However,

many large scale conformational changes occur on timescales still not accessible to most time-dependent theoretical methods, even on the best computational hardware. Therefore, one must resort to the use of a time-independent approach. This approach is NMA.

NMA is a harmonic approximation of the vibrational motion of an oscillating system in the immediate vicinity of it equilibrium. The motions studied are primarily the slowest motions in a macromolecular system and are calculated by assuming a simple harmonic form of the potential that is approximated by a quadratic function.

Calculating Normal Modes

A protein macromolecule can be modelled as a network of simple harmonic oscillators connecting point masses, vibrating around their equilibrium position. This system is comprised of N atoms under a given force field, which describes the interaction potential between atoms, and a set of Cartesian coordinates located at positions r = ($r_1, \ldots, r_n$).

The total energy of the system over time is defined by the Hamiltonian, which is the sum of the kinetic K(r) and the potential U(r) energy:

$$\mathscr{H} \cong K(r) + U(r) \tag{1.10}$$

Taylor expansion of the potential energy function around an equilibrium conformation (r^0) gives:

$$U(r) \cong U(r^0) + \sum_{i=1}^{3N} \frac{\partial U}{\partial r_i}\bigg|_{r=r^0} (r_i - r^0) + \frac{1}{2!} \sum_{i}^{3N} \sum_{j}^{3N} \frac{\partial U}{\partial r_i \partial r_j}\bigg|_{r=r^0} (r_i - r_i^0)(r_j - r_j^0) + \ldots \tag{1.11}$$

After defining the potential energy of the reference structure as $U(r^0)$ = 0, the potential energy function becomes:

$$U(r) \cong \frac{1}{2} \sum_{i}^{3N} \sum_{j}^{3N} \frac{\partial^2 U}{\partial r_i \partial r_j}\bigg|_{r=r^0} (r_i - r_i^0)(r_j - r_j^0) \tag{1.12}$$

The kinetic energy function is defined by:

$$K(r) = \frac{1}{2}\sum_{i}^{3N} m_i \Big(\frac{dr_i}{dt}\Big)^2 \tag{1.13}$$

The Hamiltonian re-written in mass-weighted coordinates, $X_i = m_i^{1/2}(r_i - r_i^0)$:

$$\mathscr{H} \cong \frac{1}{2}\sum_{i}^{3N} \dot{X}_i^2 + \frac{1}{2}\sum_{i}^{3N}\sum_{j}^{3N} \frac{\partial^2 U}{\partial X_i \partial X_j}\bigg|_{X=X^0} X_i X_j \tag{1.14}$$

The coupled oscillatory motions can be reformulated by choosing appropriate normal mode coordinates (q). The mass-weighted Cartesian and normal mode coordinates are linearly related by:

$$X = Aq \tag{1.15}$$

Using these new coordinates the Hamiltonian of Eq. 1.14 can be further simplified.

$$K(q) = \frac{1}{2}\sum_{i}^{3N} \dot{X}_i^2 = \frac{1}{2}\dot{X}^T\dot{X} = \frac{1}{2}\dot{q}^T A^T A\dot{q} = \frac{1}{2}\dot{q}^T\dot{q} \tag{1.16}$$

To simplify the potential energy, the Hessian matrix $H(X)_{ij} = \partial^2 U/\partial X_i \partial X_j$ can be transformed into the normal mode coordinates using A:

$$H(q) = A^T H(X) A = \Lambda \tag{1.17}$$

The potential energy in normal mode basis becomes:

$$U(q) = \frac{1}{2}\sum_{k}^{3N}\sum_{l}^{3N} q_k A^T H(X) A q_l = \frac{1}{2} q^T \Lambda q \tag{1.18}$$

Finally the Hamiltonian uncouples as a set of independent harmonic oscillators:

$$\mathscr{H} \cong \frac{1}{2}\sum_{k}^{3N} \dot{q}_k^2 + \frac{1}{2}\sum_{k}^{3N} \lambda_k q_k^2 \tag{1.19}$$

By solving the standard eigenvalue problem, that is, the result from diagonalising the Hessian matrix H(X):

$$HA = \Lambda A \tag{1.20}$$

The resulting pairs of eigenvalues and eigenvectors are known as the normal modes.

1.4.4 Brownian Dynamics

The computation of a protein's diffusional association rate is important, as these events are ubiquitous in biological systems. The Brownian Dynamics (BD) simulation methodology may be used to simulate diffusion-limited, protein-protein encounters and their association rates.

Ermak-McCammon Algorithm

BD is a molecular simulation technique that describes the motion of molecules in which explicit solvent is replaced by a stochastic force. Mathematically, over long timescales, this diffusional motion is best described by the Langevin equation. The most common and base solution to the integration of this equation for BD is the Ermak and McCammon algorithm (Ermak and McCammon, 1978).

In general, the theory of Brownian motion describes the movement of colloidal particles whose mass and size are much larger than those of the host medium. This results in a distribution of momenta that relaxes to an equilibrium distribution more rapidly than the distribution of particle positions. Consequently, for applications of the theory, such as in BD simulations, one is only concerned with time intervals that are larger than the momentum relaxation time i.e. the diffusive regime. These assumptions allow the definition of a diffu-

sive displacement equation where near identical equations govern both translational and rotational Brownian motion as follows:

$$r = r_0 + \frac{D_t \Delta t}{k_b T} \cdot F + S, \tag{1.21}$$

where F is the intermolecular force, R is the random vector satisfying $\langle R \rangle = 0$ and $\langle R^2 \rangle = 6D_t\Delta t$.

$$w = w_0 + \frac{D_r \Delta t}{k_b T} \cdot T + W, \tag{1.22}$$

where T is the torque acting on the particle, D_r is the rotational diffusion constant, W is a random angle satisfying $\langle W \rangle = 0$ and $\langle W^2 \rangle = 6D_r\Delta t$.

Diffusion coefficients for translation D_t and rotation D_r are given by the Stokes-Einstein relations:

$$D_t = \frac{kT}{6\pi\eta r} \tag{1.23}$$

$$D_r = \frac{kT}{8\pi\eta r^3} \tag{1.24}$$

NAM Algorithm

Although the previous method is sufficiently generalised to model Brownian motion under hydrodynamics and external force, one is predominately interested in simulating a biomolecular rate constant, in order to make comparisons to experimental studies. The Northrup, Allison and McCammon algorithm (NAM) resolves this issue by deriving an expression which relates the collision probability, obtained from BD, to a rate constant (Northrup et al., 1984).

Conceptually, the algorithm divides the configuration space of a particle pair into two distinct regions (Fig. 1.16). A sphere of radius b divides the space r into an outer region (r>b) and an

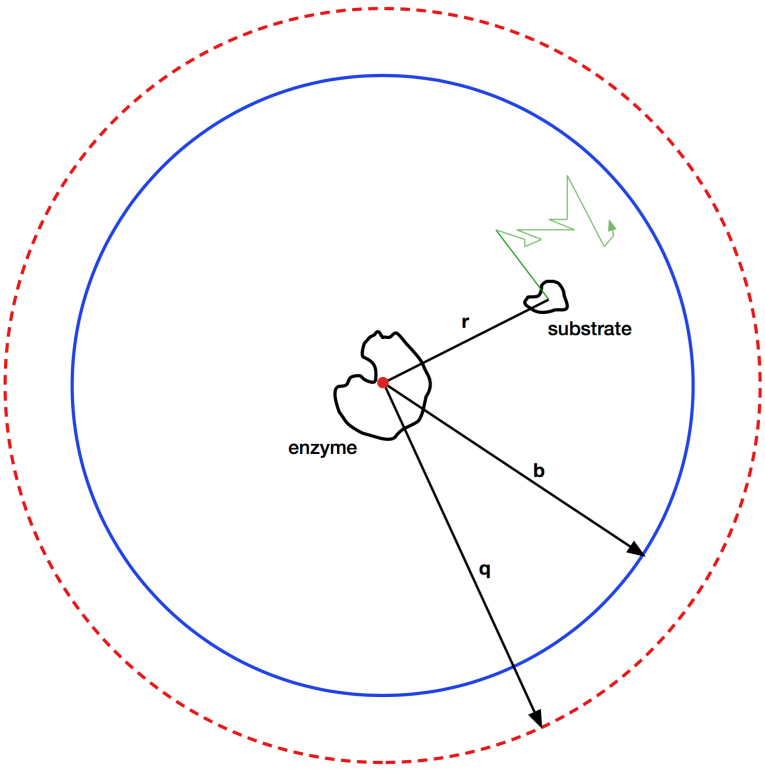


Figure 1.16: The division of the BD reaction scheme into an inner region (b) and an isotropic outer region; terminated at the quit-sphere of radius q. The substrate diffuses towards the enzyme for a successful collision or terminates its trajectory at the distance q.

inner region (r<b). The value b is chosen to be sufficiently large that the following conditions hold:

1. Interparticle forces between particles are centrosymmetric. Thus any charge asymmetry has negligible effect on diffusion at distances r>b.

2. The ensemble flux through the surface (r=b) has no angle dependence. However, angle asymmetry in the flux is isolated to the inner region.

The following relationship can be used to determine the association rate k from a BD simulation:

$$k = k_D(b)\beta^{\infty} \tag{1.25}$$

where $k_D(b)$ is the rate at which the two particle approach a separation of distance b and β^{∞} is the probability that the particles will go on to react rather an diffuse to infinite separation. $k_D(b)$ can be solve analytically, but β^{∞} must be computed by BD simulation.

If no forces or hydrodynamic interactions exist for r>b, $k_D(b)$ is given by the Smoluchowski result:

$$k_D = 4\pi Db \tag{1.26}$$

When forces and hydrodynamic interactions exist for r>b, but these forces are centrosymmetric, the following applies:

$$k_D(b) = 4\pi \left[\int_{\infty}^{b} \left(\frac{e^{\{E(r)/k_b T\}}}{r^2 D} \right) dr \right]^{-1} \tag{1.27}$$

The remaining and complicated asymmetric part denoted by β^{∞}, which is the probability that particles at separation r=b will react, involves diffusion through the inner region where asymmetric forces and reaction boundaries exist. This quantity is determined by running many BD trajectories and is given by:

$$\beta^{\infty} = \beta[1 - (1 - \beta)\Omega], \tag{1.28}$$

where Ω=$k_D(b)/k_D(q)$ and β is the fraction of trajectories in which encounter complex formation occurs before the particles diffuse to a separation distance of q. The β is a correction factor to account for the truncation of trajectories at a finite separation distance, q.

Analytical Treatment of Outer Surface Cut-off

Although the NAM algorithm allows the calculation of diffusion influenced reaction rates, this reaction scheme requires the distance to q to be an arbitrarily large value so that the particle flux is spherically symmetric. Subsequently, the NAM algorithm was extended in which the estimated value of q is replace by an analytical solution (Luty et al., 1992).

The division of the configuration space was slightly altered to eliminate the value of q. Simply, trajectories starting on b undergo diffusive motion until they either react or led to a surface of larger radius m. At the m-surface, a fraction P_{esc} (escape probability), is discontinued while the remaining particles are placed back on the b-surface according to a pre-computed distribution $\omega(\theta,\varphi)$. This cycle is repeated until all trajectories react or es-

cape to infinity. The fraction of reacting trajectories give β^{∞} and the rate constant is again calculated using the relationship of Eq 1.25.

1.5 Motivation

The formation of enzyme substrate complexes is fundamental for the processes of life. Electrostatic forces are essential for the interaction of virtually all biomolecules and can guide and steer incoming molecules towards their targets, in some cases over large distances (Hemsath et al., 2005).

The recent development of a technique for purifying tubulin using its binding affinity for TOG-domains has highlighted the possible role of electrostatics in TOG-tubulin complex formation (Widlund et al., 2012). A protocol was developed for the production of an affinity matrix comprised of TOG domains covalently coupled to a sepharose support. The resin has a high capacity to specifically bind to tubulin from crude cell extracts and elute tubulin under conditions of high salt. This may suggest that the binding affinity of TOG domains is modulated by its electrostatic environment.

In this thesis we explore the electrostatic characteristics of TOG domains and the role of electrostatic interactions in the formation of TOG-tubulin complexes. The two main areas of study are:

1. The electrostatic similarity analysis of TOG domain homologs and the effect of single-point mutations on the overall electrostatic field.

2. The role of electrostatics in the steered-diffusion of TOG-domains to the tubulin-binding site.

2 STRUCTURAL & ELECTROSTATIC ANALYSIS

2.1 Introduction

MT polymerases from the XMAP215/Dis1 family are known to promote fast MT assembly and dynamics through interactions with conserved TOG domains in their N-termini. These domains can potentially exert their influence on MT dynamics by selecting for distinct conformations of tubulin. TOG domains share a common architecture. Each domain consists of six HEAT-repeats (HRs) (Andrade and Bork, 1995)(Andrade et al., 2001) stacked side-by-side to form a helix-turn-helix motif. HR structure differs from similar repeats found in other proteins by its N-terminal helix lacking an orthogonal kink, and this results in the domain adopting a relatively straight, paddle-like topology (Al-Bassam et al., 2007).

TOG domains bind to tubulin through intra HEAT-repeat loops on one face of the domain (Fig. 2.1 C) that interact with a concave, negatively-charged region on the tubulin dimer (Fig. 2.1 B). Residues in these loop regions are highly conserved and single point-mutations are known to disrupt tubulin binding (Al-Bassam et al., 2007). TOG domains engage tubulin across its intra-dimer interface (Fig. 2.1 A) in an architecture that precludes the binding of a second subunit (Ayaz et al., 2012). TOG domains bind to unpolymerised tubulin in a nucleotide independent manner (Ayaz et al., 2014).

Cells comprise of hundreds of thousands of different types of biological macromolecules that interact with each other to maintain the function of the cell. The interactions between these molecules are highly orchestrated and specific. Protein macromolecules are able to recognise binding partners in a highly crowded environment with many competing substrates (Zhou, 2013). Often this recognition process is fast and may be guided over a long distance. The most straight-forward contender for such a long-range force is electrostatics.

Electrostatic forces are thought to be essential for virtually all interactions between biological macromolecules. Long-range electrostatic interactions can steer protein molecules towards their pre-binding orientations. Relatively small nudges can be surprisingly effec-

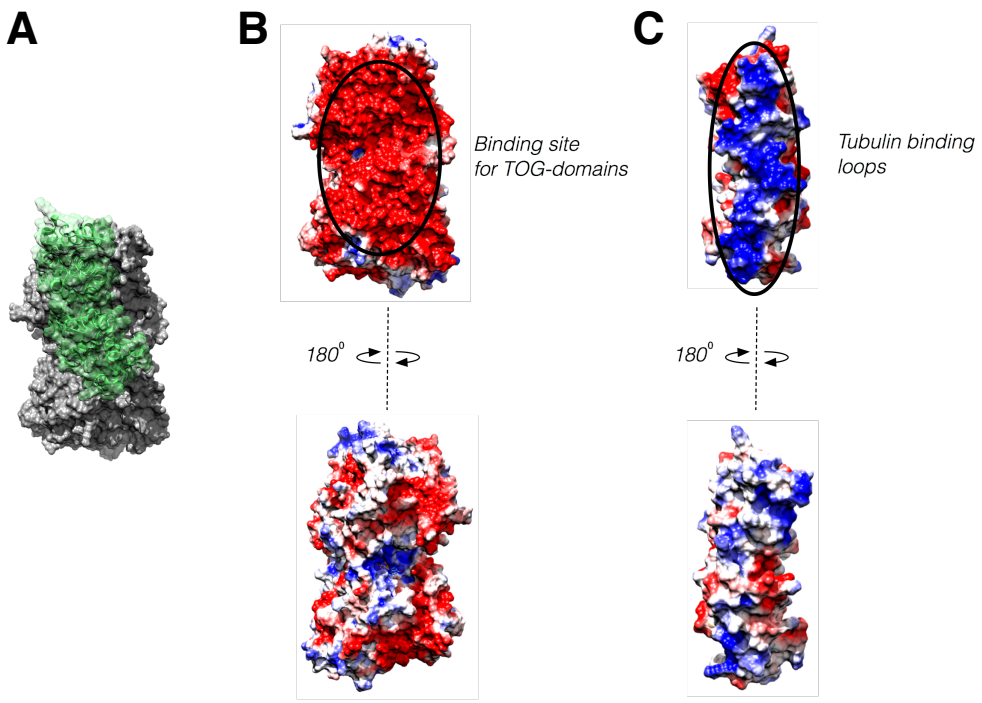


Figure 2.1: Molecular graphics representation of the electrostatic potential of the TOG:$\alpha\beta$-tubulin complex. The (**A**) tubulin and TOG domain (**B**) are rotated about the vertical axis by 180° to illustrate their electrostatic complementarity at each interface, each highlighted inside the ellipses. Molecular surfaces are coloured based on electrostatic potential values, with a gradient from -5 k_bT/e to +5 k_bT/e (blue).

tive. For example, the rate of enzyme degradation of some neurotransmitters is known to increase significantly as a result of electrostatic steering of negatively-charged inhibitors towards the active site region of the enzyme (Meltzer et al., 2006). Other examples include the action of snake venom on the muscle action of its prey and diffusion-controlled polymerisation processes such as the polymerisation of actin (Sept et al., 1999).

The biological functions of macromolecules originate with thermal fluctuations and energy-dependent conformational rearrangements. A detailed knowledge of the dynamics of macromolecules enables a more thorough understanding of biological processes, including the potentially conformation-dependent role of electrostatic steering. These protein dynamics cover a broad timescale still not accessible to most time-dependent theoretical methods. One way to address this is to study large-scale protein flexibility using a time-independent approach such as normal mode analysis (NMA).

TOG tubulin binding is known to be modulated by ionic interactions, a feature recently exploited in the purification of tubulin via its affinity for TOG domains. Using a single species TOG domain, the yield and purity of tubulin captured from crude extracts varied significantly (Widlund et al., 2012). This indicates that the TOG domain is binding different tubulin

species with differing affinity. Although tubulin is structurally similar across species, its electrostatic field can vary due to varying isoforms and post-translational modifications in cells (Tuszynski et al., 2006). TOG domains, therefore, might have evolved distinct functions to match or compensate for the diversity seen in tubulin species. Studying the electrostatics of TOG domains may potentially shed light on their individual species-specific functions.

One putative mechanism of TOG tubulin binding is that TOG domains bind to unpolymerised tubulin with equal affinity, independent of its nucleotide state and are released upon incorporation into the MT lattice, as a result of the tubulin becoming straight. This conceptual mechanism agrees with recent TOG tubulin crystal structures which suggest that TOG domains are rigid and do not change conformation upon tubulin binding. Nonetheless, flexibility in the TOG domains might potentially aid in the capture of free tubulin. Studying the degree of flexibility in different TOG domains within and across species can address the question of whether TOG flexibility plays a role in the specificity and dynamics of TOG tubulin binding.

2.2 Aims

1. Does electrostatic analysis identify invariant regions that are resistant to mutation?

2. Do different classes of TOG domains have distinct electrostatic fields, reflecting for example their array position?

3. Which residues contribute to flexibility within TOG domains?

4. Which residues in each TOG domain are important for enhancing or inhibiting tubulin binding?

2.3 Results & Discussion

2.3.1 Cumulative electrostatic similarity distribution for 14 TOG homologues projected on to the surface of S. cerevisiae TOG1

Conserved electrostatic potentials are known to be responsible for the strengthening and accelerating association of protein-protein interactions. Areas of highly conserved potential are often indicative of the location of functional sites.

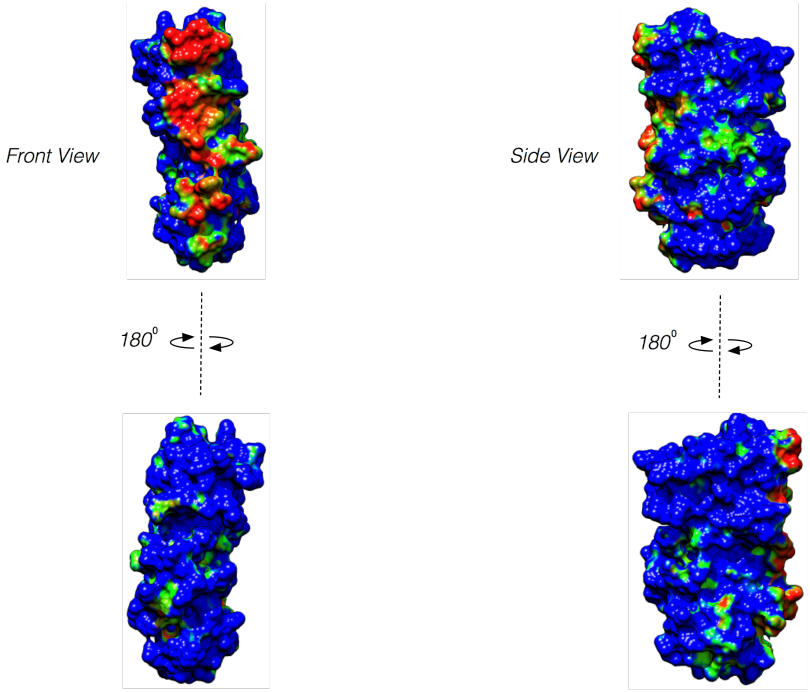


Figure 2.2: Conservation of electrostatic potential for homologues of TOG-domains. The cumulative electrostatic similarity distribution projected onto the surface of a representative TOG domain. The colour transitions from blue-green-red correspond to ESI values of 0.10-0.15-0.40.

Electrostatic calculations were performed on a combination of available TOG domain crystal structures and homology based models. This data set was chosen to span a range of species from the lower to higher eukaryotes in which most of the full-length proteins containing the domains have undergone some form of functional characterisation. The chosen domains are relatively diverse in net charge (-3 to +6) and span a sequence identity of ~13% to 53%. Each calculated field was scored by an electrostatic similarity index (ESI) (see Methods). The cumulative distribution of the ESI scores is illustrated in the surface projection in Fig. 2.2. ESI calculations spanning a range of eukaryotic TOG-family mem-

bers reveal conserved patterns of surface charge distribution, despite large variations in the electrostatic potentials of the studied homologs.

The ESI distribution illustrates that homologs possess invariant regions of positive charge (Fig. 2.2, highlighted in red), which correspond to the functional site of the TOG-domain and its established role in binding to negatively-charged tubulin dimers. The loops in this region contain many conserved residues of lysine and arginine, making the net charge in this region highly positive. The highest ESI scores are located in HR A-D that engages with β-tubulin. Specifically, the highest conservation values are found in the T1 binding loop as one moves towards the second α-helix of HR A. In many TOG-structures, this position is the site of a solvent exposed tryptophan that when mutated to alanine inhibits tubulin binding (Slep and Vale, 2007).

The electrostatic potential surrounding the functional site shows little or no conservation, which suggests that distal residues have distinct functions within the TOG-array. Overall the ESI distribution reveals a low level of conserved electrostatic potential outside of this known functional site, as shown by the blue surface projections.

2.3.2 Perturbation maps of TOG electrostatic fields based on theoretical alanine-scanning mutagenesis of charged residues.

Over the course of evolution, it is likely that the electrostatic field of TOG domains have been tuned by natural selection. On this basis it would be reasonable to expect that mutations to key residues that are critical to the protein's function would be the most resistant to perturbation of its field. Otherwise, these mutations would manifest themselves as a reduction in, or loss of protein function.

To test this idea, computational perturbations of the electrostatic potential were performed on each TOG domain by replacing charged residues with alanine, a charge-neutral residue. Cumulative ESI distributions were calculated for each domain based upon a range of mutant potential maps. The resulting ESI distributions were used to identify regions of high electrostatic similarity, i.e. those least affected by mutation.

All perturbation maps identified similar characteristic regions of high similarity, with most of the resistance to perturbation aligning to the front (MT binding) face of the TOG domain and to a lesser extent to its rear. In comparison to the overall field of Fig. 2.3, conservation along the tubulin binding region is broader and encompasses both the loops and flanking regions that connect to the HR helices.

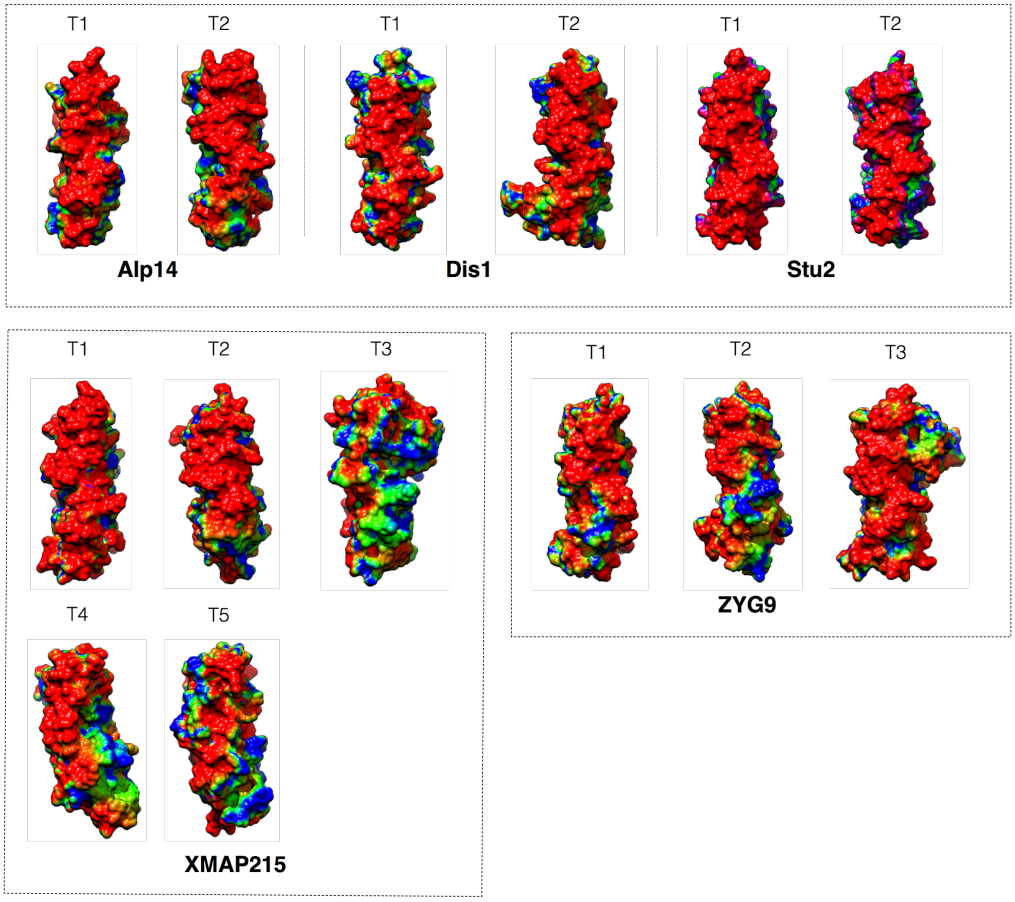


Figure 2.3: Mutation-based perturbation maps for 14 TOG homologs. The cumulative electrostatic similarities for alanine perturbations are projected on to the surface of each representative TOG structure. Perturbation map colour scheme is blue-green-red; low to high similarity, which corresponds to ESI values of 0.5-0.7-0.9.

2.3.3 Electrostatic clustering analysis of available and homology-derived TOG domains.

Electrostatic maps were computed for TOG domains at 0 mM ionic strength and clustered in order to distinguish differences between the potentials of each structure. The calculations were performed by pairwise comparison of an electrostatic similarity distance metric (see Methods). The data set consists of a combination of homology-derived structures and their corresponding experimentally determined templates.

Overall in the calculated dendrogram of Fig 2.4, TOG domain structures are observed to fall into 5 distinct clusters; signified by the colour of each branch. All TOG1 domains form a single cluster irrespective of species. Within this grouping Stu2, XMAP215 and Alp14 share the highest similarity, followed closely by Zyg9 and Dis1. Similarly, the Stu2 TOG2 domains, crystallised alone (2QK1) and in complex with tubulin (4U3J), show the highest levels of similarity with all other domains and associate especially closely with the remaining TOG2 domains, with the exception of Dis1 and Zyg9. In vitro reconstitution experiments using purified XMAP215 have shown the contributions of individual TOG domains to microtubule growth. TOG1 and TOG2 are thought to bind tubulin more strongly and promote polymerase activity better than TOG3 and TOG4, whereas TOG5 is thought to have little or no effect (Widlund et al., 2011). The as yet undetermined function of TOG5 and its possible divergence from tubulin binding might explain why TOG5 forms a distinct branch in the dendrogram and does not cluster well alongside the other structures. Interestingly, Zyg9 TOG3 shares most similarity to Msps TOG4, whereas the remaining TOG3 domains from MSPS and XMAP215 cluster into a region surrounded by TOG1 structures. This pattern mirrors to some extent the results of sequence and structural analysis of the Msps TOG3 domain. This study found that Zyg9 TOG3 and Msps TOG3 share low-levels of sequence identity and Msps TOG3 maintains a TOG1- and TOG2-like tubulin binding surface (Howard et al., 2015). On the other hand, Zyg9 TOG3 is most similar to TOG5 from pentameric arrays but is still structurally distinct from any other TOGs in having an additional N-terminal HR. Dis1, the homologue in yeast of the Alp14, is thought to have a divergent function. Dis1 in *S. pombe* cells is implicated in contributing towards interphase microtubule bundling activity. Also, Dis1 together with the Dam1 complex, plays a role in the depolymerisation of microtubule plus-ends.

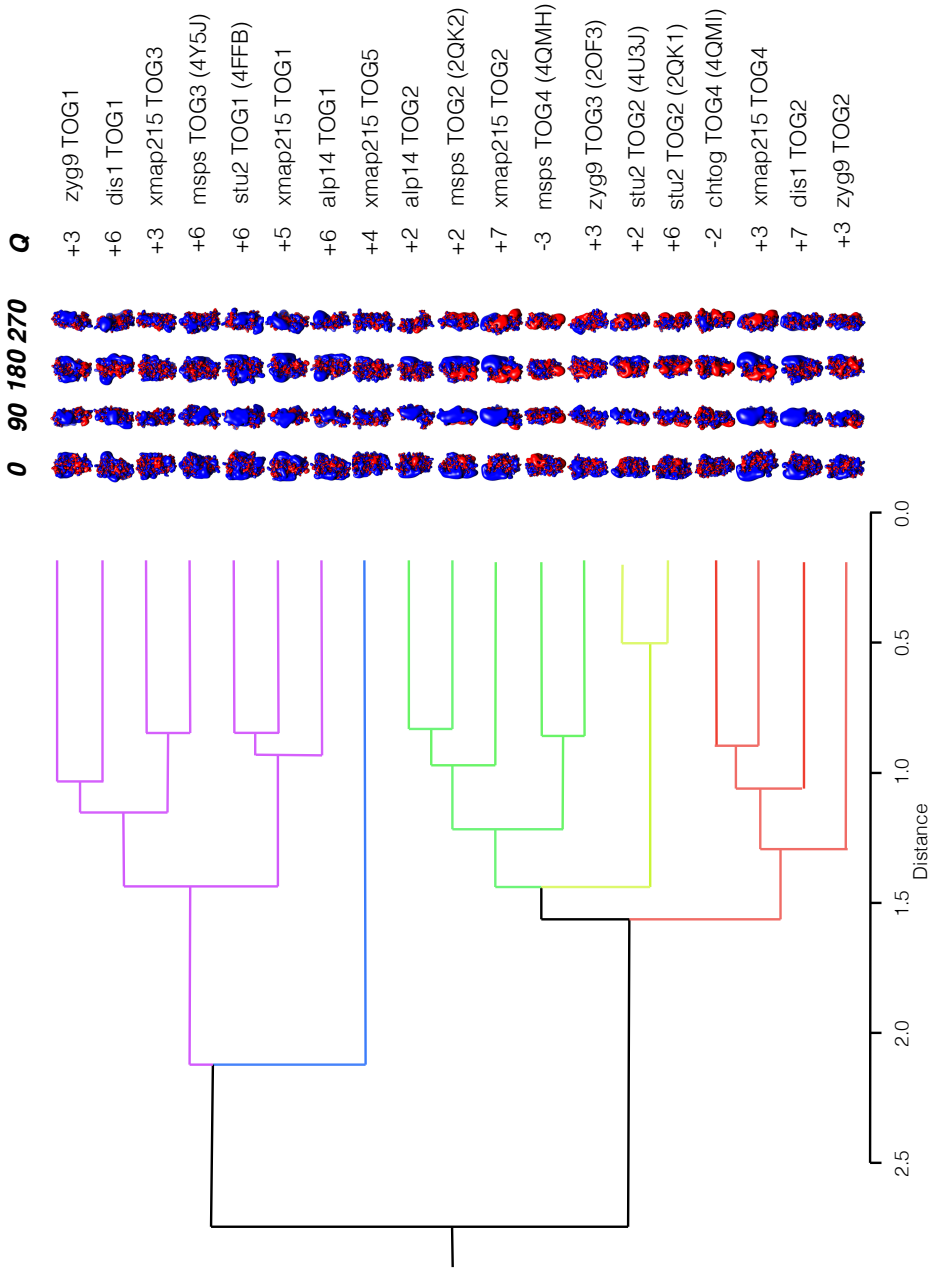


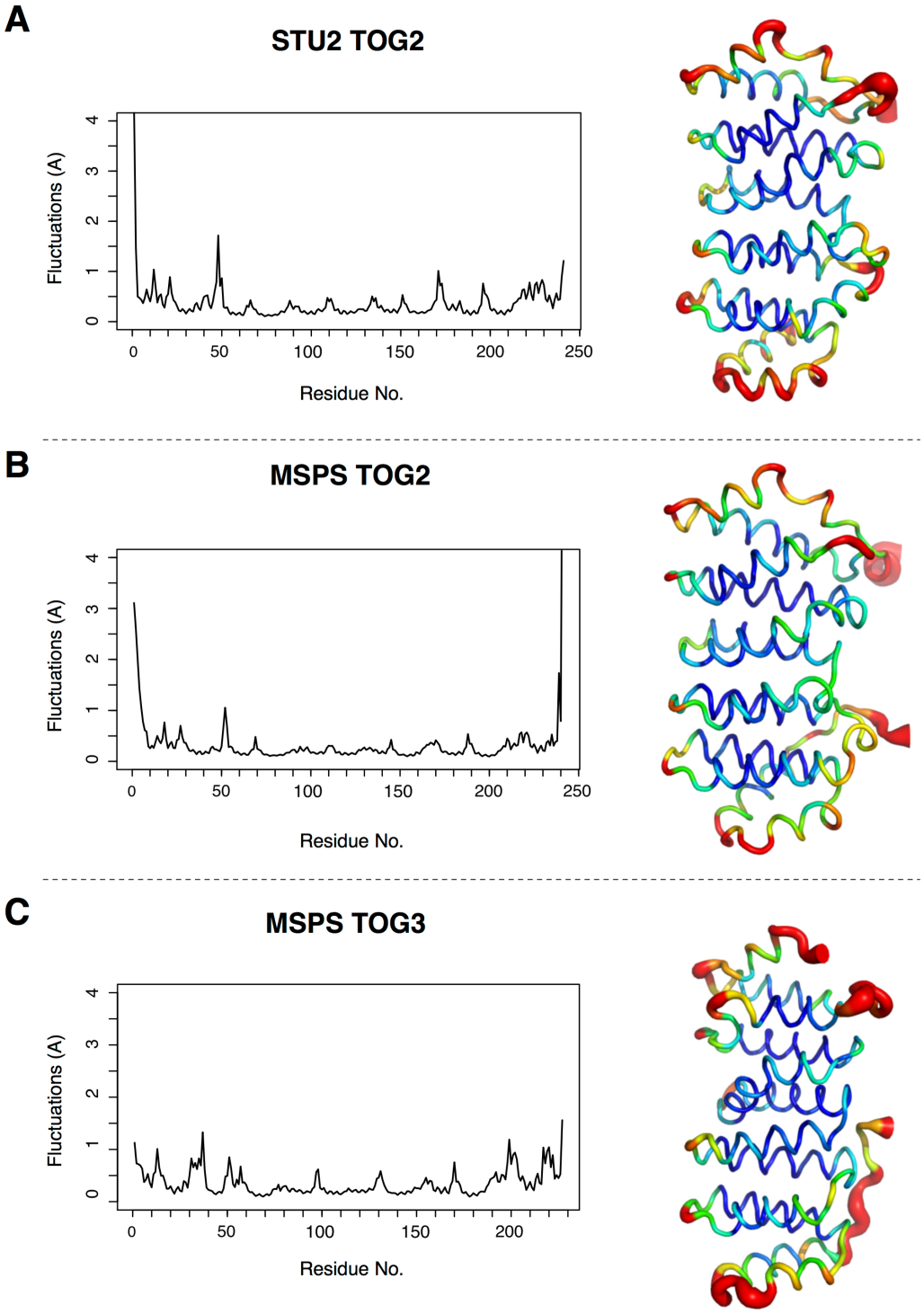
Figure 2.4: Electrostatic clustering analysis of TOG-domains. The horizontal axis on the dendrogram represents electrostatic similarity distance. Electrostatic potentials were calculated using an ionic strength of 0 mM salt concentration. Isopotential contours are present in four different orientations and plotted with thresholds of +/- 1 kbT/e, with blue and red corresponding to positive and negative electrostatic potential. The net charges are indicated in the figure for each subtype. Coloured branches in the directed graph represent clusters of similar electrostatic potential.

2.3.4 Normal Mode Analysis to characterise the intrinsic dynamics of TOG-homologues.

The molecular recognition and catalytic activity of a number of proteins are thought to be subtly modulated by protein dynamics. Many enzymes show significant variability in substrate and catalytic function, despite high-levels of conservation in their overall fold. Equally, functionally homologous enzymes often share low sequence similarity and identity which can make sequence-based studies of phylogeny and binding motif identification difficult. TOG domains share highly conserved structural features in terms of their overall fold but can be very diverse in terms of sequence conservation. In the following experiments, normal mode analysis (NMA) was undertaken in order to characterise the differences in residue fluctuations between different TOG domains that might be crucial for their selectivity to their tubulin substrate.

Analysis of the atomic fluctuation profiles of the six representative TOG domain proteins reveals some variability in their overall flexibility patterns (Fig 2.5). Helices and loop regions close to the amino termini exhibited the highest values of atomic displacements, as these regions contain the beginnings of the long linker regions that connect TOG domains in full-length proteins. The loop regions interconnecting the HRs on the front face the domain make up the tubulin binding loops, and are found in areas of high fluctuation. More specifically, the T1 and T3 binding loops that contain conserved residues that engage with β- and α-tubulin show higher levels of displacement than the other loops.

Flexibility of the active site in most proteins is considered a requirement to reduce the free energy barrier and accelerate enzymatic reactions. The range of flexibility seen in TOG domains may help to improve the domains affinity to different tubulin conformations. Residues in the tubulin binding loops are all solvent-exposed and therefore subject to higher motilities. No direct experimental evidence suggests that TOG domains undergo conformational change upon tubulin binding. This fact is indicative of the low degree of flexibility exhibited elsewhere in these domains across species.


A
STU2 TOG2
Fluctuations (A)
0
1
2
3
4
0
50
100
150
200
250
Residue No.
B
MSPS TOG2
Fluctuations (A)
0
1
2
3
4
0
50
100
150
200
250
Residue No.
C
MSPS TOG3
Fluctuations (A)
0
1
2
3
4
0
50
100
150
200
Residue No.

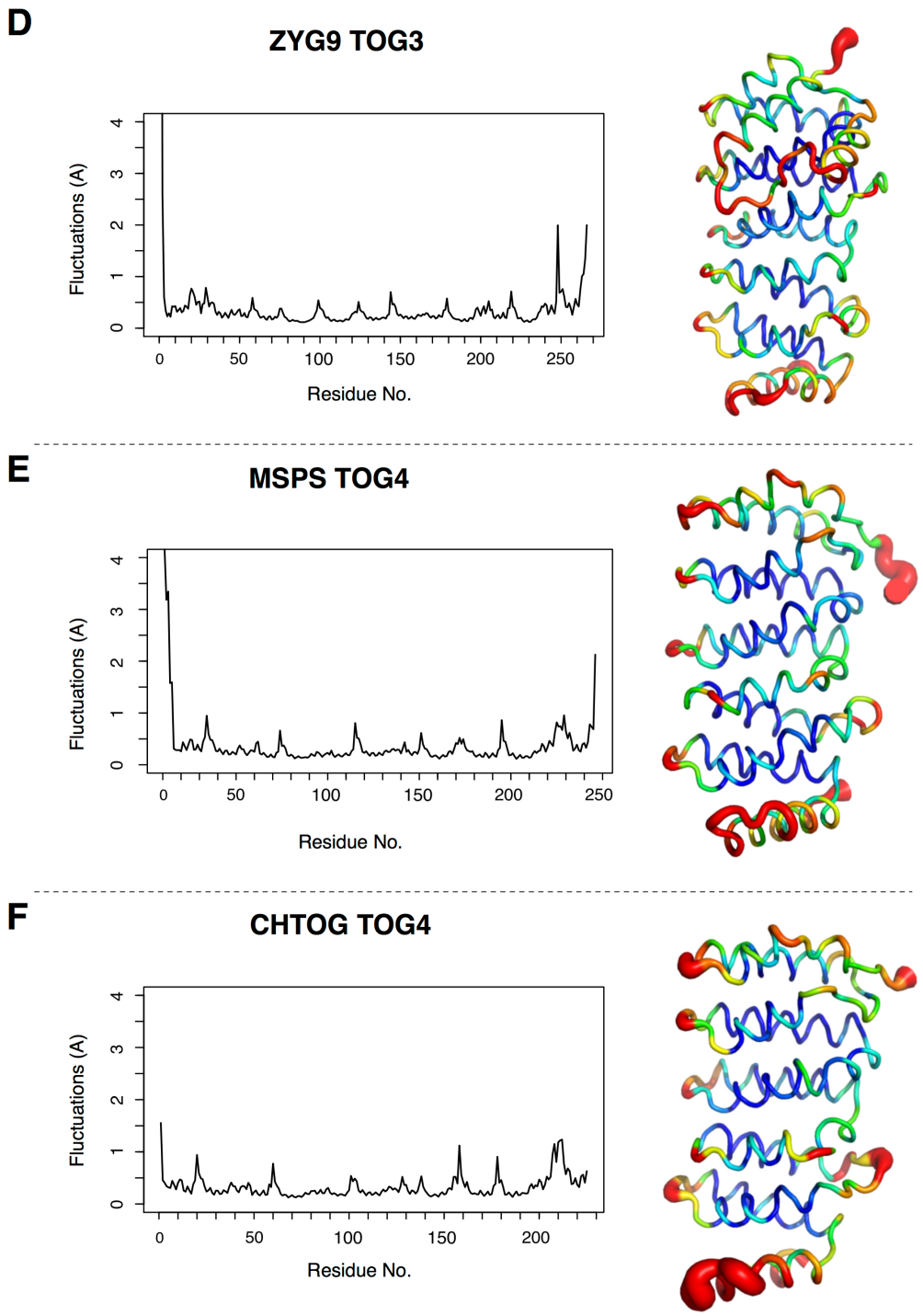


Figure 2.5: Plot and graphical representation of the atomic fluctuations of six TOG domain crystal structures taken directly from the protein databank. The plots show atomic fluctuations vs. residue number. For the visual representation of mode fluctuations, magnitude is represented by a thin to thick tube coloured blue (low fluctuations), white (moderate fluctuations) and red (large fluctuations).

2.3.5 Dynamic Cross Correlation Matrix Analysis of Individual TOGs

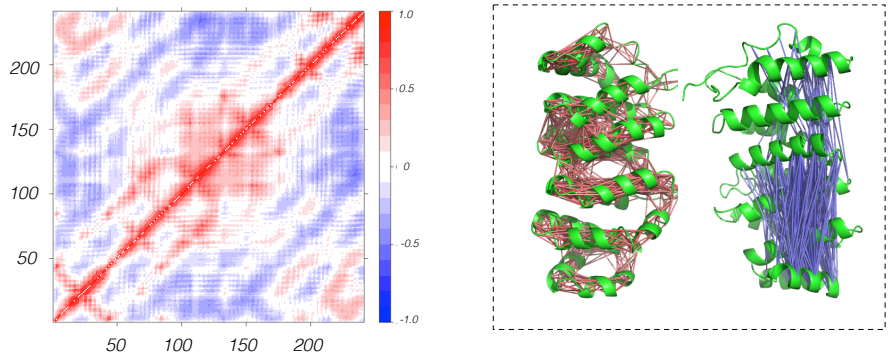
The atoms of a particular domain are likely to move with concerted, well-correlated motions. Structural domains and subdomains can in principle be identified by analysing the extent to which these motions are correlated with one another. These relationships can be expressed in matrix form by examining the magnitude of all pairwise cross-correlation coefficients for Cα atoms. Dynamic cross-correlations maps (DCCM) were computed for each representative TOG domain (Fig 2.6).

Most correlations are seen along the inner core of the large number of α-helices that make up the repeated nature of the TOG domain. Correlations are seen in the loop regions between HRs. However, this is less pronounced than the correlation seen in loops regions within individual HRs. This evidence indicates that the adjacent α-helix may be acting to stabilise the loop conformation and help to maintain its binding characteristics.

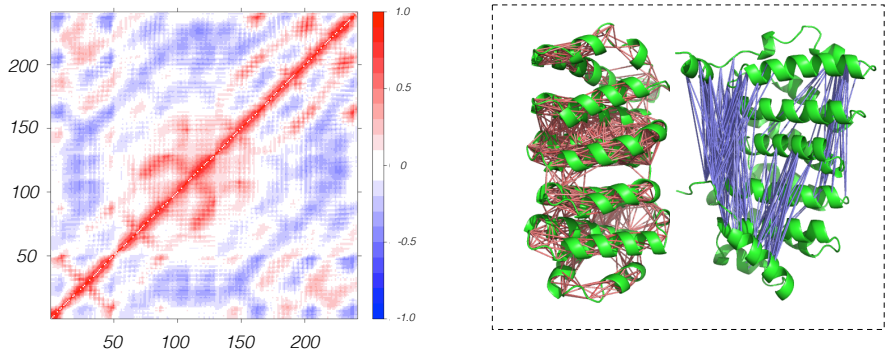
The most striking anti-correlation is seen between the N-terminal α-helix in the centre of the TOG domain in Stu2 TOG2, Msps TOG3 and Zyg9 TOG3. The remaining TOGs are only weakly anti-correlated between helices.

One explanation for this observation could be that Stu2 TOG2 and Msps TOG3 are using the anti-correlated movement of their helices to position the N-terminal helix in a conformation such that the T1 helix is optimally positioned for tubulin binding. The T1 loop contains a conserved tryptophan residue that is necessary for binding. This tryptophan residue is rendered immobile by a nearby salt bridge between R295 and D331, limiting its ability to bind. The anti-correlated movement enables a rotation invariant amino acid to sample a range of conformations therefore increasing the chance of tubulin binding.

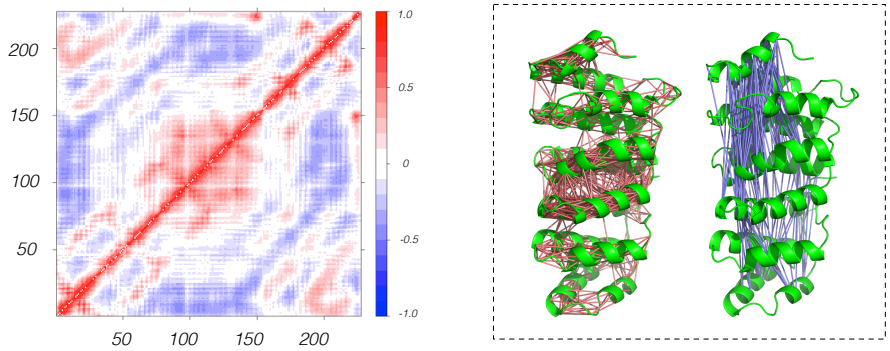
A **STU2 TOG2**



B **MSPS TOG2**



C **MSPS TOG3**



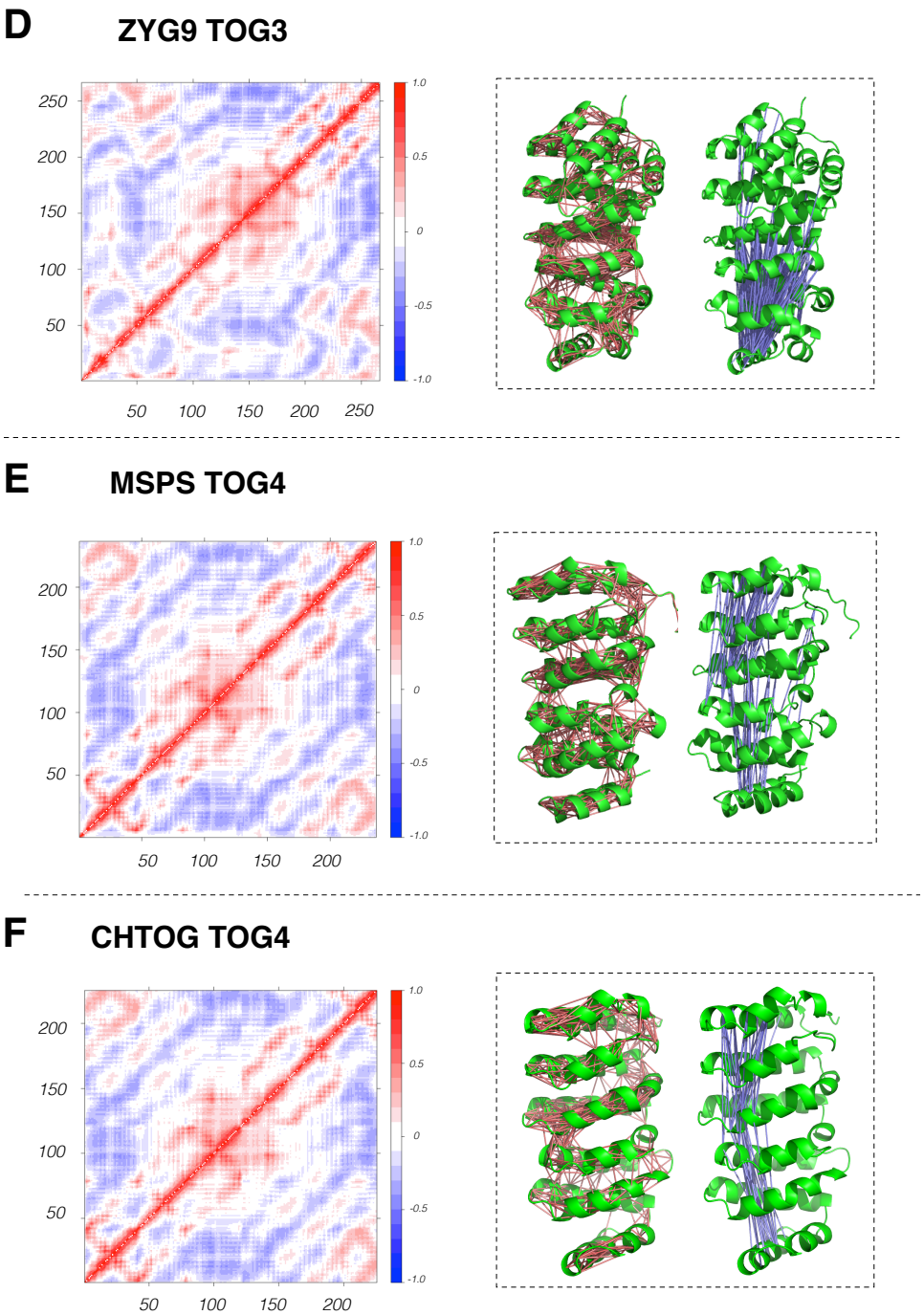


Figure 2.6: Correlation maps and 3D visualisation of residue-residue cross-correlation, showing correlated (red) and anti-correlated regions (blue) in TOG-domain structure. The values for each 3D structure range from +/- 0.4 to +/- 0.6.

2.3.6 Principal Component Analysis of TOG-Structures.

PCA analysis was undertaken on known crystal structures to compare structural characteristics of each domain. First, each TOG domain was aligned to an invariant core and pairwise comparisons were made of each protein domain by calculating their RMSD. These results are presented in the clustered-heatmap of Figure 2.7B. Following core identification and super-position, PCA was employed to examine the relationship between equivalent structures based on the xyz coordinates of their constituent residues. In short, the resulting principal components describe the axes of maximum variance amongst the structures in terms of the displacement of their atom positions. Figure 2.7A shows the projection of the structures plotted onto the two largest principal components. This lower dimensional representation groups the structures into three distinct clusters on the basis of their total mean squared displacement of their atom positions. TOGs 2-3 of Msps and Stu2 form a single cluster whereas Zyg9 is more closely related to the TOG4 of Msps and ch-TOG.

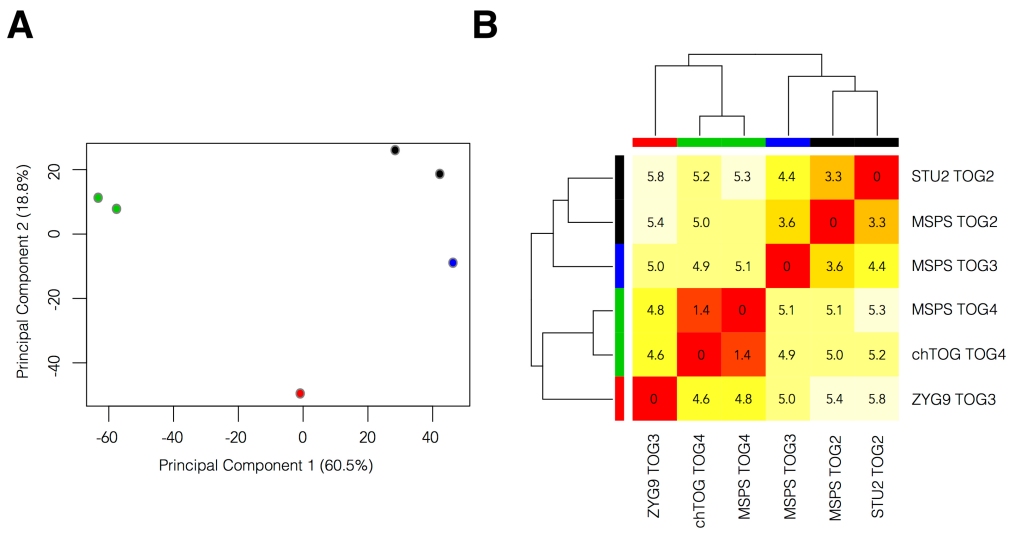


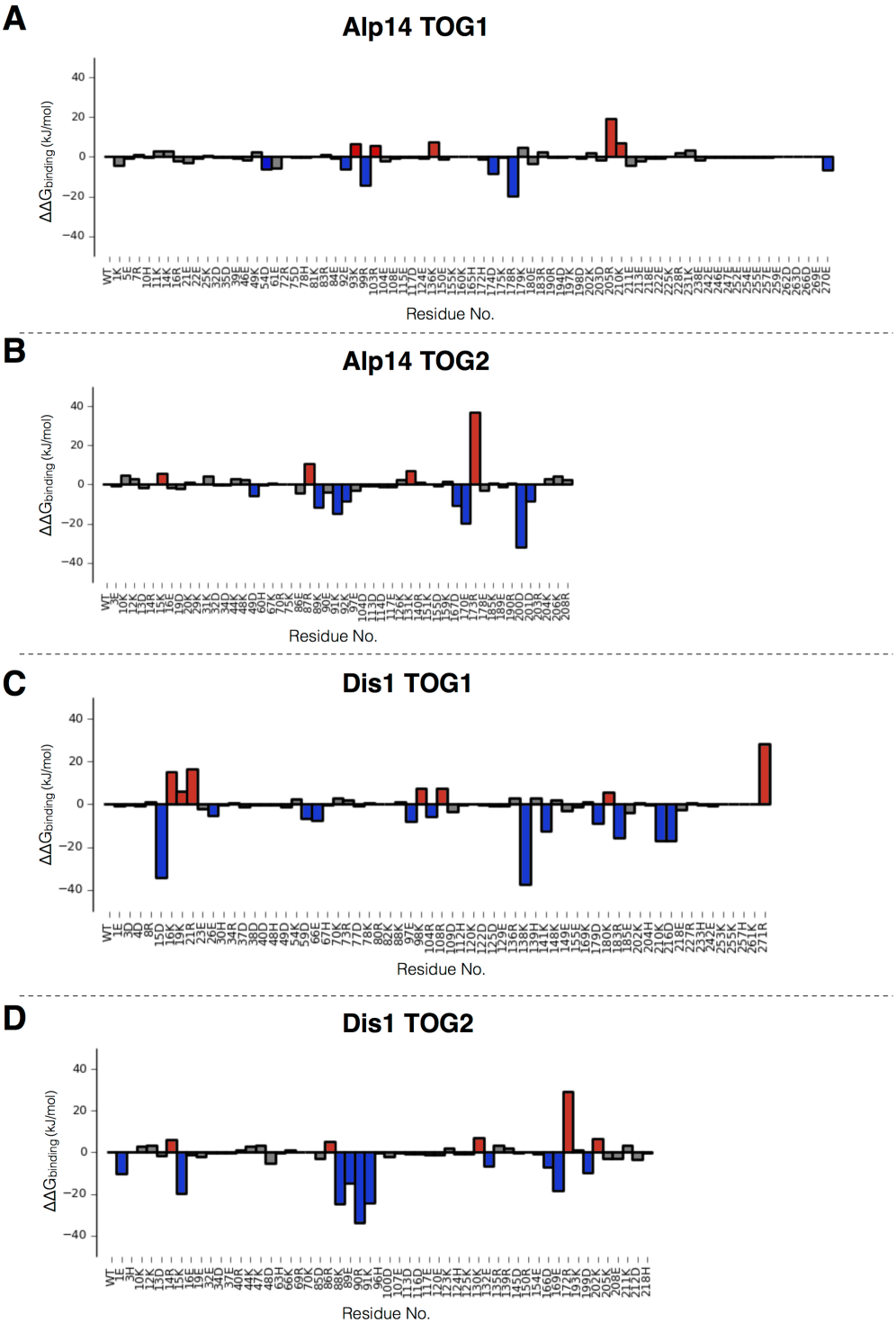
Figure 2.7: Principal component analysis of TOG-domains, structures form into four distinct groups - blue, black, red and green. (**A**) Clustering based on PC1-PC2 with percentages indicating the proportion of total variance. (**B**) Heatmap based on RMSD between structures, with calculated distances in angstroms displayed in each box.

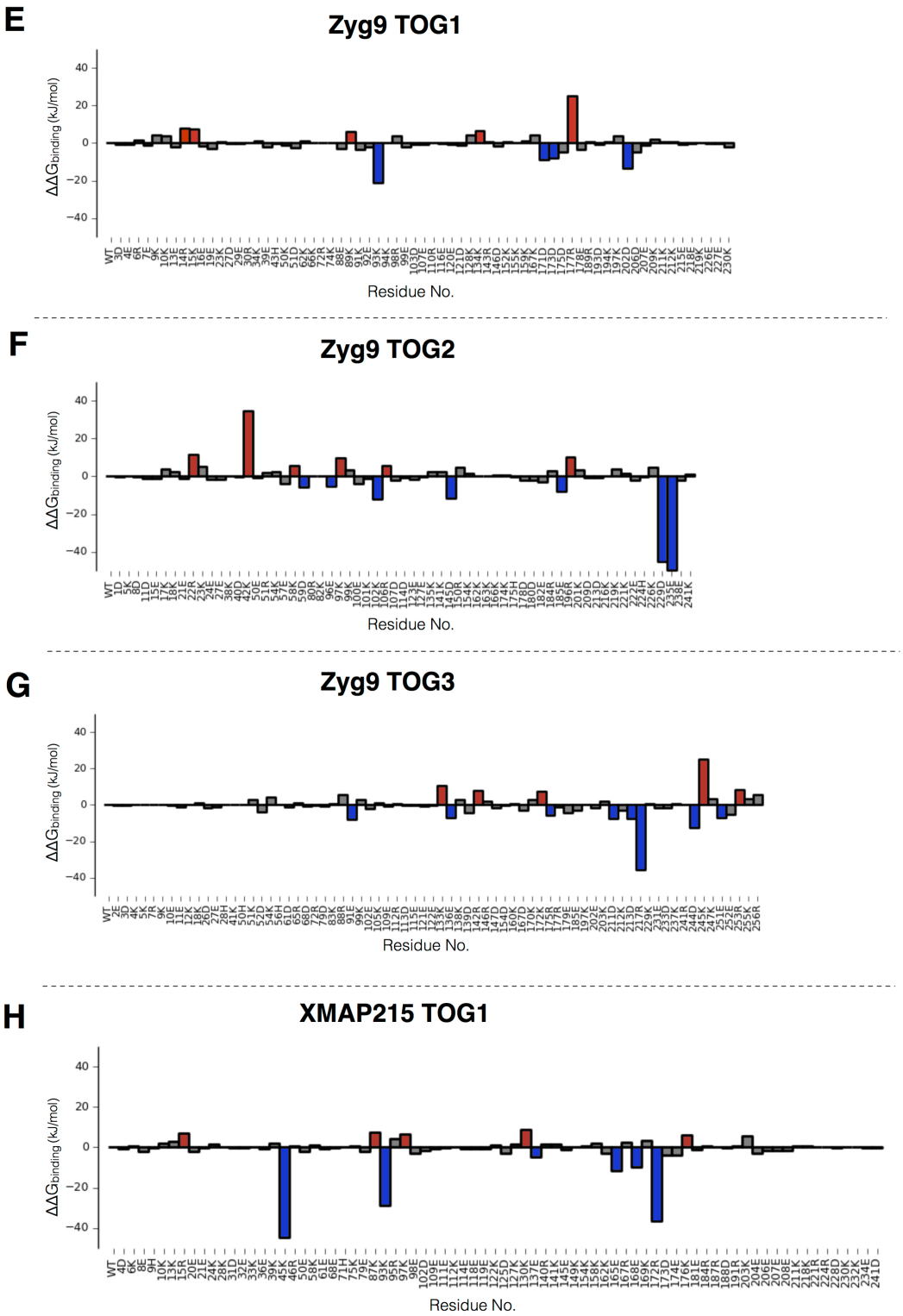
This analysis shows that TOG domains can be successfully clustered by a simple structural measure which also shows correlations to their putative functional roles. Most structural differences occur in the N-terminal region of the domain. This region of TOGs is the site of interaction with β-tubulin, which undergoes a number of structural rearrangements during MT assembly and disassembly. If the function of the different TOG domains is position-

dependent within the TOG array, and this relates to tubulin binding activity, then it is unsurprising that the N-termini of the TOG-domains show structural divergence.

2.3.7 Electrostatic free energies of association.

To examine the stability of TOG-tubulin binding, electrostatic free energies for the TOG:$\alpha\beta$-tubulin complexes were calculated based on the magnitudes and perturbation of these energies when residues were mutated to alanine. Each structure underwent MD energy minimisation and refinement (see Methods) as a complex. Individual residues were identified that led to a >5 kJ/mol change in the free energy. Positive $\Delta G_{binding}$ values correspond to a gain the free energy of association, while negative values correspond to a loss for that residue. Subsequently, residues were ranked and two mutations from each TOG domain were identified as inputs for Brownian Dynamics simulations performed in Chapter 3.


A
Alp14 TOG1
ΔΔGbinding (kJ/mol)
Residue No.
B
Alp14 TOG2
ΔΔGbinding (kJ/mol)
Residue No.
C
Dis1 TOG1
ΔΔGbinding (kJ/mol)
Residue No.
D
Dis1 TOG2
ΔΔGbinding (kJ/mol)
Residue No.


E
Zyg9 TOG1
ΔΔGbinding (kJ/mol)
Residue No.
F
Zyg9 TOG2
ΔΔGbinding (kJ/mol)
Residue No.
G
Zyg9 TOG3
ΔΔGbinding (kJ/mol)
Residue No.
H
XMAP215 TOG1
ΔΔGbinding (kJ/mol)
Residue No.

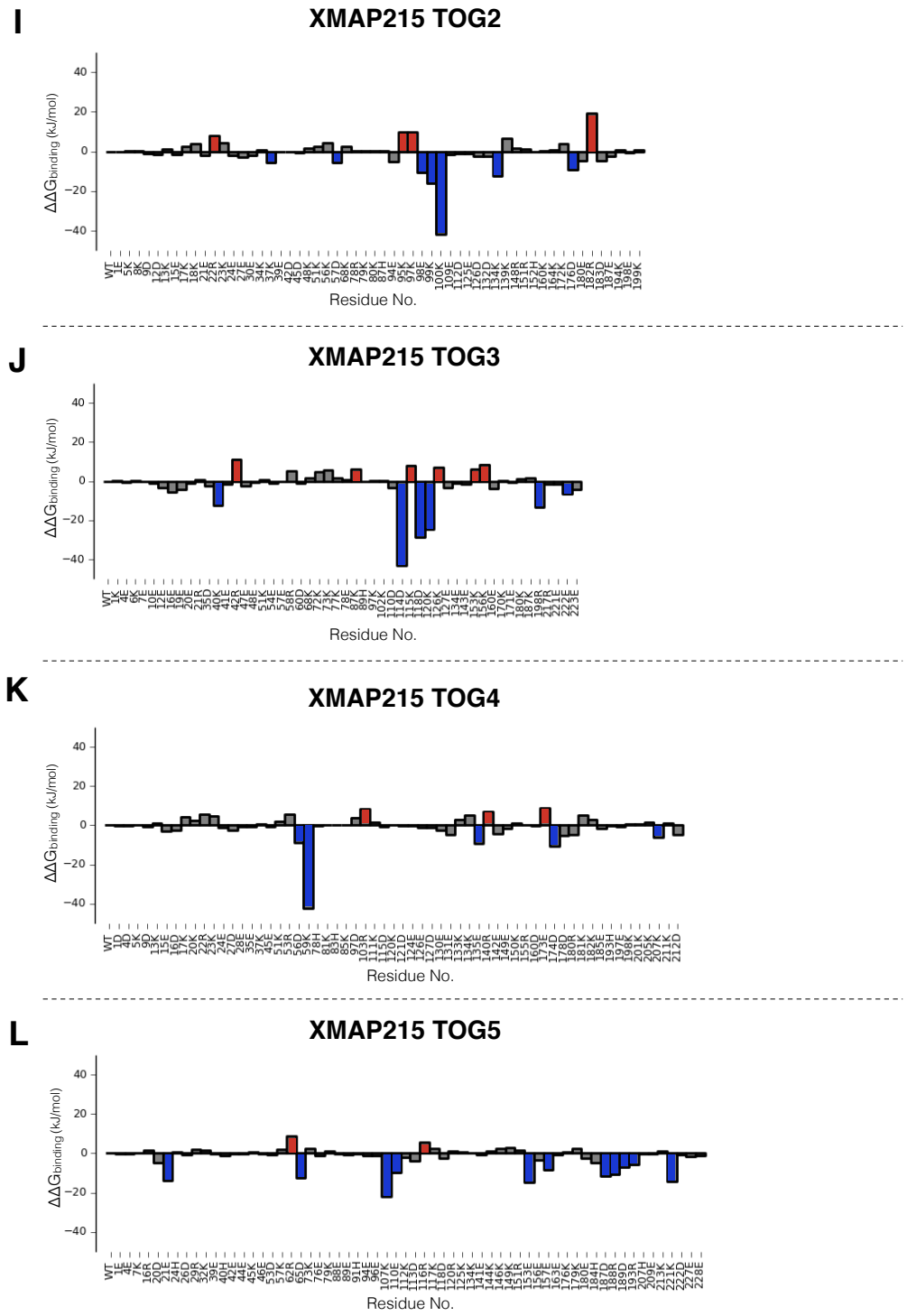
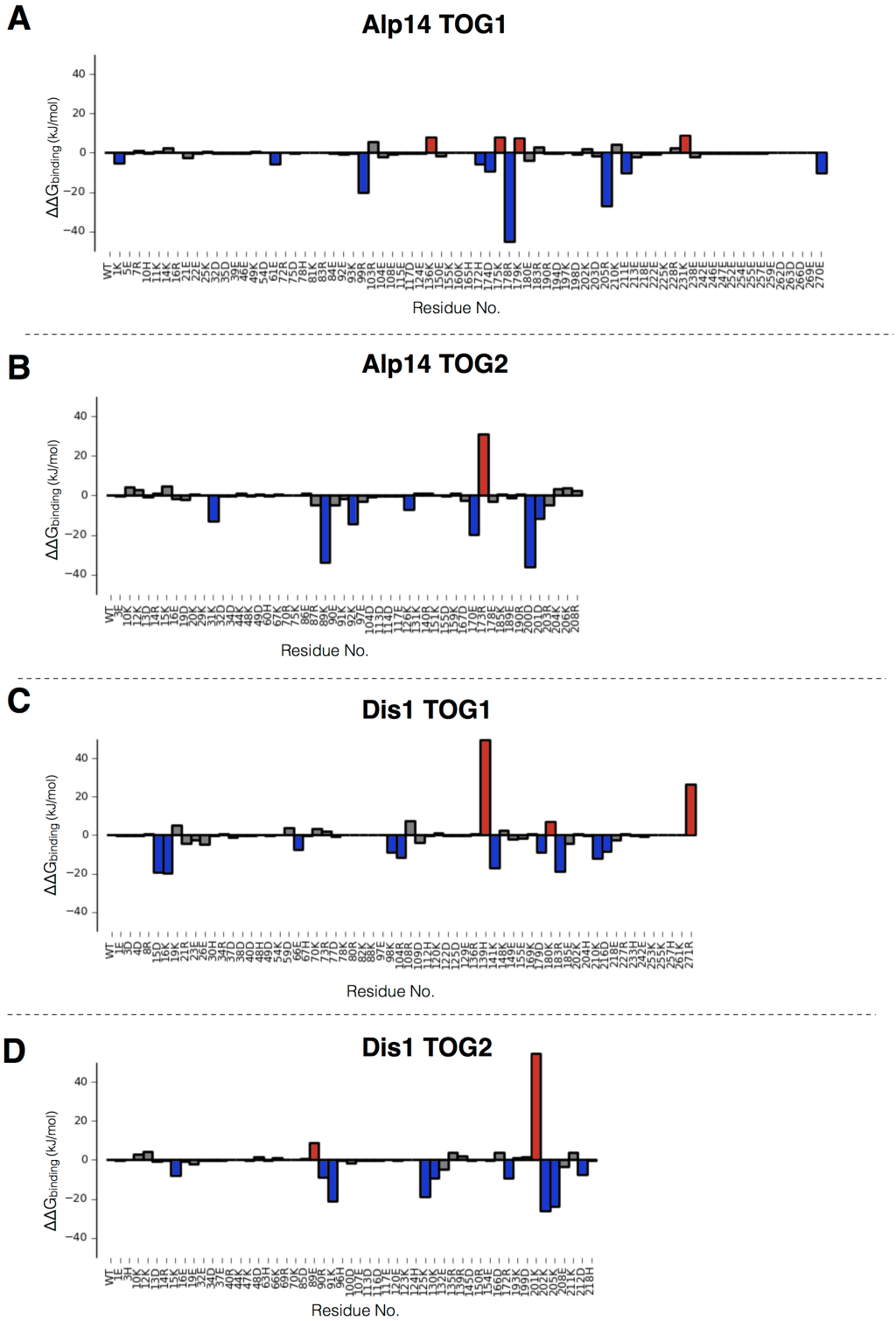
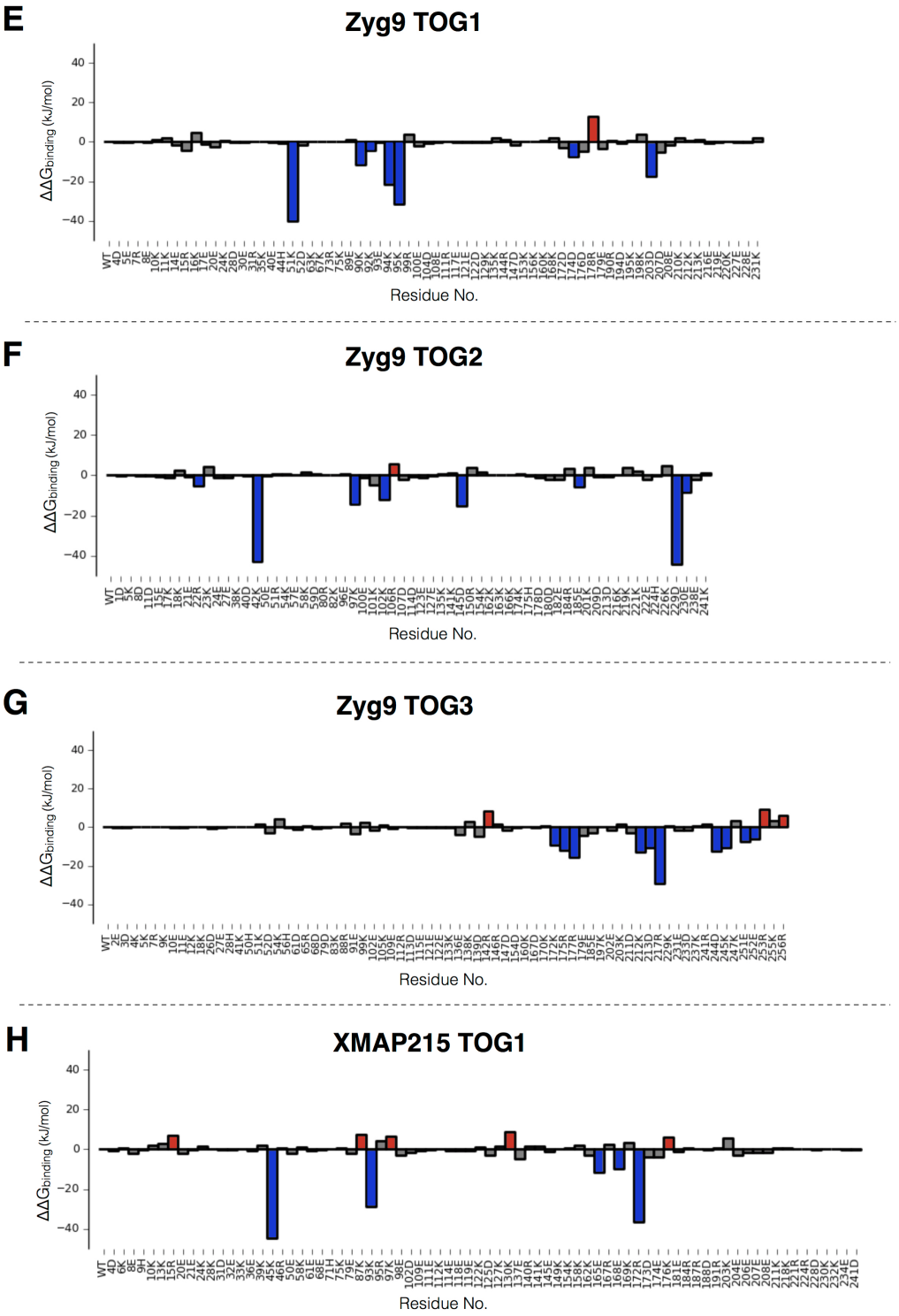


Figure 2.8: Electrostatic binding free energies of TOG-αβ-tubulin complex with cognate tubulin. A positive value indicates a loss of binding (red) and negative value indicate a gain in binding (blue).


A
Alp14 TOG1
ΔΔGbinding (kJ/mol)
Residue No.
B
Alp14 TOG2
ΔΔGbinding (kJ/mol)
Residue No.
C
Dis1 TOG1
ΔΔGbinding (kJ/mol)
Residue No.
D
Dis1 TOG2
ΔΔGbinding (kJ/mol)
Residue No.


E
Zyg9 TOG1
ΔΔGbinding (kJ/mol)
Residue No.
F
Zyg9 TOG2
ΔΔGbinding (kJ/mol)
Residue No.
G
Zyg9 TOG3
ΔΔGbinding (kJ/mol)
Residue No.
H
XMAP215 TOG1
ΔΔGbinding (kJ/mol)
Residue No.

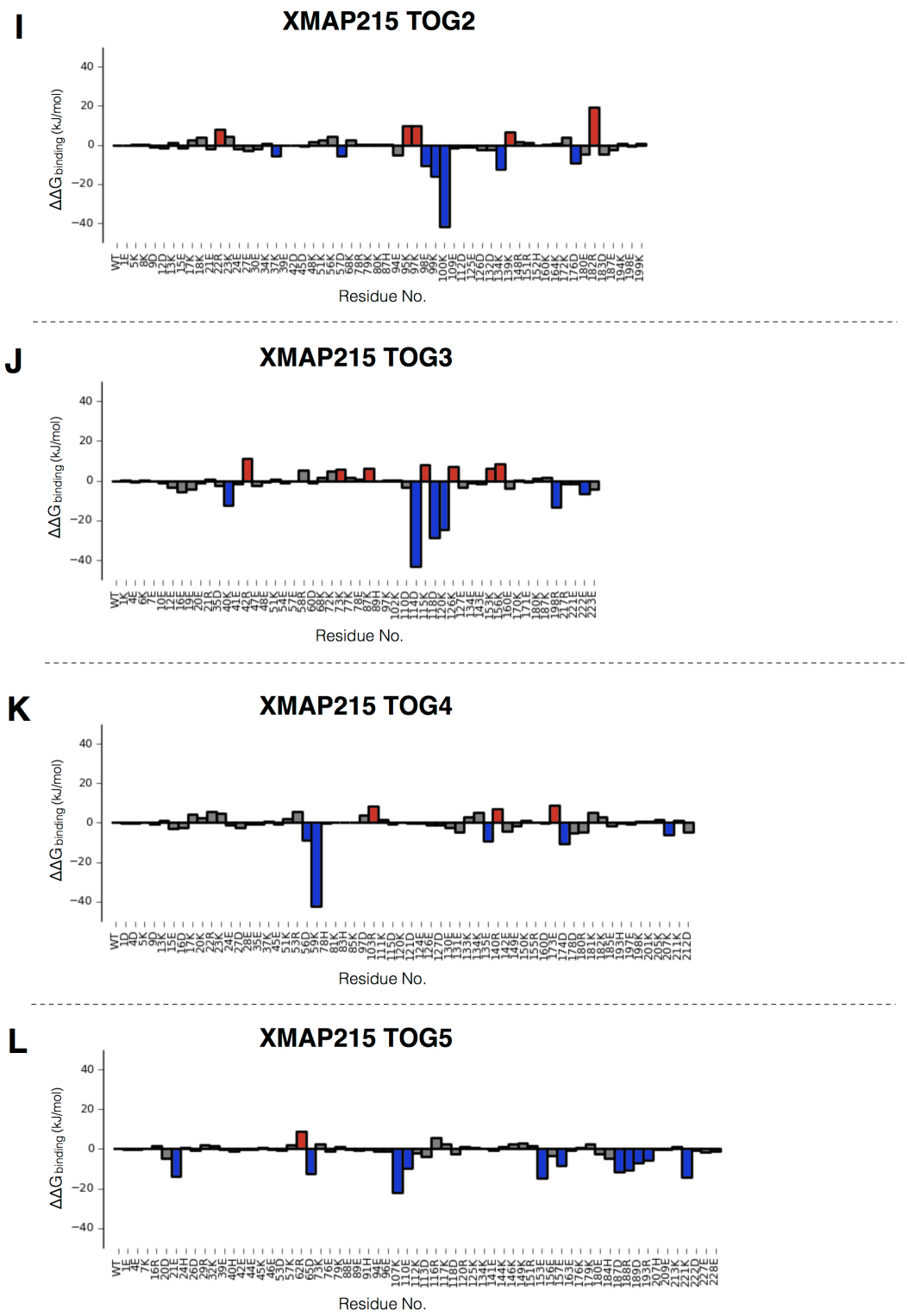


Figure 2.9: Electrostatic binding free energies of TOG-αβ-tubulin complex with porcine tubulin. A positive value indicates a loss of binding (red) and negative value indicate a gain in binding (blue).

2.4 Methods

2.4.1 Structure Preparation

Protein structures were acquired from the Protein Data Bank (PDB), see Table 2.1. Missing loop regions and homology models were created and refined using MODELLER (version 9.14) (Sali and Blundell, 1993). Template structures for comparative modelling were selected by multiple sequence alignment. The template with the highest identity score to the sequence was selected for modelling. Each structure underwent MD refinement, see section 3.4.2 for details.

Name	PDB Code	Type	Template
Alp14 TOG1		Homology model	4FFB
Alp14 TOG2		Homology model	4U3J
Dis1 TOG1		Homology model	4FFB
Dis1 TOG2		Homology model	4U3J
Stu2 TOG1	4FFB	Native	
Stu2 TOG2	4U3J	Native	
Zyg9 TOG1		Homology model	4U3J
Zyg9 TOG2		Homology model	4U3J
Zyg9 TOG3	2OF3	Native	
XMAP215 TOG1		Homology model	4FFB
XMAP215 TOG2		Homology model	4U3J
XMAP215 TOG3		Homology model	2OF3
XMAP215 TOG4		Homology model	4QMI
XMAP215 TOG5		Homology model	4QMI
MSPS TOG2	4QK2	Native	
MSPS TOG3	4Y5J	Native	
MSPS TOG4	4QMH	Native	
ch-TOG TOG4	4QMI	Native	

Table 2.1: List of TOG domain structures used in the electrostatic analysis and as templates in homology modelling. The structures highlighted in bold form the core model structures of the dataset.

Electrostatic Calculations

Electrostatic calculations were performed with APBS (1.4.1) (Baker et al., 2001), using AMBER charges and radii at pH 6.9 obtained from PDB2PQR (version 1.9.1). The full, nonlinear Poisson-Boltzmann (PB) equation was solved for each protein. Atomic charges were mapped to grid points via B-spline discretisation. The protein boundary was delineated by a dielectric boundary between solute (ε_p of 2) and solvent (ε_s of 78.54). The surface was defined as the molecular surface (srfm: smol and srad: 1.40). The APBS calculations were performed on a grid that provided a minimum resolution of not greater than 0.5 Å.

Electrostatic Similarity Analysis

Electrostatic potentials for TOG domains were analysed using the analysis of electrostatic similarity of proteins (AESOP) framework (Gorham et al., 2011). AESOP was used to quantify the effects of charged residues on TOG-tubulin association in combination with custom python scripts.

The PIPSA (Protein Interaction Property Similarity Analysis) software was used to compare the TOG domains with respect to their electrostatic potentials (Blomberg et al., 1999). Distance matrices were computed at 0mM salt concentration and domains grouped into dendrograms using complete hierarchical clustering in R.

The similarity indexes (SI) for each pair of surface electrostatic potentials were calculated according to the following formula:

$$SI_{12} = \frac{2(P_1, P_2)}{(P_1, P_1) + (P_2, P_2)} \tag{2.1}$$

$$(P_1, P_2) = \sum_{i,j,k} \phi_1(i,j,j)\phi_2(i,j,k) \tag{2.2}$$

The distance is given by: $D_{a,b} = \sqrt{2-2SI_{1,2}}$

Perturbation maps were generated for each TOG domain by mutating charged residues to alanine. For each set of electrostatic potentials, cumulative distributions of electrostatic similarity (ESI) were calculated, according to the following expression:

$$ESI = \frac{1}{N} \frac{|\phi_A(i,j,k) - \phi_B(i,j,k)|}{max(|\phi_A(i,j,k), \phi_B(i,j,k)|)} \tag{2.3}$$

For this averaged normalised difference, ϕ_A represents the parent electrostatic potential to which all other potentials are compared (parent), while $\phi_{B,n}$ represents the Nth member of the family of electrostatic potentials for comparison. Surface projections of electrostatic similarity were output to OpenDX files and rendered using UCSF Chimera (version 1.10.1) (Pettersen et al., 2004).

2.4.2 Electrostatic Free Energies of Association

Theoretical alanine scan mutagenesis was performed for all charged residues in each TOG domain, one at a time. Electrostatic free energies of association were calculated for each TOG mutation, according to a thermodynamic cycle that includes association in both a solvent and vacuum reference states:

$$\Delta G_{assoc} = \Delta G_{coul} + \Delta G_{solv} \tag{2.4}$$

The electrostatic free energies of mutants are represented relative to the parent proteins as described:

$$\Delta\Delta G_{binding} = \Delta G_{assoc}^{mutant} - \Delta G_{assoc}^{parent} \tag{2.5}$$

All steps in this cycle were automated through the use of Python scripts and binding plots were generated using the python libraries: pandas and matplotlib.

2.4.3 Normal Mode Analysis

Normal mode analyses were performed using the Bio3d package (version 2.2.-2) (Grant et al., 2006) for each domain. In Bio3d, the low frequency normal modes are calculated using the coarse-graining elastic network model (ENM). The harmonic potential mode used in these analyses was the residue specific force field (sdenm) developed from NMR datasets (Dehouck and Mikhailov, 2013). The correlated and anticorrelated motions between residue pairs of distance structural elements were determined from the combined 200 modes.

2.4.4 Principal Component Analysis

PCA was performed on all Cα positions from individual TOG domains to examine the differences in conformation between structures. Prior to dimensionality reduction, structures

were subjected to iterated rounds of structural superposition to identify the most structurally invariant region. This region was used as the reference frame for the alignment of all crystal structures. From the superposed structures, Cartesian coordinates were extracted to define the elements of a covariance matrix. The covariance matrix was then diagonalised to derive principal components with the eigenvalues providing the associated variances. The TOG structures were then projected on to a subspace defined by principal components with the two largest eigenvalues. All processing steps were undertaken in R using the Bio3d library. PyMol was used for molecular figure generation.

2.5 Summary

Electrostatic analysis successfully identified the invariant regions of TOG domain potential as being located in the tubulin binding loops. Regions outside of these loops showed no ability to perturb the overall field. Clustering of TOGs by their potentials resulted in distinct groupings of TOGs from nearby positions in the array, reflecting their known functional roles. NMA further refined the importance the tubulin binding loops. Loop regions close to interacting with β-tubulin showed greater amplitudes of fluctuation. This might reflect the greater structural diversity sees in the N-termini of TOGs, which could be important for regulating tubulin selectivity.

3

BROWNIAN DYNAMICS OF TOG DOMAIN ASSOCIATION

3.1 Introduction

Fast protein-protein association is at the core of a diverse range of biological processes, many involving enzyme catalysis and inhibition. Many proteins do not exist in isolation but function in the context of multi-component complexes. Association rates often play a critical role in regulating these processes.

Experimentally observed rate constants of protein association span a wide range from 10^3 to 10^9 $M^{-1}s^{-1}$ (Schlosshauer and Baker, 2004) (Fig. 3.1) . In general, the maximum rate of protein association is limited by diffusion and the geometric constraints to the binding site. In purely diffusion-controlled association, proteins typically undergo only fast, local conformational changes between their unbound and bound state. For association rates $<10^6$ $M^{-1}s^{-1}$ the association rate is typically limited by the slow formation and exchange of bonds between the two proteins, and/or by a requirement for conformational change(s). For association rates $>10^6$ $M^{-1}s^{-1}$, diffusion must be accelerated by the presence of intermolecular forces. The dominant long-range force in protein-protein association is provided by electrostatic interactions. This steering force is mediated by highly complementary electrostatic surfaces between the two protein partners.

The rate limiting steps of protein-protein association can be understood by considering the binding process between two proteins, A and B, in the following reaction scheme:

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A \cdot B \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C \tag{3.1}$$

A·B denotes a transient complex, where the two proteins form similar contacts and orientations to the final, native complex C. The formation of the transient complex is governed by both the translational and rotational diffusion of the molecules.

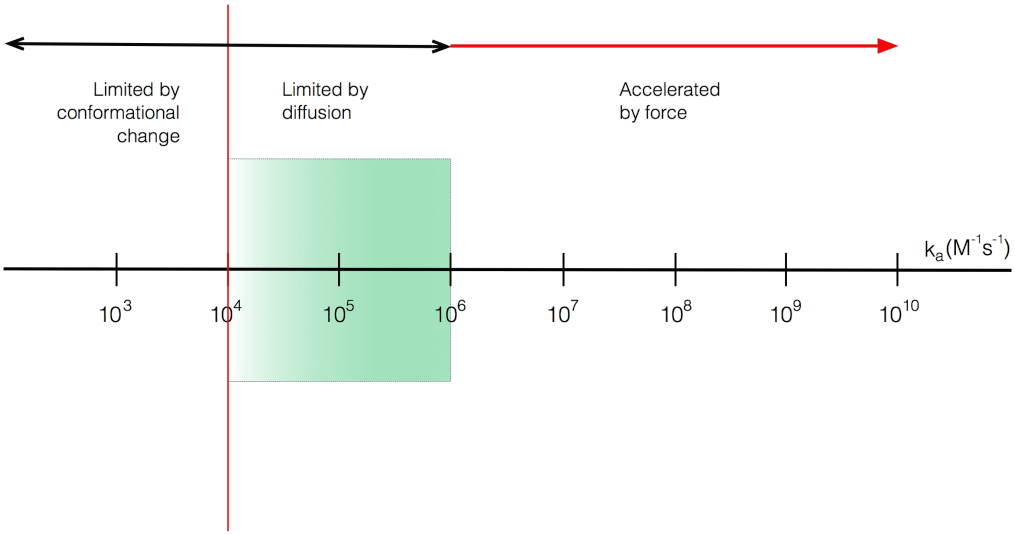


Figure 3.1: The spectrum of association rates. The areas covering conformational and diffusion-limited regimes are demarcated by the red line and the shaded area in green marks the absence of long-range forces. Figure adapted from (Alsallaq and Zhou, 2008)

The second step in the binding process is overcoming the energy barrier from A·B to form the low-energy bound state, C (Fig. 3.2). Complex formation involves the forming of short-range interactions such as hydrogen bonds, removal of water from the binding site, and a range of structural rearrangements to allow the proteins to adapt themselves to fit the final binding interface.

Under the assumptions of the steady-state approximation, the overall rate constants for association and dissociation are:

$$k_a = \frac{k_1 k_2}{k_{-1} + k_2} \tag{3.2}$$

$$k_d = \frac{k_{-1} k_{-2}}{k_{-1} + k_2} \tag{3.3}$$

The association rate is diffusion limited when conformational rearrangement to form the final complex is fast relative to dissociation of the transient complex ($k_2 > k_{-1}$). This leads to a situation where the overall rate constant of association is tightly constrained by the

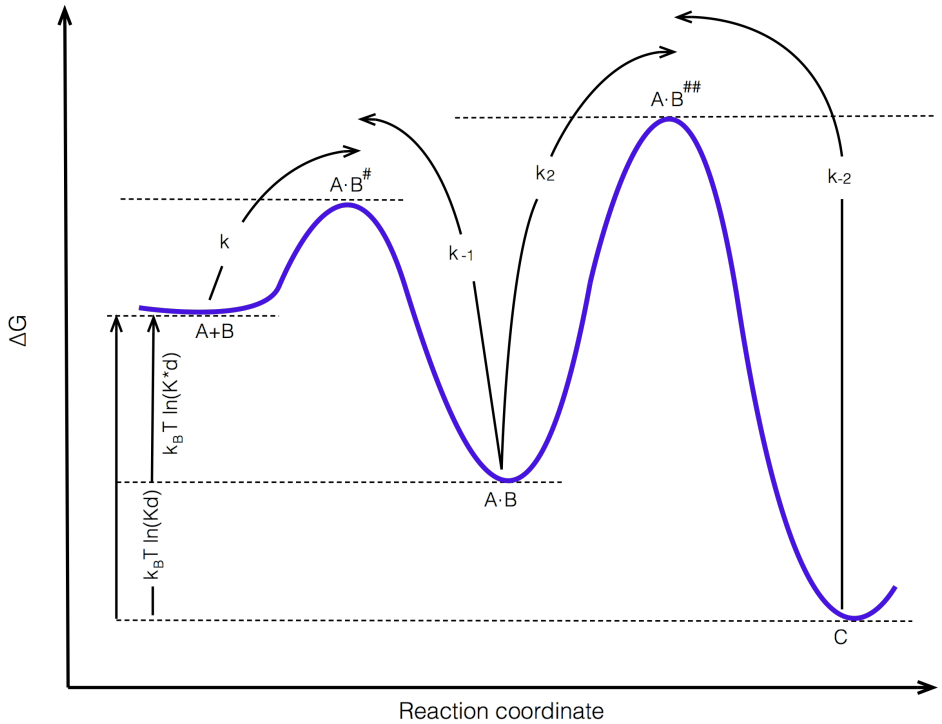


Figure 3.2: Two-step binding model for a bimolecular reaction. The plot shows a description of the energy landscape for the binding of two proteins. The heights of these barriers are related to the kinetic parameters. Protein binding begins with a diffusion-controlled association step of A to B resulting in the formation of a diffusional encounter complex (A·B). The binding energy of this step is related to an equilibrium dissociation constant K*d = k_1/k_{-1}[M]. This transient encounter complex is not fully ordered and solvated but is relatively stable and similar to the final binding state. The second step of the binding process is associated with overcoming an energetic barrier (A·B$^{\#\#}$) to achieve a low-energy bound state (C). Figure adapted from (Romanowska et al., 2015)

on-rate for transient complex formation. In contrast, when complex formation is limited by conformational change, the association rate can be approximated by:

$$k_a \approx (k_1/k_{-1}) \cdot k_2 \tag{3.4}$$

Protein-protein association can be modeled by simulating the translational and rotational Brownian motion of rigid protein subunits. By omitting conformational fluctuations, timescales are achievable that are far beyond that of molecular dynamics, making the calculation of diffusion-controlled rate constants tractable.

3.2 Aims

1. Based on electrostatic steering, predict the likely position of the TOG domain binding site on the tubulin heterodimer.

2. Examine the differences between association rates for different TOG domain, and relate this to their array position in the native protein.

3. Examine the differences in TOG domain binding kinetics for porcine versus cognate tubulin.

4. Determine the residues that are important for modulating association rates using in silico mutagenesis experiments.

5. Build a general description of the TOG domain association pathway and the role of electrostatic guidance.

3.3 Results & Discussion

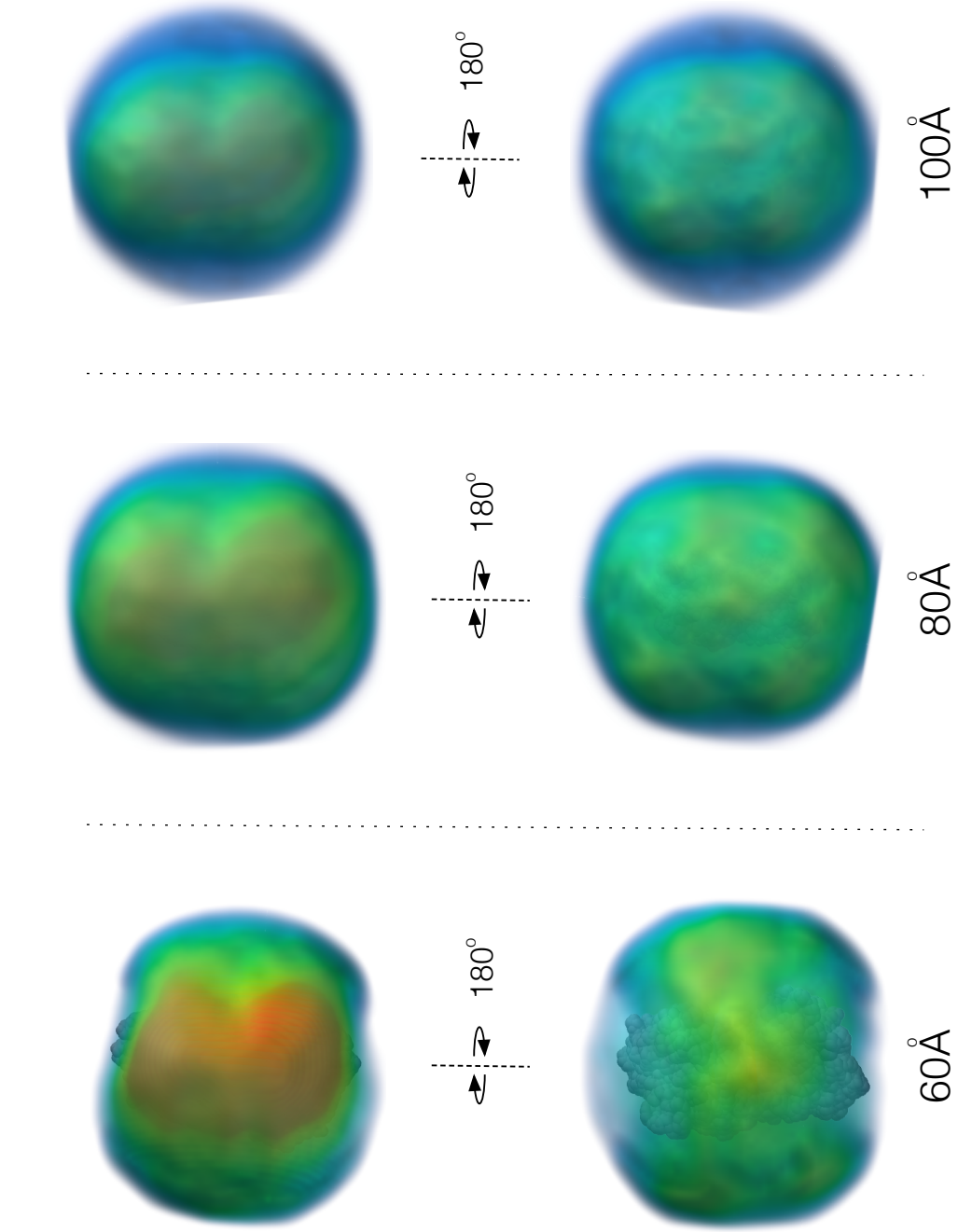
3.3.1 Brownian Dynamics density maps show a clear TOG domain binding site on the tubulin heterodimer

BD trajectories of Stu2 TOG2 diffusing in the electrostatic field of tubulin with and without the presence of the tubulin C-terminal tails were generated at 100 mM monovalent ionic strength to investigate the TOG-tubulin association pathway. These density maps, showing the positions of the TOG domain at every time point in a trajectory, from all trajectories, were generated to show their positions at 65 Å, 85 Å and 100 Å from the centre of the tubulin dimer. Trajectories were generated with no reaction criteria until escape in order to fully sample the electrostatic field (see Methods). Stu2 TOG2 was chosen because of the existence of a high-resolution crystal structure of Stu2 TOG2 bound to tubulin, and extensive biochemical characterisation, both of which support the notion that Stu2 TOG2 binds a single tubulin heterodimer (Ayaz et al., 2012)(Ayaz et al., 2014). A random subset of 20 000 trajectories from a total of 500 000 was used to prepare the density maps (Figure 3.3).

This study was undertaken using the straight, taxol-stabilised, GDP-tubulin structure (PDB: 1JFF) instead of using the tubulin dimer from the TOG-tubulin crystal complex (PDB: 4U3J). The taxol was removed from the structure and energy minimisation and refinement performed (see Methods). The use of this energy-minimised straight structure minimised the potential bias of the complex structure.

Overall, the TOG domain shows a clear preference for a single binding site for a straight tubulin structure. This hotspot straddles the α- and β-tubulin monomers. The hotspot for the binding of TOGs to tubulin without its C-terminal tails is larger (Fig. 3.3A) and consequently less dense in comparison to the tubulin with the tails included (Fig. 3.3B). This implies that at close range, when the TOG domain is close to its final binding position, the inclusion of the tail focuses the TOG domain to a narrower binding site. This trend is reflected through the less dense regions seen outside the binding hotspot. Here the tubulin dimer without tails shows higher levels of density in all regions away from the binding site when compared with the same regions around the tubulin dimer with tails (Figure 3.3B). As the radius of the density maps was increased, the tubulin without tails showed little change in the shape of the binding hotspot, with similar levels of density around the rest of the dimer. In contrast, for the tubulin dimer with tails, the binding hotspot at high radius expanded and began to encapsulate the C-terminal extension of the tubulin, with low density in its surrounding regions.

Initial studies of Stu2p suggested that TOG domain proteins potentially undergo a large conformational transition from an open to closed structure, enwrapping the free tubulin dimer upon its capture (Al-Bassam et al., 2006). This capture model implied that tubulin possesses more than one TOG binding site. More recent crystallographic studies, at higher resolution, have definitively shown that single TOG domains bind a single tubulin heterodimer (Ayaz et al., 2012)(Ayaz et al., 2014). The results in this chapter strongly disfavour models with multiple TOG binding sites on a single tubulin heterodimer, Instead, the density maps show very strong electrostatic guidance controlling the diffusional approach trajectory of TOGs to a single position on the tubulin heterodimer.


A
180°
180°
180°
60Å
80Å
100Å

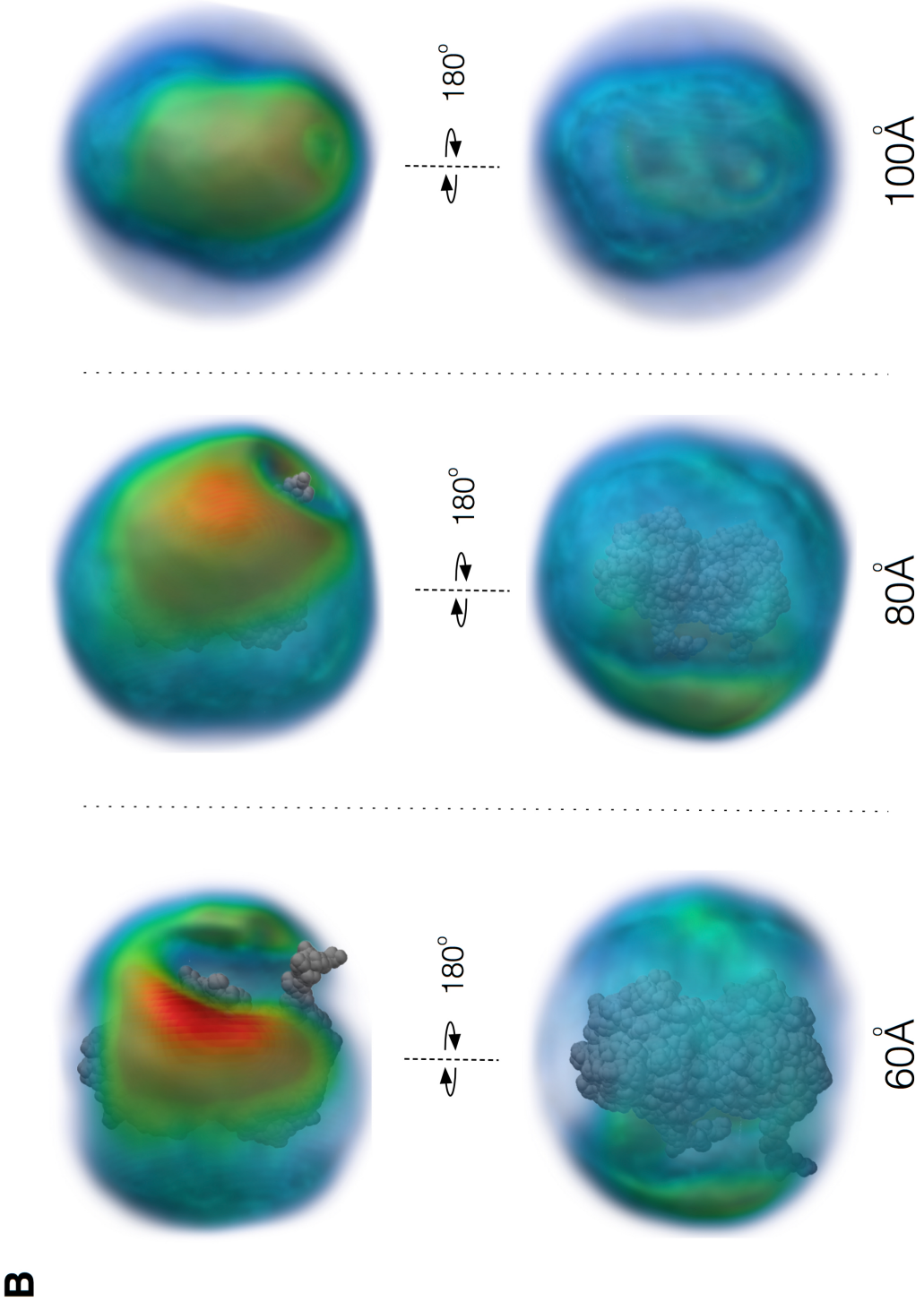


Figure 3.3: Density maps of the TOG domain diffusion pathway around a tubulin dimer (**A**) with no C-terminal tail and (**B**) with a C-terminal tail. The colour change represents the average occupancy value for each position in 3D space, with a low point density being represented by green and high in red.

3.3.2 BD simulations of TOG domain association show species-specific association rates

BD simulations were run for a subset of TOG domains diffusing towards both porcine and the cognate tubulin species for each TOG domain family member. In this context, cognate tubulin refers to the specific tubulin for that species, e.g. Xenopus tubulin for XMAP215. Reaction criteria were selected based on a pilot comparative study on the native protein complexes of Stu2 TOG1 (PDB: 4U3J) and Stu2 TOG2 (PDB: 1FFB) (see Methods). The simulations were tuned to obtain an association rate that was representative of electrostatically-driven diffusion and where the final bound complex showed a low RMSD to the known crystallographic complex. Pairwise residue interactions were calculated from energy minimised structures for each complex using a cut-off distance of 0.5Å. Successful reactions for each BD simulation required the formation of three interactions from the pairwise list with inter-atomic distances of 5Å. The top two mutations from the electrostatic binding analysis of Chapter 2 were then examined to determine the effect of these mutations on the association rate for cognate tubulin.

When association rates were calculated (Fig. 3.4, Tables 3.1, 3.2), it was found that Alp14 TOG1 and TOG2 associate to the cognate tubulin faster than to porcine. Individually, Alp14 TOG1 associated faster than TOG2, suggesting elevated binding to tubulin. For Dis1, TOG1 associated more slowly to *S. pombe* tubulin than porcine tubulin. This pattern is a reversal of that seen for Alp14 TOG1 and may be indicative of different functional roles for these TOG isoforms in *S. pombe* cells. Alp14 is predominantly associated with enhancing MT growth and Dis1 alongside the Ndc80 kinetochore protein has been implicated in MT bundling activities. Dis1 TOG1 associated more slowly than Dis1 TOG2 when binding to its cognate tubulin, whereas both Dis1 TOG1 and Dis 1 TOG2 showed elevated, roughly equal association rates when binding to porcine tubulin. This result was surprising and might reflect a more specialised role for Dis1 in *S. pombe* cells that does not obviously equate to a role in a mammalian system.

Association rates of Zyg9 TOGs to *C. elegans* tubulin varied according to the position of the TOG within the native array, in the order TOG1 >TOG2 > TOG3. Zyg9 TOG1 associated faster to *C. elegans* tubulin than to porcine tubulin, whereas Zyg9 TOG2 associated more quickly to porcine tubulin. Zyg9 TOG3 associated to both tubulin species at similar rates. Zyg9 TOG3 is unusual in that an N-terminal HR binds laterally across the domain. This arrangement differs markedly from that of the TOG1 and TOG2 binding loops and may explain why its association rate is insensitive to the tubulin species.

XMAP215 TOG1 showed a high association rate for Xenopus tubulin while the remaining TOG domains (2-5) showed a similar and slower association rate. Again for porcine tubulin there is no dependence of association rate on the TOG position in the native array. The high association rate for XMAP215 TOG1 is most likely caused by its significantly higher overall negative charge when compared to the other XMAP215 TOG domains. This suggests that the magnitude of the charge plays an important role in the binding of TOG domains to the tubulin heterodimer. Furthermore, the association rate of XMAP215 TOG1 binding to porcine tubulin is significantly less than binding to its cognate tubulin. TOG2 also shows a low association rate binding to porcine tubulin. According to *in vitro* studies of MT polymerisation, XMAP215 TOG1 and TOG2 are functionally equivalent in terms of tubulin binding and polymerisation activity (Widlund et al., 2011). However, these studies were performed using mammalian tubulin instead of the native Xenopus tubulin and the association rates shown in Figure 3.4 show clear differences in TOG binding to the different tubulin species.

Overall, the data show a marked difference between the association rates calculated for cognate and mammalian tubulin. Although mammalian tubulin is the most commonly used form of tubulin for biochemical assays due to its ease of purification, these results show that this could have a significant effect on putative function of TOG domains.

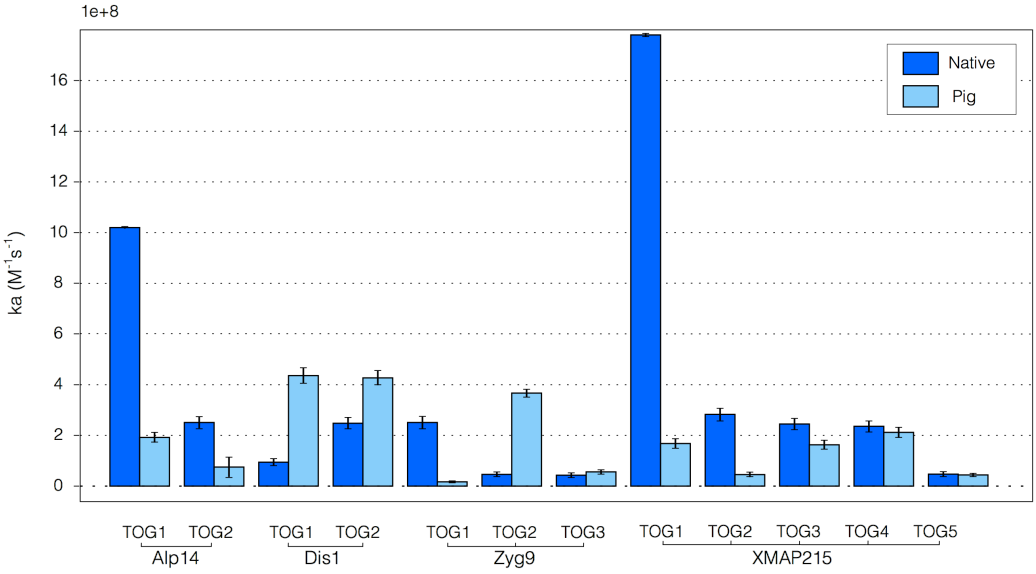


Figure 3.4: TOG domain family association rates from BD simulations. Each TOG domain was simulated with cognate or porcine tubulin. The rates for native tubulin are shown in dark blue and those for pig tubulin in light blue. The rates are shown with 95% confidence intervals.

	TOG1	TOG2	TOG3	TOG4	TOG5
Alp14	1.92e+7 $\pm$ 0.20	7.52e+6 $\pm$ 0.39			
Dis1	4.35e+7 $\pm$ 0.31	4.27e+7 $\pm$ 0.29			
Zyg9	1.69e+6 $\pm$ 0.37	3.67e+7 $\pm$ 0.15	5.62e+6 $\pm$ 0.81		
XMAP215	1.68e+7 $\pm$ 0.19	4.55e+6 $\pm$ 0.97	1.63e+7 $\pm$ 0.18	2.12e+7 $\pm$ 0.2	4.41e+6 $\pm$ 0.63

Table 3.1: Porcine association rate constants, with 95% confidence intervals.

	TOG1	TOG2	TOG3	TOG4	TOG5
Alp14	1.02e+8 $\pm$ 0.04	2.51e+7 $\pm$ 0.23			
Dis1	9.43e+6 $\pm$ 0.14	2.48e+7 $\pm$ 0.23			
Zyg9	2.51e+7 $\pm$ 0.24	4.62e+6 $\pm$ 0.97	4.30e+6 $\pm$ 0.93		
XMAP215	1.78e+8 $\pm$ 0.06	2.83e+7 $\pm$ 0.24	2.45e+7 $\pm$ 0.22	2.36e+7 $\pm$ 0.21	4.72e+6 $\pm$ 0.98

Table 3.2: Cognate association rate constants, with 95% confidence intervals.

3.3.3 Alanine mutations of electrostatically-favourable residues within TOG domains reduced the association rate to tubulin

Electrostatically-favourable residues used in these simulations were identified in Chapter 2. These charged residues were mutated to alanine and TOG domain diffusion was simulated as before to calculate the effect on the association rate to cognate tubulin species (Fig. 3.5, Table 3.3).

Alanine mutations for each native complex (See Table 3) had the effect of reducing the association rate constant in the majority of cases. For those that did not reduce the rate, there was a negligible effect. This effect is likely due to the net charge between the tubulin dimer and the TOG domain for highly charged TOG domains, a single point mutation is unlikely to perturb the overall electrostatic field between the two diffusing proteins. This result shows that those mutations which have the greatest effect on electrostatic binding energy, also modify the association rate constant for TOG domains to the greatest extent.

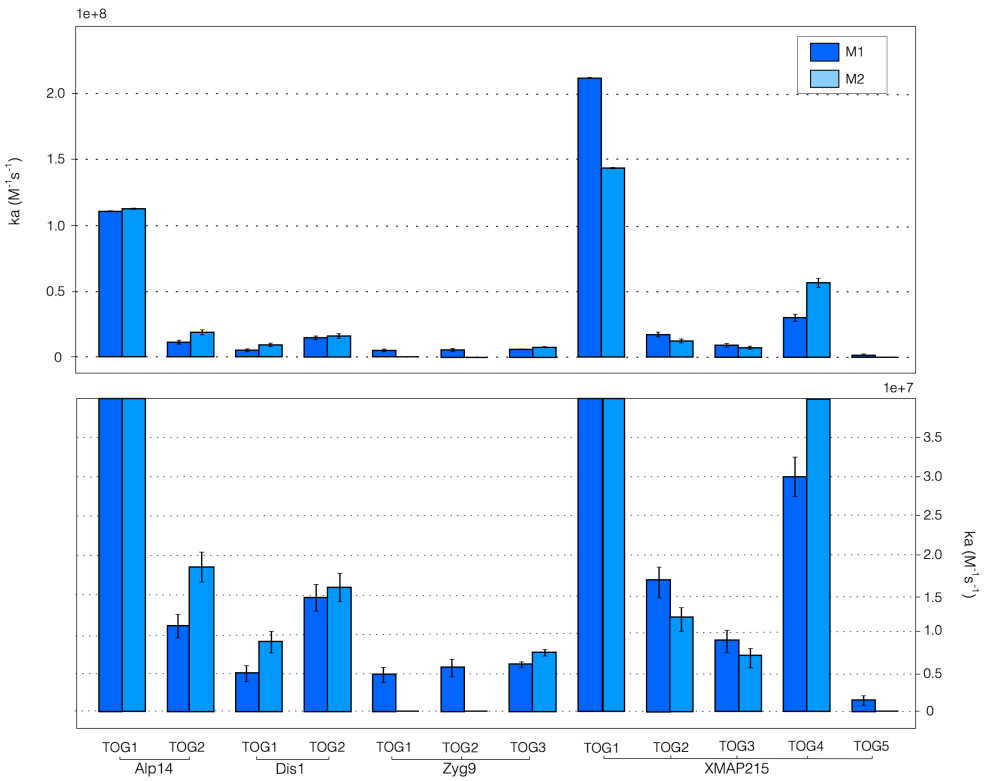


Figure 3.5: Native association rates for mutations for each TOG domain from each species. The two mutations with greatest influence on electrostatic binding analysis where selected with a cutoff of >5kJ/mol. The association rates are presented with 95% confidence intervals. The upper panel shows the overall distribution of rates while the lower panel is scaled to identify smaller values.

		TOG1	TOG2	TOG3	TOG4	TOG5
Alp14	M1	1.10e+8 $\pm$ 0.05	1.10e+7 $\pm$ 0.15			
	M2	1.12e+8 $\pm$ 0.05	1.85e+7 $\pm$ 0.19			
Dis1	M1	4.97e+6 $\pm$ 0.10	1.46e+7 $\pm$ 0.17			
	M2	8.98e+6 $\pm$ 0.14	1.59e+7 $\pm$ 0.18			
Zyg9	M1	4.80e+6 $\pm$ 0.97	5.70e+6 $\pm$ 0.11	6.1e+5 $\pm$ 0.35		
	M2			7.6e+5 $\pm$ 0.39		
XMAP215	M1	2.11e+8 $\pm$ 0.06	1.69e+7 $\pm$ 0.19	9.12e+6 $\pm$ 0.14	3.0e+7 $\pm$ 0.25	1.5e+5 $\pm$ 0.06
	M2	1.43e+8 $\pm$ 0.05	1.21e+7 $\pm$ 0.15	7.17e+6 $\pm$ 0.12	5.64e+7 $\pm$ 0.34	

Table 3.3: Association rates of TOG domains to cognate tubulin after alanine mutation of key charged residues, with 95% confidence intervals.

		TOG1	TOG2	TOG3	TOG4	TOG5
Alp14	M1	R103A	R87A			
	M2	K93A	R173A			
Dis1	M1	R21A	K130A			
	M2	R271A	R172A			
Zyg9	M1	R177A	R196A	K172A		
	M2			K133A		
XMAP215	M1	K176A	R182A	K115A	R103A	R62A
	M2	R15A	R22A	R42A	E173A	R116A

Table 3.4: List of charged residues mutated for each TOG domain.

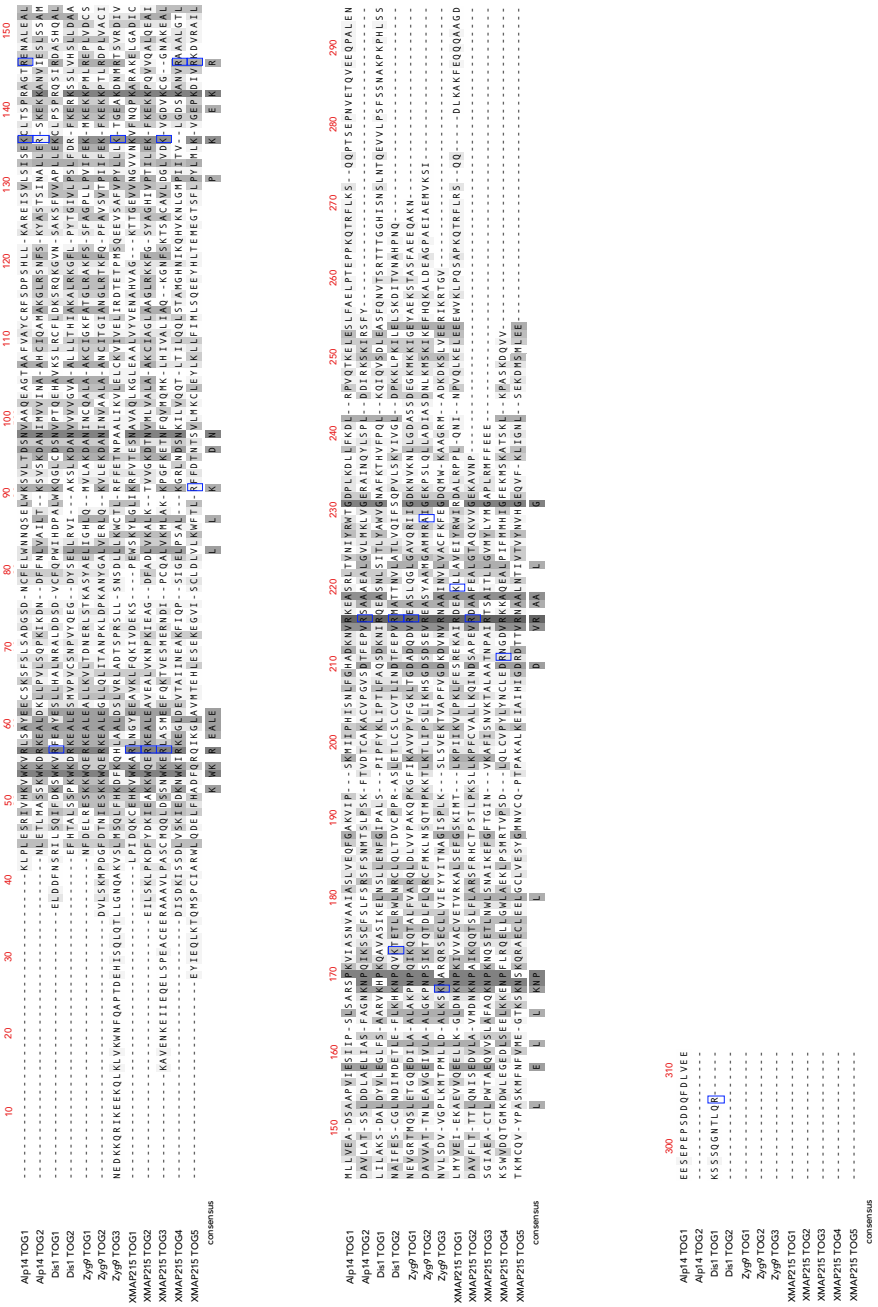


Figure 3.6: Multiple sequence alignment of 12 TOG-domains. The alignment is coloured by residue conservation, from white for low identity (0-9%) to dark grey for high identity (90-100%). Predominately, these mutations cluster within regions of high conservation and are found in residues at the exposed end of each α-helix, close to the start of the tubulin binding loops.

3.3.4 The translational and orientational steering of TOG domains by long-range electrostatics is dependent on species and position within the native TOG array

The trajectories from BD simulations of TOG-tubulin encounters were analysed in order to clarify the nature of their encounter state. As explained in the methods section, the occupancy and conformational space was computed based on the following coordinate space: one set of 3D coordinates for translation and one set of rotational coordinates. The translation and orientation maps are both displayed as a function of centre-to-centre position between each protein

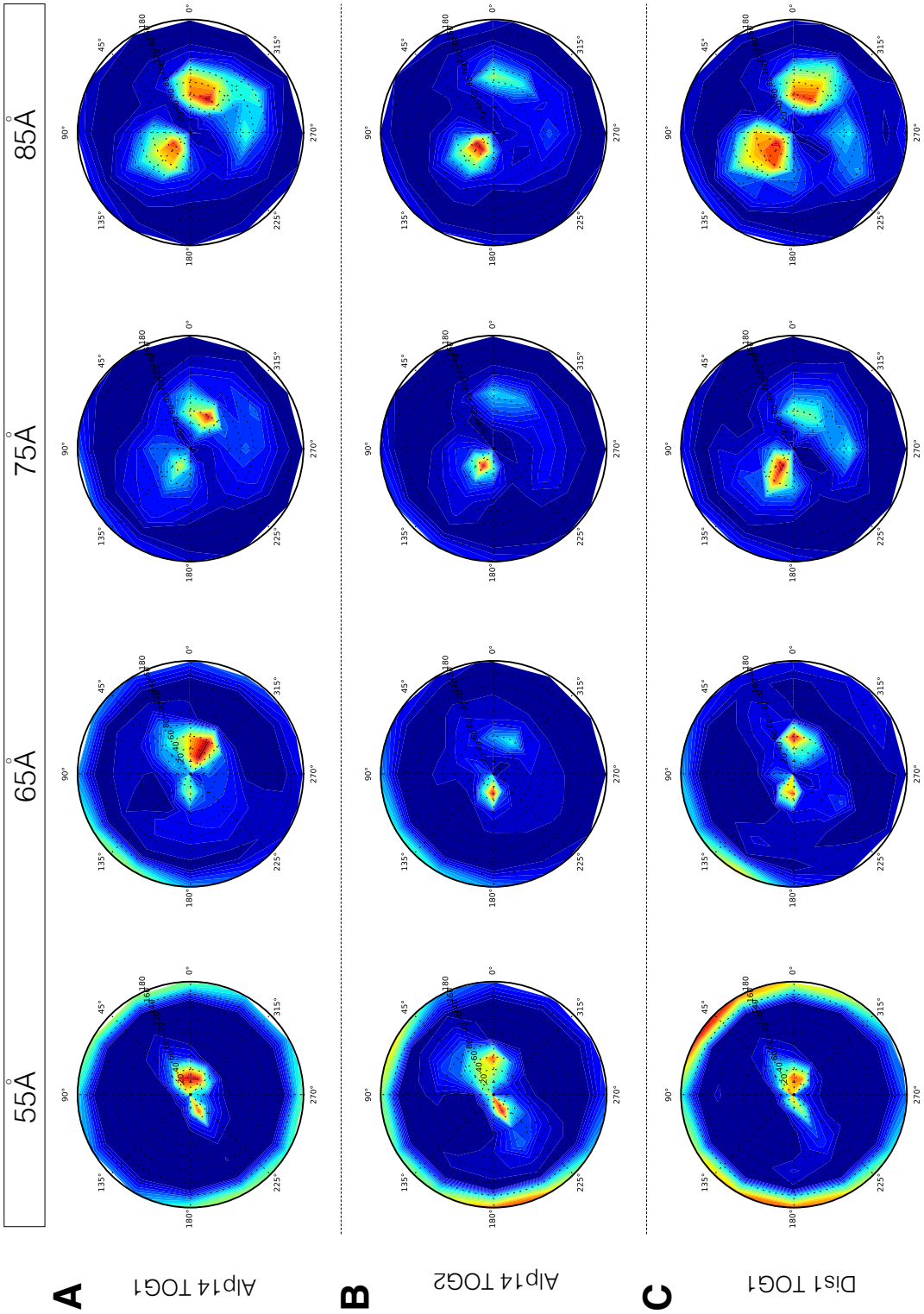
Overall, for short distances from the binding site the highest densities are found in the positions closest to the bound state, while at increasing distance the occupancy cloud becomes more anisotropic with respect to both θ and ϕ angles. This pattern occurs for both translation and orientation maps (Figs. 3.7 & 3.8). This is unsurprising as at higher contact distances, the protein experiences weaker long-range electrostatic forces and fewer geometric constraints that could restrict translational and rotational movement.

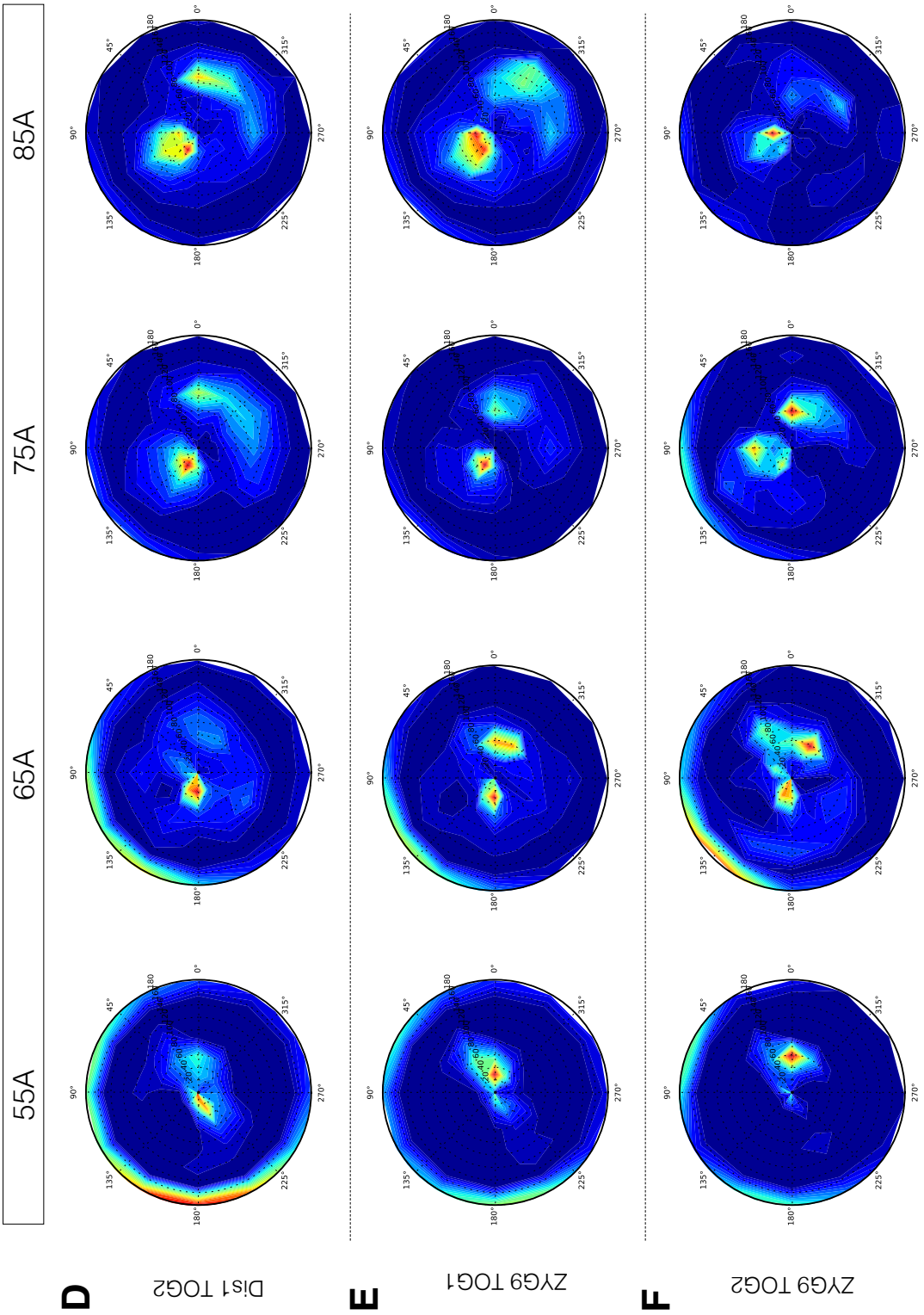
Translation

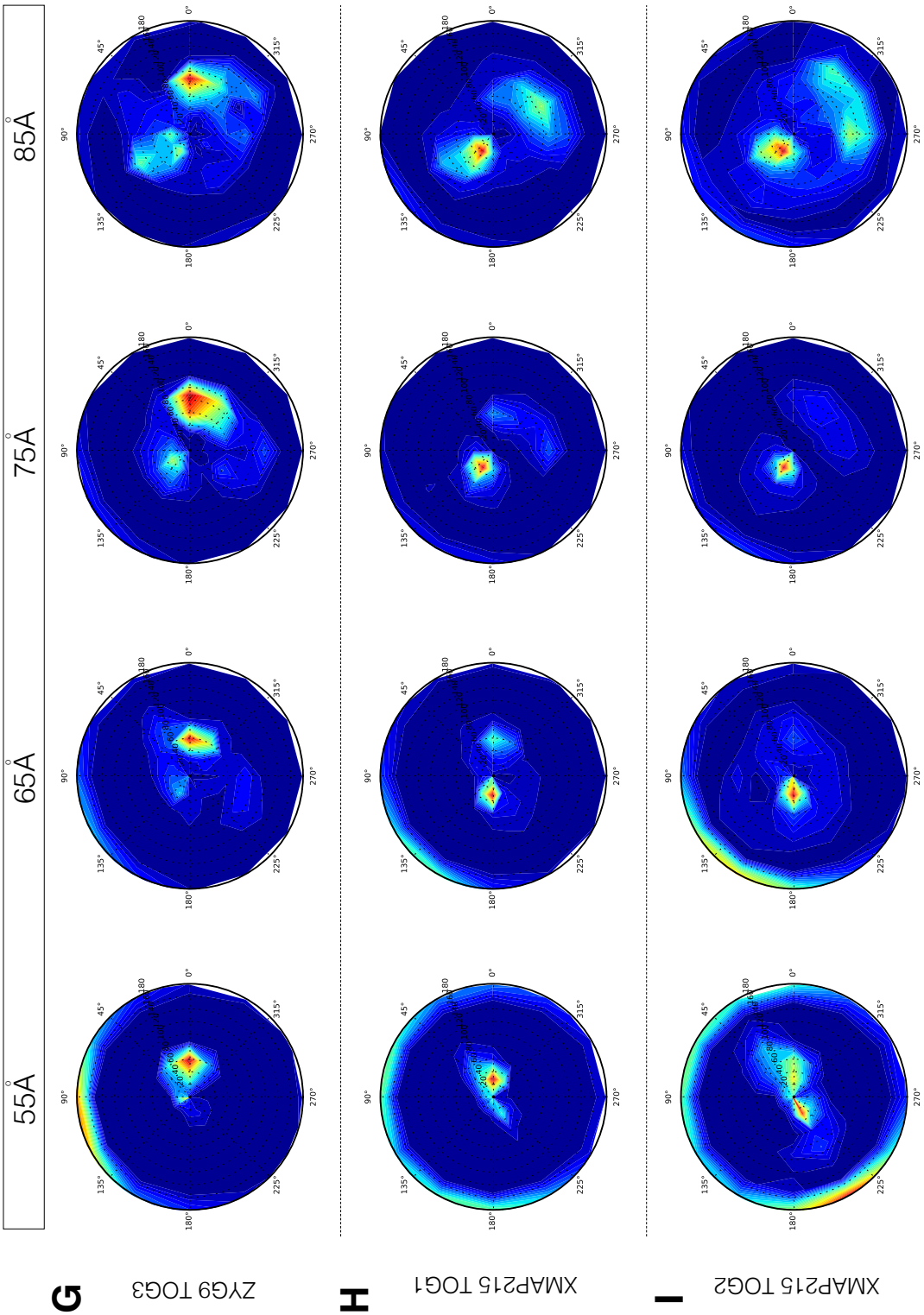
The association rate of Alp14 TOG1 is significantly faster than Alp14 TOG2, and this is reflected in the translation maps (Fig. 3.7A & 3.7B). Alp14 TOG1 shows a hotspot close to the binding site that extends outwards over 5-10 Å covering a small range in θ. In contrast, Alp14 TOG2 shows a similar hotspot at close range but with lower density, which dissipates faster at higher distances. The Alp14 TOG2 displays its highest level of translational density at a site distal to the binding site. This hotspot persists over many angstroms unlike the hotspot seen at the binding interface. The patterns of translational diffusion for Dis1 TOG1 are similar to its homolog Alp14 TOG1. However, the maps for Dis1 also show the presence of high levels of density at elevations close to 180°. This could explain the lower association rates seen with these TOGs, as at close range (55 Å) there is a high proportion of protein on the opposite side of the binding site of the tubulin dimer.

TOG1, TOG2 and TOG3 of Zyg9 show similar patterns of translational diffusion. However, TOG1 has an increased association rate reflected in the faster reduction in ϕ as the distance between the TOG and the tubulin dimer is reduced, from 85 Å to 55 Å.

For XMAP215, TOG1 has the highest association rate and shows a tight hotspot close to the binding interface. TOG2-4 have similar patterns of translational diffusion, albeit with differing densities, and this is reflected in their association rates which are consistent with each other. TOG5 is unusual and shows hotspots at various locations. By 65 Å, the protein has begun to congress at the binding site, but the density is still spread over a wide range of angles around the equator of the simulation sphere.


55Å
65Å
75Å
85Å
A
Alp14 TOG1
B
Alp14 TOG2
C
Dis1 TOG1


55A
65A
75A
85A
D
Dis1 TOG2
E
ZYG9 TOG1
F
ZYG9 TOG2


55Å
65Å
75Å
85Å
G
ZYG9 TOG3
H
XMAP215 TOG1
I
XMAP215 TOG2

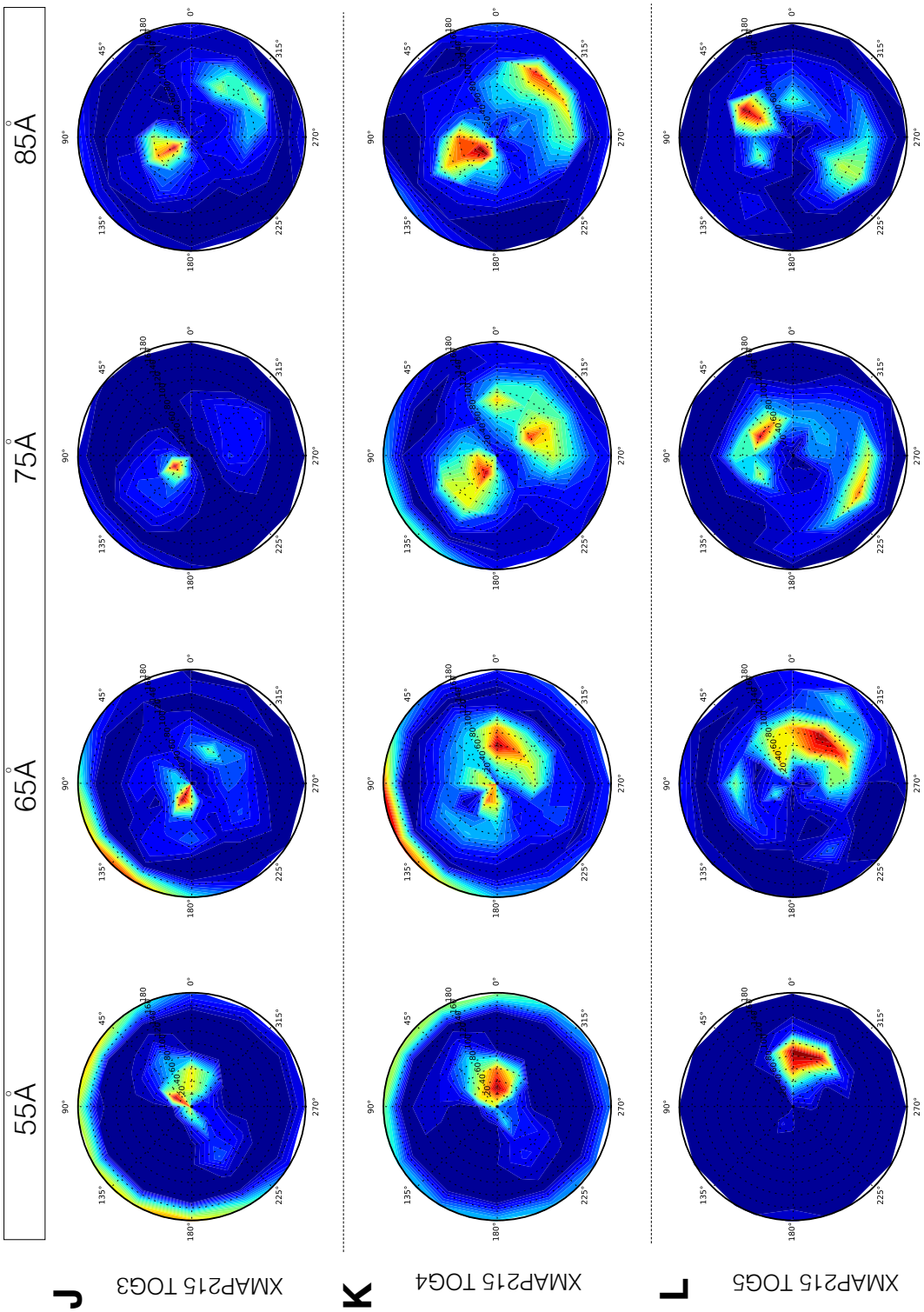


Figure 3.7: Occupancy maps of associated trajectories as a function of contact distance for translational coordinates. The scale is the minimum and maximum number of angles for each distance. Red shows high occupancy and blue shows low occupancy. The angles are defined as follows: θ is the azimuthal angle in the xy-plane from the x-axis with $0 \leq \theta \leq 2\pi$, ϕ is the zenith angle from the positive z-axis with $0 \leq \phi \leq \pi$.

Rotation

Rotational density was assessed separately to translational density. Alp14 TOG1 exhibits a small region of high rotational density close to the binding site at a radius of 55-65 Å, which dissipates at higher contact distance. The shape of the overall hotspot at close range indicates restriction the rotational movement of the TOG domain on the final approach to the binding site. In contrast, Alp14 TOG2 also possesses a hotspot at close range but this density is more distributed in theta and its range is shorter than Alp14 TOG1. This might explain why Alp14 TOG2 has a simulated association rate constant that is an order of magnitude slower than Alp14 TOG1.

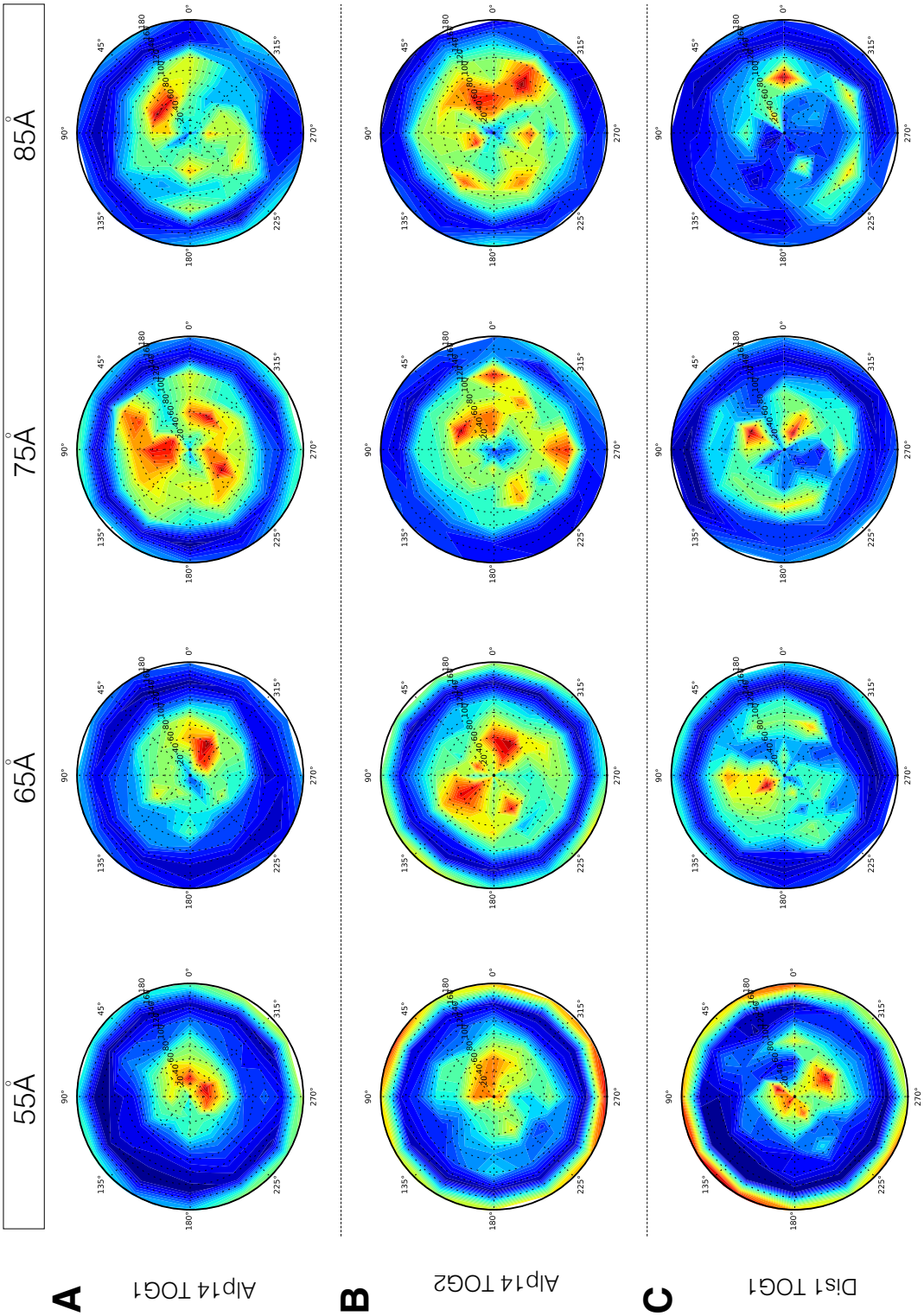
Dis1 TOG1 and Dis1 TOG2 share a high degree of rotational freedom at distances greater than 55 Å from the TOG binding site, with no clear patterns of correlation. Surprisingly, Dis1 TOG1 showed a very disperse rotational hotspot when close to its binding site that spans 0-60° of θ. By contrast Dis1 TOG1 domain spends most of its time rotated by 20° from its binding site. The ability of Dis1 TOG2 to restrict rotational movement over the binding site may account for its increased association constant.

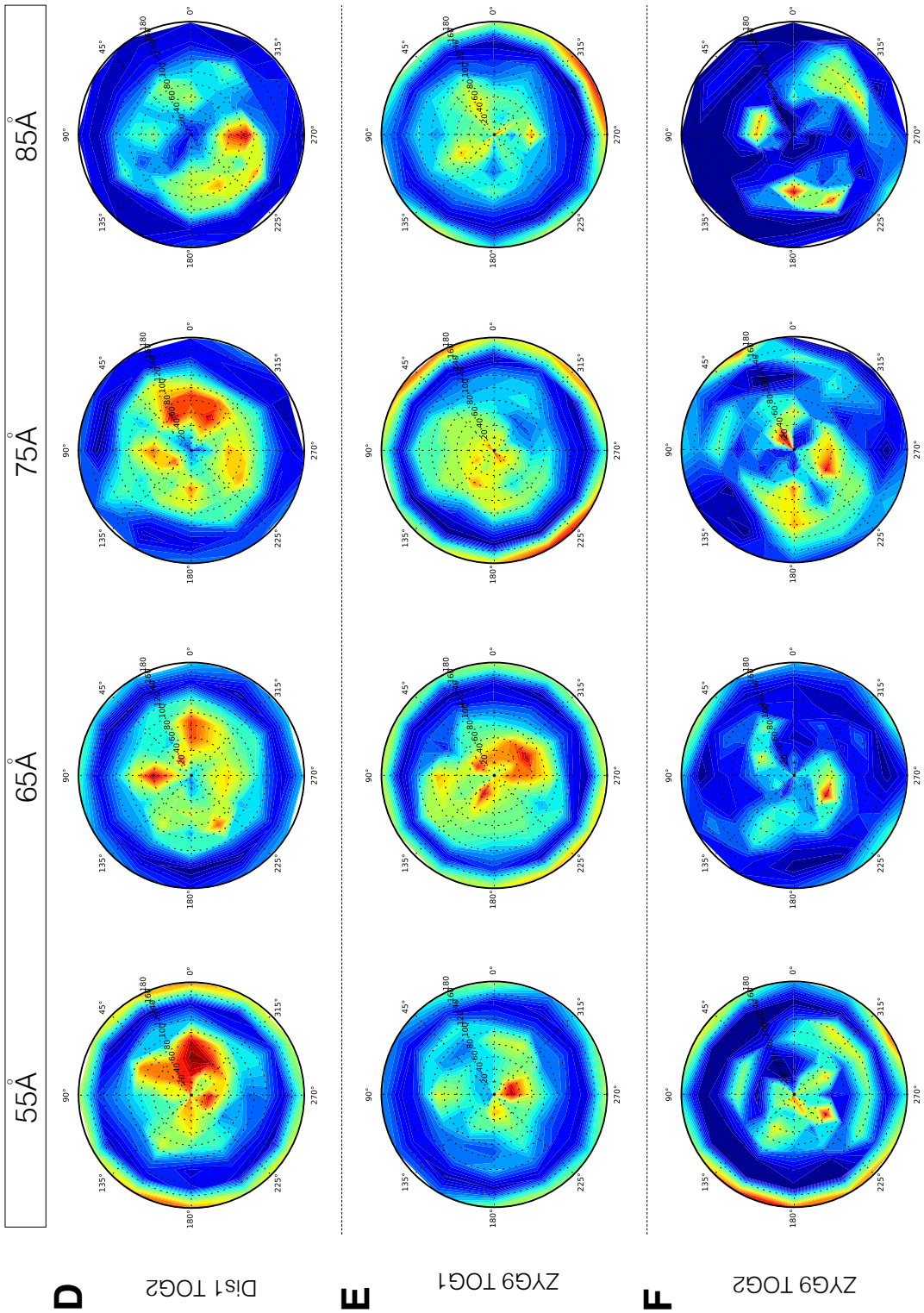
Zyg9 TOG1 showed rotational steering at 65 Å which dissipated at closer range. Zyg9 TOG2 showed little rotational restriction at all distances. Zyg9 TOG3 at 55-65 Å shows a clear rotational hotspot over the binding site that extends over 80° of θ at close range. The highest rate constant from simulations was found for Zyg9 TOG1. Although this TOG domain shows little rotational restriction at close range compared to Zyg9 TOG3, it does possess high levels of rotational conservation in an arc which spans over 90° at 65 Å. Zyg9 TOG1 also shows high levels of rotational restriction at increasing distance that is not present in the other ZYG9 TOG domains. Consistent with this, TOG1 shows a higher association rate constant.

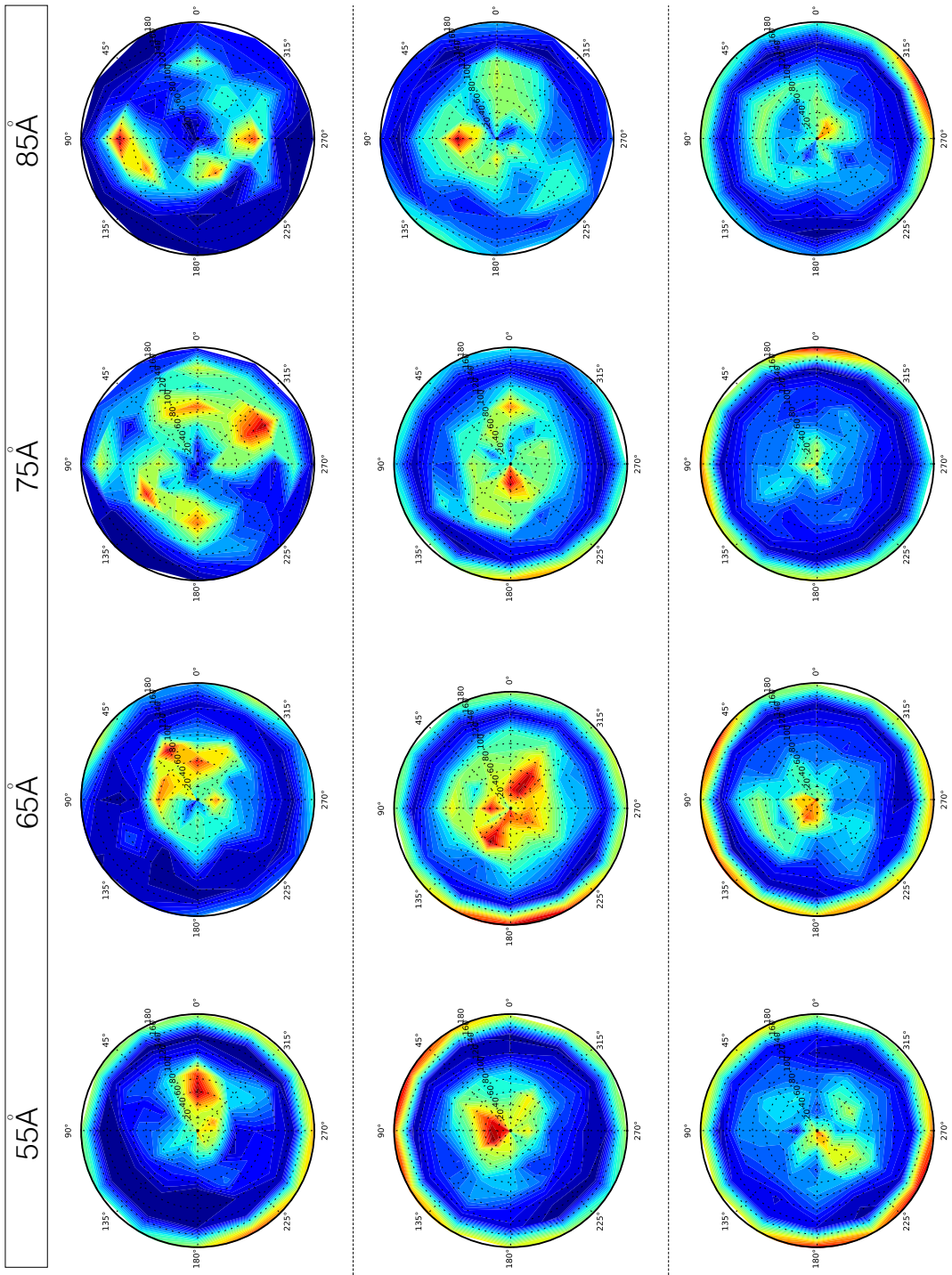
XMAP215 TOG1 displayed the fastest association rate constant. It has a hotspot over the binding site in the 55-65 Å range. This pattern is not so obvious for TOG2-TOG4 and may be a reason for their reduced association rate constants. Interestingly, TOG5 is anisotropic at all distances and its rate constant is the slowest of all domains.

In summary, with decreasing distance from the binding site occupancy clouds become more anisotropic with respect to the interfacial normal. Overall, it is thus very clear from these maps that there is substantial electrostatic steering of the TOG domains towards the tubulin before binding. Even at high center-to-center distance, many of the TOG domains are partially orientated and pointing roughly towards the binding interface. Translational

electrostatic steering extends over many angstroms. By contrast rotational restriction tends to occur close to the final binding site and may help to rotate the TOG domain into the correct position in the final moments before binding. There is a strong positive correlation between the levels of translational and rotational electrostatic steering and the calculated association rates.


55Å
65Å
75Å
85Å
A
Alp14 TOG1
B
Alp14 TOG2
C
Dis1 TOG1


55Å
65Å
75Å
85Å
D
Dis1 TOG2
E
Zyg9 TOG1
F
Zyg9 TOG2


55Å
65Å
75Å
85Å
G
ZYG9 TOG3
H
XMAP215 TOG1
I
XMAP215 TOG2

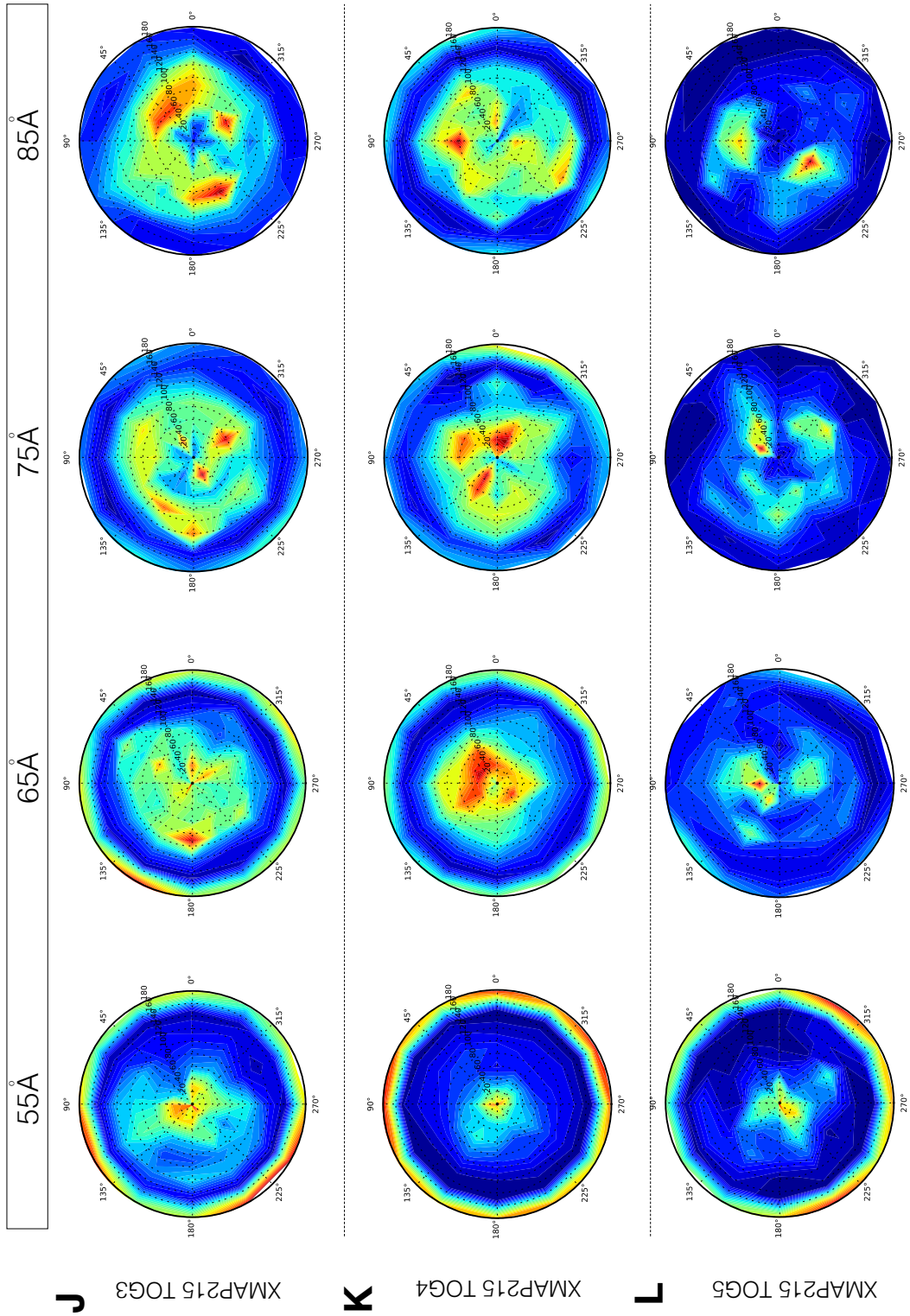


Figure 3.8: The computed occupancy maps of associated trajectories as a function of contact distance for the orientational coordinates. The scale is the minimum and maximum number of angles for each distance. Red shows high occupancy and blue shows low occupancy. The angles are defined as follows: θ is the azimuthal angle in the xy-plane from the x-axis with $0 \leq \theta \leq 2\pi$, ϕ is the zenith angle from the positive z-axis with $0 \leq \phi \leq \pi$.

3.3.5 The TransComp method shows the association-rate of TOG domains to tubulin is electrostatically driven

Transient complex theory is based on the analysis of the interaction energy landscape of associating proteins. The theory uses the association constant of the transient complex (k_1) as a close approximation for the overall association rate constant (k_a), given the difficulty of calculating the association rate of complex formation (k_2). Consequently, the transient complex must be placed close enough to the native complex in order that the resulting k_1 is a useful approximation for k_a.

The native complex sits in a deep well in the interaction energy landscape. The optimal definition for the transient complex ensemble is at the outer boundary of the native-complex energy well (Fig. 3.9).

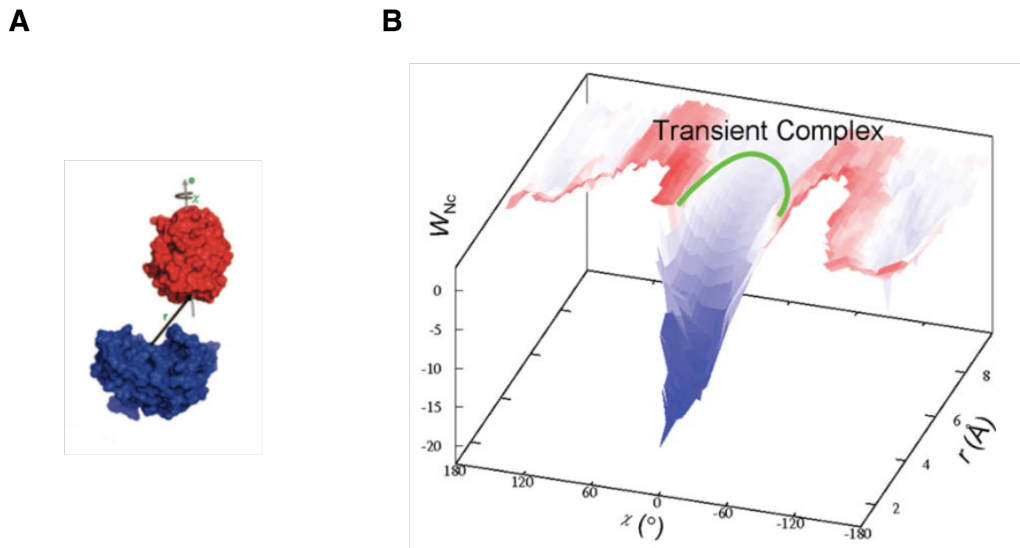


Figure 3.9: Location of the transient complex within the interaction energy landscape. (**A**) The binding proteins are described in terms of translational (r) and rotational (e, χ) coordinates. (**B**) The transient-complex ensemble is indicated by the green ring. The complex is identified where σ_x undergoes a shape increase with decreasing N_c. Figures was adapted from (Alsallaq and Zhou, 2008)

The outer boundary of the native complex energy well was found to coincide with a reduction in translational and rotational freedom. The short-range interaction energy, which stabilises the native complex, was modeled by the number of contacts (N_c) formed and the translational-rotational freedom measured by $\sigma_x(N_c)$, the standard deviation of the rotational angle χ in configurations at a given contact level.

The association rate constant is predicted as:

$$k_a = k_{a0}\, exp(\frac{-\Delta G^*_{el}}{k_b T}) \tag{3.5}$$

where k_{a0} is the basal rate constant at which the two proteins reach the transient complex by free diffusion, and the Boltzmann factor captures the rate enhancement by inter-protein electrostatic attraction. The basal rate captures all the configurations of short-range interactions and configurational entropy such as differences in the sizes and shapes of the protein interface.

The reaction criteria for the prior simulations were quite restrictive in order to position the TOG domain in an orientation that is representative of the domain's final bound state, rather than that of a transition complex. For this investigation, the transition complex method (Alsallaq and Zhou, 2007a)(Alsallaq and Zhou, 2007b) was applied to calculate an association rate for each cognate complex. This method uses a more rigorous definition of the reaction criteria based upon transition complex theory. A basal association rate is calculated using BD simulations in the absence of electrostatic force together with its electrostatic potential calculated from an ensemble of TOG domain configurations. The basal rate and electrostatic potential are combined to predict changes in association rate constant. The method is advantageous in removing the potential bias of user-defined reaction criteria.

Upon application of the TransComp method, the fungal species showed similar association rates for TOG1 and TOG2 with a slightly elevated rate for Dis1 TOG1 (Fig. 3.11, Table 3.5). Zyg9 shows asymmetry in its association rates, with TOG3 showing a higher association rate than TOG2. XMAP215 TOG1 showed a significantly higher association rate than all other TOGs which correlates with its high negative charge. The remaining TOG domains, with the exception of TOG5, show similar association rates. TOG5 associates poorly to tubulin, which is consistent with its undetermined role in MT polymerisation.

Mutations were undertaken using the TransComp methodology for a single Alp14 TOG1 domain. Alanine mutations were carried out on charged residues within 10 Å of the tubulin binding loops. This was done to replicate the investigation in Figure 3.5 to see if similar inhibition occurred. Only a single TOG domain was chosen, with a reduced number of mutations, due to the computationally-intensive nature of these calculations. The alanine mutations showed that the association rate was most prominently inhibited by mutating charged residues in the loop regions and residues flanking the α-helix (Fig. 3.11).

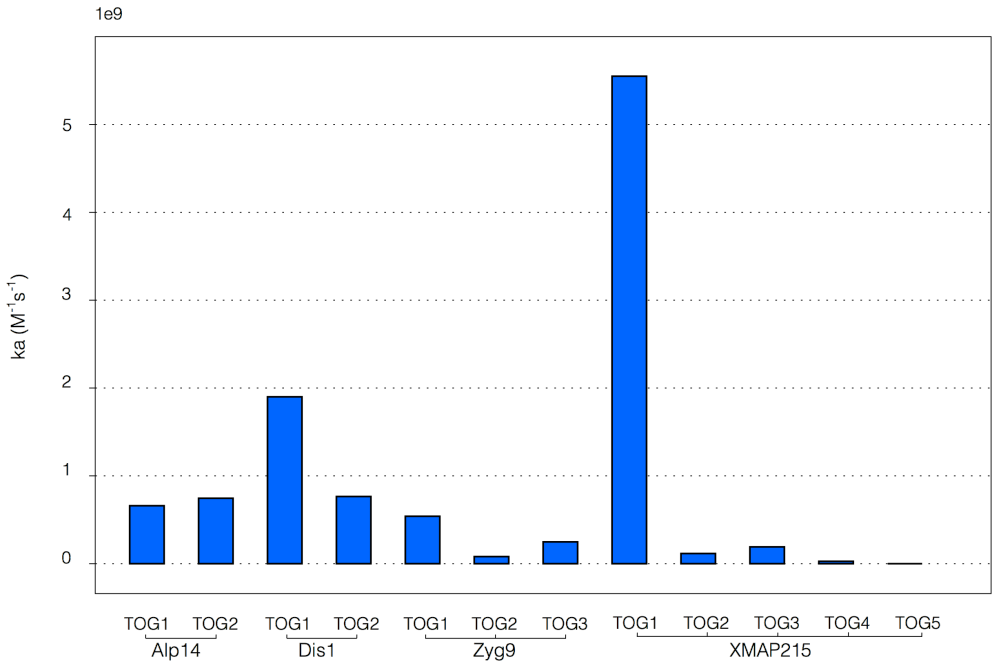


Figure 3.10: Predicted association rates for TOG family members.

	TOG1	TOG2	TOG3	TOG4	TOG5
Alp14	1.30e+09	1.49e+09			
Dis1	3.80e+09	1.53e+09			
Zyg9	1.08e+09	1.62e+08	4.97e+08		
XMAP215	1.11e+10	2.32e+08	2.83e+08	5.46e+07	5.53e-01

Table 3.5: Table of association rates predict using the TransComp method.

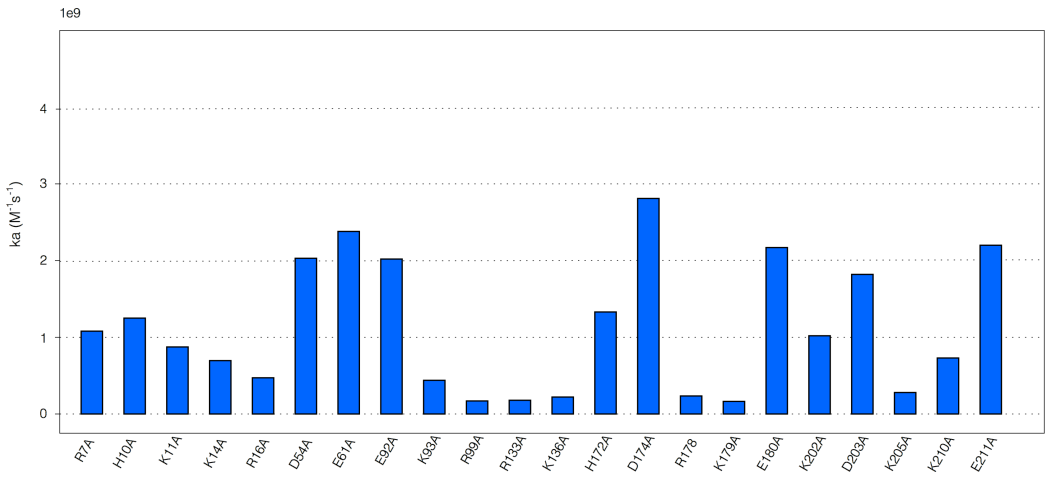


Figure 3.11: Predicted association rates for selected alanine mutations in Alp14 TOG1.

Mutated Residue	Association Rate
R7A	1.08e+09
H10A	1.25e+09
K11A	8.74e+08
K14A	6.96e+08
R16A	4.72e+08
D54A	2.03e+09
E61A	2.38e+09
E92A	2.02e+09
K93A	4.39e+08
R99A	1.70e+08
R133A	1.79e+08
R136A	2.21e+08
H172A	1.33e+09
D174A	2.81e+09
R178A	2.35e+08
K179A	1.64e+08
E180A	2.17e+09
K202A	1.02e+09
K203A	1.82e+09
K205A	2.80e+08
E210A	7.30e+08
E211A	2.20e+09

Table 3.6: Table of predicted association rates for alanine point mutations using the TransComp method.

3.3.6 Protein-protein docking shows an energetically favourable binding position that correlates with the crystallographic TOG-array position

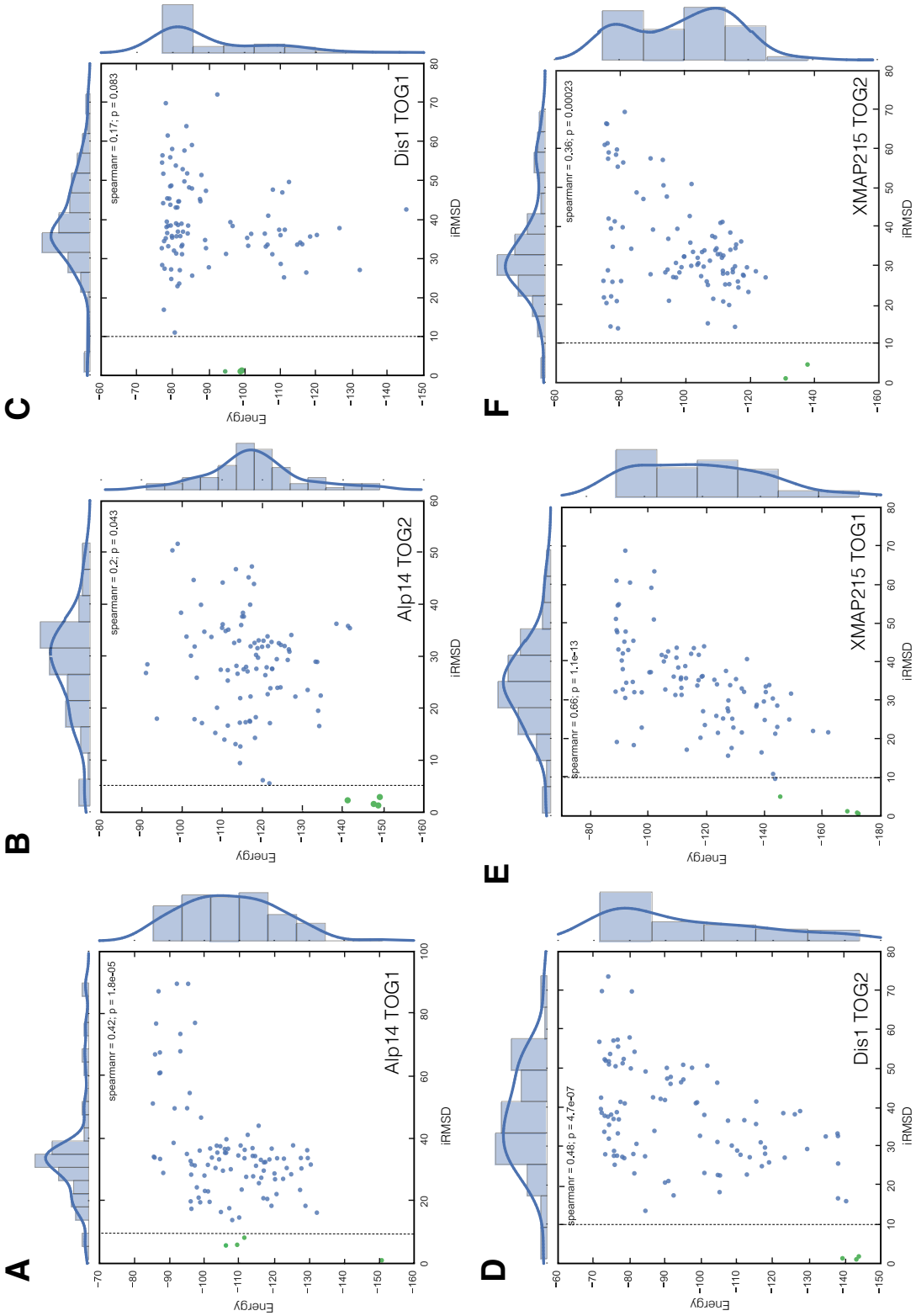
The top 100 structures ranked by their energy score were extracted from 10,000 docking runs. A single distance restraint was applied to improve the docking results. This consisted of a single residue constraint in the conserved tryptophan residue of the first HR of each TOG domain.

The docking results show a clear relationship between the Spearman rank coefficient and the energy score, which in turn relates back to the simulated association rate (Fig. 3.12, Table 3.7). This relationship shows that higher values of the correlation coefficient are associated with a higher on-rate. This relationship holds within species. For instance, the correlation coefficient identifies the difference in association rate between Alp14 TOGs and Dis1 TOGs. Unlike Alp14, Dis1 TOG2 has a higher association rate than its TOG1 domain, and this reversal in association rate in comparison to Alp14 is captured by the correlation coefficient. This trend is further seen in Zyg9 and in XMAP215. Specifically, the correlation coefficient identifies that XMAP215 TOG5 is electrostatically unfavourable when it comes to binding a tubulin dimer, and unsurprisingly shows the lowest rate within the XMAP215 TOG array.

Using this relatively simple method, without defining binding criteria, the results show that TOG domains can bind tubulin in near-native conformation, using an electrostatics and desolvation scoring function alone. Results using pyDock thus further support the hypothesis that electrostatics are critical determinants for the binding kinetics of TOGs to tubulins.

	Spearman Rank	**p-value**	**BD k_a**
Alp14 TOG1	0.42	1.80e-5	1.02e+08
Alp14 TOG2	0.2	0.043	2.51e+07
Dis1 TOG1	0.17	0.083	9.43e+06
Dis1 TOG2	0.48	4.70e-07	2.48e+07
Zyg9 TOG1	0.66	1.10e-13	1.78e+08
Zyg9 TOG2	0.36	0.00023	2.83e+07
Zyg9 TOG3	0.5	1.50e-07	2.36e+07
XMAP215 TOG1	0.45	2.60e-06	2.36e+07
XMAP215 TOG2	0.17	0.099	4.72e+08
XMAP215 TOG3	0.37	0.0017	2.51e+07
XMAP215 TOG4	0.35	0.00041	4.62e+06
XMAP215 TOG5	0.29	0.003	4.30e+06

Table 3.7: Correlations between PyDock energy score and iRMSD.


A
Alp14 TOG1
spearmanr = 0.42; p = 1.8e-05
B
Alp14 TOG2
spearmanr = 0.2; p = 0.043
C
Dis1 TOG1
spearmanr = 0.17; p = 0.083
D
Dis1 TOG2
spearmanr = 0.48; p = 4.7e-07
E
XMAP215 TOG1
spearmanr = 0.66; p = 1.1e-13
F
XMAP215 TOG2
spearmanr = 0.36; p = 0.00023
Energy
iRMSD

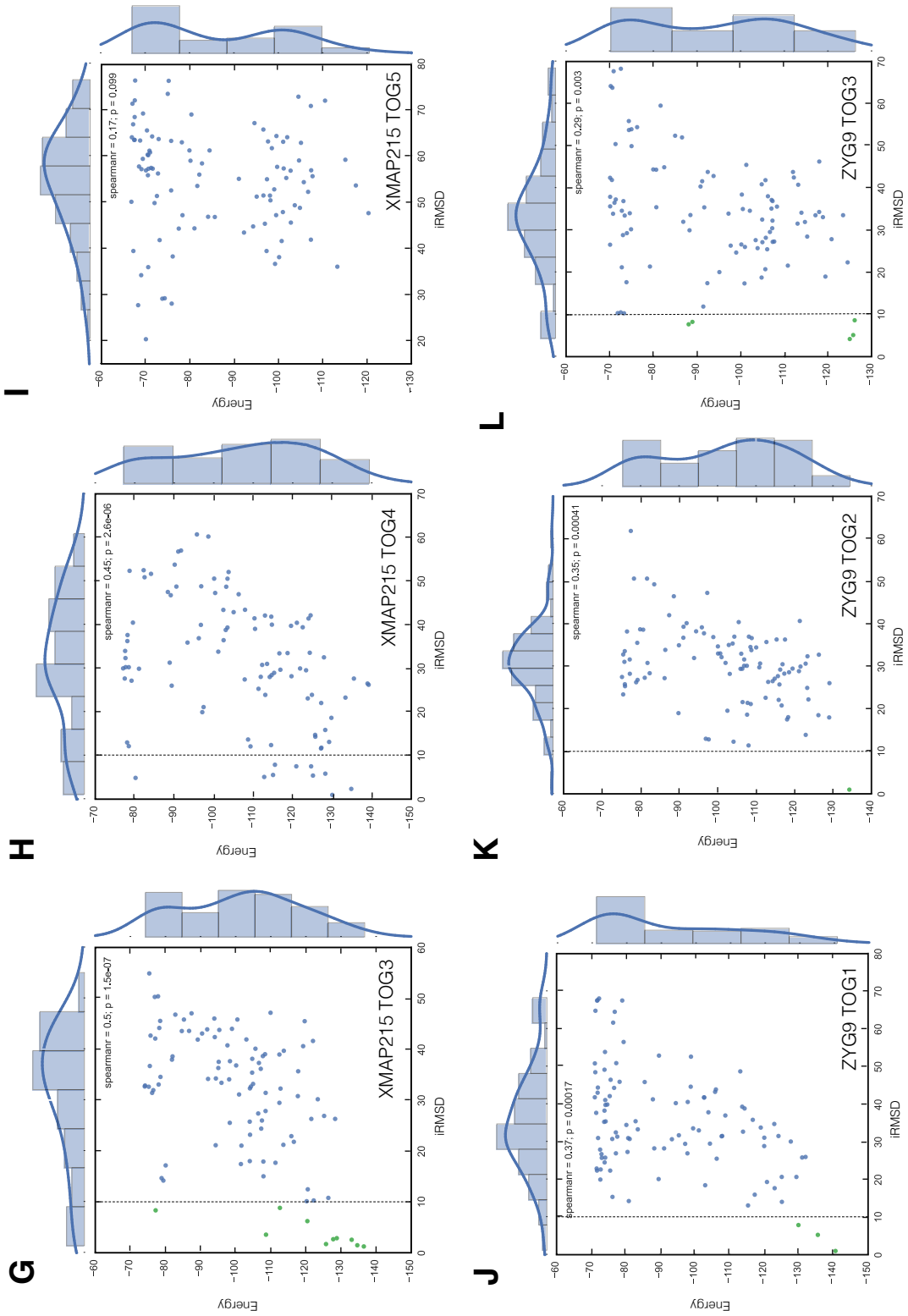


Figure 3.12: Plots of TOG domain docking results showing pyDock energy versus iRMSD with marginal for each axis together with Spearman Rank coefficient to assess correlation.

3.4 Methods

3.4.1 Brownian Dynamics

Browndye software (version 17-June-15) (Huber and McCammon, 2010) was used to compute brownian dynamics trajectories to calculate TOG domain association rates. The PQR files of atomic charges and radii were generated from PDB structures by pdb2pqr (version 1.9) (Dolinsky et al., 2007), using the AMBER force field at pH 6.9. APBS was used to generate electrostatic fields by solving the non-linear Poisson-Boltzmann equation. The contact distance was determined from a pilot study of the association of Stu2TOG1-2:$\alpha\beta$-tubulin after MD refinement. The calculated rate constants were 3.17e+8 $\pm$ 0.08 for Stu2 TOG1 and 9.43e+7 $\pm$ 0.17 for Stu2 TOG2. These simulations defined a contact distance of 5Å. Reactions were considered successful when simulations formed at least three interactions from the pairwise list. 200,000-500,0000 single trajectory simulations according to the NAM algorithm determined each association rate.

3.4.2 Molecular Dynamics Refinement

Initial atomic models for each TOG-Tubulin complex underwent energy minimization and equilibration via the amber11 molecular dynamics package. The LEaP modules of the AMBER 11 package (were used for the addition of missing hydrogen atoms. The system was neutralised by Na+ counter ions and solvated with TIP3P waters using an octahedral box of 12 Å. All simulations were performed with the Sander module of the Amber 11 package and the ff99SB all-atom force field and obtained parameters for guanine nucleotides plus magnesium. Energy minimization was carried out with decreasing constraints on the protein structure, followed by constant volume heating to 300K for ~10 ps and a 200 ps constant temperature and pressure equilibration. The SHAKE algorithm was used to constrain all covalent-bonds using hydrogen atoms and short-ranged non-bonded interactions were truncated at 10 Å with the Particle-Ewald method.

3.4.3 Binding Maps

Coordinates for each BD trajectory were extracted using the auxiliary program process_ trajectories and stored in a HDF5 table using PyTables. Density maps were constructed by kernel density estimation from the SciPy package and visualized in 3D using Mayavi.

3.4.4 Occupancy Maps

After each timestep of a BD trajectory, the orientational and positional coordinates of the TOG domain relative to the tubulin dimer are computed with respect to a reference coordinate system. The maps of positional and orientational coordinates were computed separately.

For the positional coordinates a spherical coordinate frame was chosen. The z-axis was defined as the vector from the center of the tubulin dimer to the center of the TOG domain in the bound state of the crystal structure. The x- and y- axis are orthogonal to the z-axis. The angle between the z-axis and the center-to-center vector for a given trajectory position is denoted as the zenith angle (θ). The azimuthal angle (ϕ) is the angle in the xy-plane from the x-axis.

The orientiational coodinates use the same spherical coordinate frame but exchange the centre-to-center vectors with normal vectors (denoted n_1 and n_2).

3.4.5 Transient Complex

The lowest energy structures were used as starting configurations for probing the transition complex. Briefly, the TOG domain was systematically translated and rotated with respect to the tubulin dimer, whilst monitoring steric clashes and inter-subunit contacts. For clash-free configurations, the number of contacts (N_c) at the onset of a sharp increase in the rotation angle (χ) was used to define the transient complex. The bound state in these terms showed numerous short-range interactions (high N_c) but restricted translational and rotational freedom (low r and χ). Alternatively, the unbound state showed at most a small number of interactions (low N_c) but expanded degrees of freedom (large r and χ).

The transient complex approach was used to examine the effect of mutations on binding affinities. The average electrostatic interaction energy was calculated based upon 100 randomly selected configurations from the transient complex ensemble.

$$\Delta G_{elec} = \Delta G_{elec}(comp) - \Delta G_{elec}(tubulin) - \Delta G_{elec}(tog) \tag{3.6}$$

$$\Delta\Delta G_{elec} = \Delta G^{*}(mut) - \Delta G^{*}(wt) \tag{3.7}$$

3.4.6 Tubulin Homology Modeling

Coordinates for the C-terminal tails of tubulin were kindly gifted by Jack Tuszynski, University of Alberta. The tails were docked to the Porcine tubulin structure and converted to electron density maps using USCF chimera. New tail sequences were constructed using modeler. These tail regions were MD refined into the EM density map using DireX (version 0.6.2) (Wang and Schröder, 2012).

3.5 Summary

Brownian dynamics simulations were able to determine the position of the TOG-binding site by sampling the conformation space under a single electrostatic force. This identified a single site that spans across both tubulin monomers that is in good agreement with the recent TOG:$\alpha\beta$-tubulin complexes. TOG domain association rates show positional dependence across the TOG-array and show differential activity on the basis of tubulin species. Mutations of residues to alanine in regions of high electrostatic potential caused a reduction in association rates. Protein docking was able to recapitulate the location of the TOG binding site.

4

THE DEVELOPMENT OF A MINIMAL +TIP SYSTEM

4.1 Introduction

MT dynamics are regulated by a conserved class of proteins called the plus-end tracking proteins (+TIPs). +TIPs display dynamic accumulation at the growing MT ends and mainly act as stabilising agents. The regulation of MT dynamics by +TIP proteins has major effects on cell functions. Usually +TIP proteins can be classified on the basis of structural motifs and their interactions with MTs.

Although many +TIPs can localise independently to MT ends in the absence of other proteins like EB1, in cells the functions of these individual proteins are likely to be determined by a combination of interactions with other +TIPs. This complex network is likely to be affected by mutual protein-protein affinity and competition for binding sites at MT tips. These interactions, even with a small number of components, are complex and a small system of +TIPs made from purified components is a necessary step to understanding the global functions of these proteins.

The main focus of my work has been an investigation of the detailed mechanism of tubulin binding by TOG domain polymerases, using computational chemistry techniques. Recent work has however revealed that the polymerase activity of XMAP215, perhaps the best-studied TOG polymerase, is substantially (synergistically) enhanced by EB-proteins (Zanic et al., 2013). The data suggested that the EB-TOG interaction is absolutely required for in vivo rates of plus end growth. The mechanism of this effect remains unclear. To investigate for the case of *S. pombe* MTs, I began work to purify and characterize the activities of *S. pombe* Mal3 and Alp14 for *S. pombe* tubulin. Time constraints meant that this work was only partially completed, but preliminary results are potentially revealing and are described below.

4.2 Aims

1. Reconstitute MT dynamics *in vitro* and image MTs under fluorescence and non-fluorescence conditions.

2. Purify baculovirus expressed TOG-polymerases at high yield and purify.

4.3 Results & Discussion

In vitro reconstitution experiments have been vital tools for dissecting the individual functions of tubulin and MT-associated proteins in absence of the complexity at the MT-tip. However, a largely overlooked issue has been the use of mammalian tubulin in the study of microtubule dynamics and in particular their regulation and interaction with MT associated proteins. The use of mammalian tubulin has a number of drawbacks. The use of heterologous components are likely to affect interactions, with brain tubulin having a number of isotypes and post-translational modifications. At present, despite recent advances in tubulin purification techniques, untagged tubulin purified from native sources has still proven difficult, especially in terms of yield and purity for biochemical assays (Minoura et al., 2013).

To avoid some of the aforementioned issues, a minimal system was developed to investigate the role of the MT polymerase Alp14 on single-isoform *S. pombe* tubulin in concert with key regulators Mal3 and Alp7.

4.3.1 Mal3 MT Plus-End Tracking *In Vitro*

Basal microtubule dynamics were established with unlabeled *S. pombe* tubulin. *S. pombe* tubulin was purified as described in Materials & Methods. This tubulin was used to establish a baseline from which parameters of MT dynamics (growth, shrinkage, catastrophe frequency and rescue) could be observed under the influence of additional proteins. MTs were imaged with dark-field microscopy (Fig. 4.1) using time-lapse acquisition of 30 mins. The tubulin concentration was titrated so that growth and shrinkage was observed from both ends of the MT without spontaneous nucleations in the field of view.

After a baseline of 4.5 μM tubulin was established, full length Mal3-GFP, with its histidine-tag (His) removed by cysteine protease cleavage from Tobacco Etch Virus (TEV), was investigated by Total Internal Reflection (TIRF) microscopy to determine its ability to bind *S. pombe*

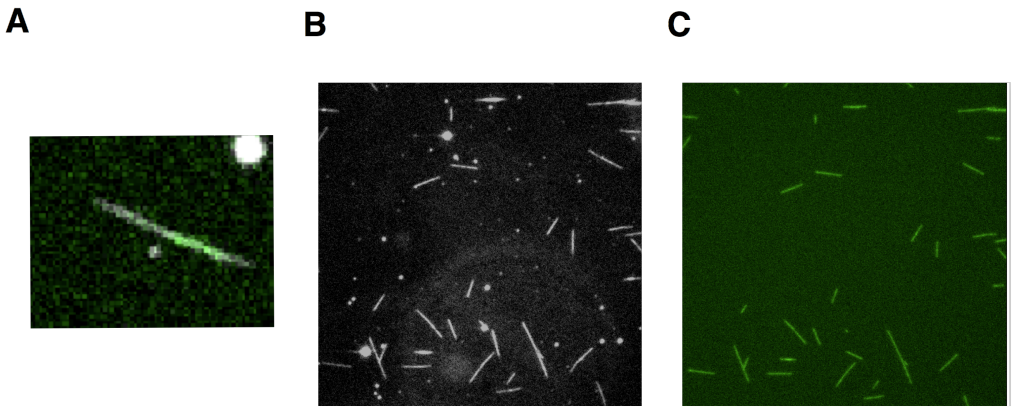


Figure 4.1: Microtubule dynamics reconstituted *in vitro* with *S. pombe*, imaged with dark-field microscopy. (**A**) Combined fluorescence and darkfield channel. (**B**) MT growth from seeds with unlabelled *S. pombe* tubulin imaged by dark field microscopy. (**C**) MT-seeds labeled with Alexa-488 fluorophore imaged by epifluorescence.

MTs and tip track at different ionic strengths. Purification of Mal3-GFP was performed as described in Materials and Methods.

Previous studies of Mal3 *in vitro* have shown the protein has enhanced affinity for the growing MT ends as opposed to the MT lattice . This affinity difference at the tip was proposed as the mechanism that allows it to track growing MT ends autonomously. Mal3 selectively accumulates at the growing MT ends at considerable ionic strength and protein concentration. Mal3 also binds weakly along the entire length of MTs (Bieling et al., 2007).

Determining the conditions for Mal3 to produce the comet-like structures on *S. pombe* tubulin as seen in cells was important because high concentrations of Mal3 could potentially decorate the MT lattice, abrogating its role with Alp7 and Alp14 at the MT tip.

Mal3-GFP was titrated under low ionic strength conditions (100 mM PIPES, 1 mM $MgSO_4$, 2 mM EGTA, adjusted to pH 6.9 with KOH) from 0.5 nM to 300 nM. Mal3 bound strongly across the entire MT lattice, including to GMPCCP stabilised seeds (Fig. 4.2). From this experiment, 50 nM Mal3 was determined to be the optimal concentration of protein to reproduce its normal action without spontaneous nucleation of new MTs.

The ionic strength was then varied in order to investigate the point at which Mal3 binds preferentially to the MT tip rather than along the entire MT lattice. In order to do this, it was necessary to increase the tubulin concentration to 14 μM because high salt concentrations cause inhibition of MT polymerisation.

For low ionic strength conditions Mal3 bound with equal occupancy to the MT lattice and tip. But when the ionic strength increased, Mal3 accumulated more strongly at the tip and the fluorescence intensity reduced on the lattice (Fig 4.3). This is an indication that Mal3 recognises a structural feature in the growing MT end region that is distinct from older, lattice-incorporated GDP tubulin (Maurer et al., 2012).

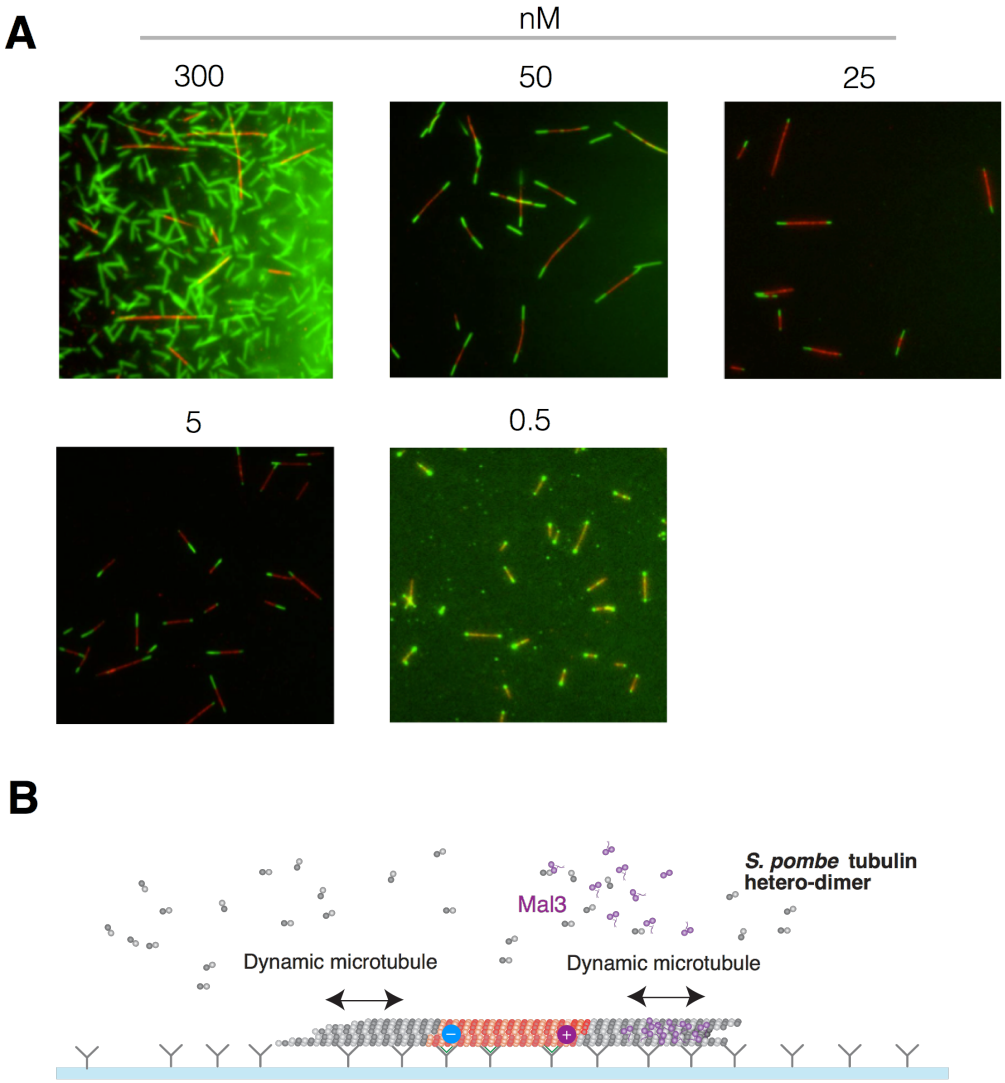


Figure 4.2: Titration of the Mal3 concentration from 300 nM to 0.5 nM with 4.5 µM *S. pombe* tubulin. (**A**) TIRF microscopy images of dynamic MTs over a range of Mal3 concentrations. (**B**) The assay consists of stabilised rhodamine-labelled seeds bound to a glass surface by antibodies. The unlabeled tubulin is add to the chamber alongside Mal3. As the *S. pombe* tubulin is unlabeled, the signal (green) comes from excitation of the fluorescent Mal3 bound to the MT.

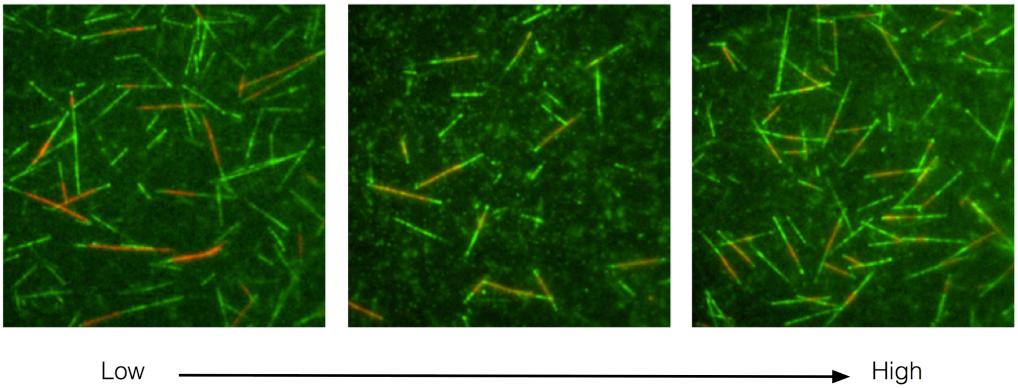


Figure 4.3: Mal3 binding over a range of ionic strength conditions for 50 nM Mal3 and 14 µM *S. pombe* tubulin. The salt concentration was titrated between 100 mM to 85 mM monovalent salt

4.3.2 Purification of Alp14-Avi

Baculovirus insect cell expression systems are widely used in industry and academia for the production of recombinant protein. Baculovirus expression systems are attractive in that they contain eukaryotic processing capability and express high-levels of protein when compared to other eukaryotic expression systems. Insect cell lines grow well in suspension culture, permitting large-scale protein purification as evidenced in the pharmaceutical industry.

Baculovirus was produced as described in the material and methods. The N-terminal Alp14-Avi vector was kindly gifted by Frauke Hussmann, formerly of the Cross Lab. To determine the protein yield before purification, Alp14-Avi was expressed in 4x 500 L cultures of Sf9 cells at 1x10^6 cells/ml infected with 400 µl of Alp14 Baculovirus Infected Cell Stocks (BIIC) (Wasilko et al., 2009) and 400 µl of BirA biotin Ligase. Cells were pelleted after 72 hours and samples retained for quantification by Western Blotting (Fig 4.4).

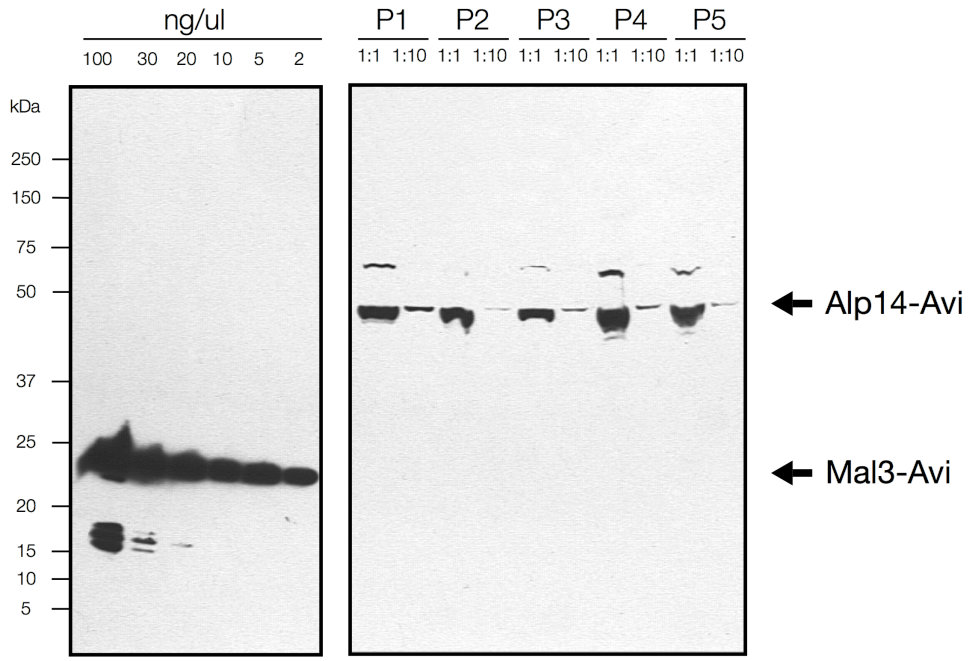


Figure 4.4: Western blot of Alp14-Avi pellets from insect cell expression. Left panel contains Mal3-Avi used as a loading standard to quantify the pellet bands from expression on the right.

Quantification of protein yield from Western blots was determined by densitometry. Films were imaged along with an optical density wedge and band intensities were converted to OD values. This determined the protein yield to be ~3 mg/L.

Initially frozen cell pellets were recovered in lysis buffer and clarified by ultracentrifugation. The supernatant was loaded onto an anion exchange column after equilibration with low and high-salt HiTrap buffer. The protein was eluted over a linear gradient and fractions identified by SDS-page and Western Blotting (Fig. 4.5).

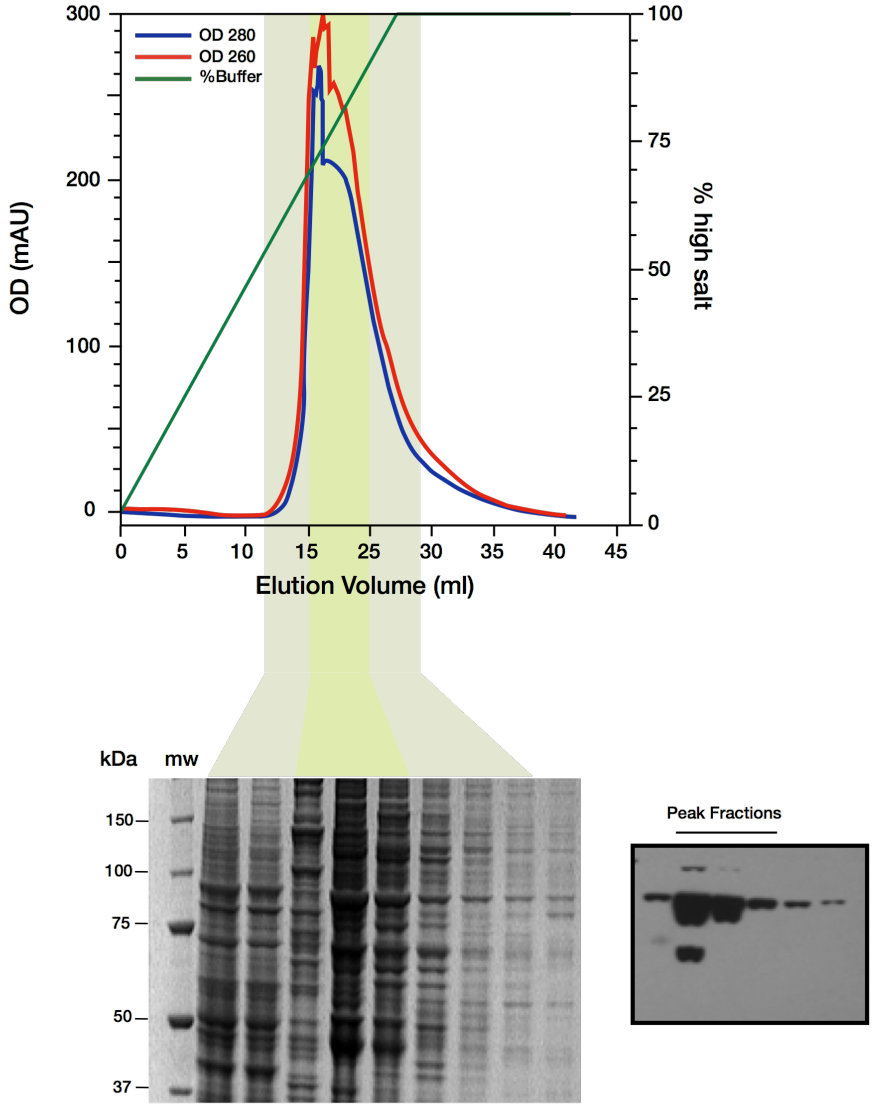


Figure 4.5: Elution profile of anion-exchange chromatograph column at OD_{260} and OD_{280} with a high salt buffer gradient. Peak fractions after staining with SimplyBlue and Western Blot against the biotin antibody. The shaded areas show the fraction analysed by SDS-page.

The pooled fractions from the anion-exchange step were loaded onto a self-packed column containing 2 ml of monomeric avidin resin. The column was equilibrated with low-salt and high-salt Avi buffer. After binding the protein to the column in low-salt buffer, the protein

was eluted in 10 mM D-Biotin + 2M Urea. This caused the protein to elute from the column in multiple peaks and to gradually leach from the column over many column volumes.

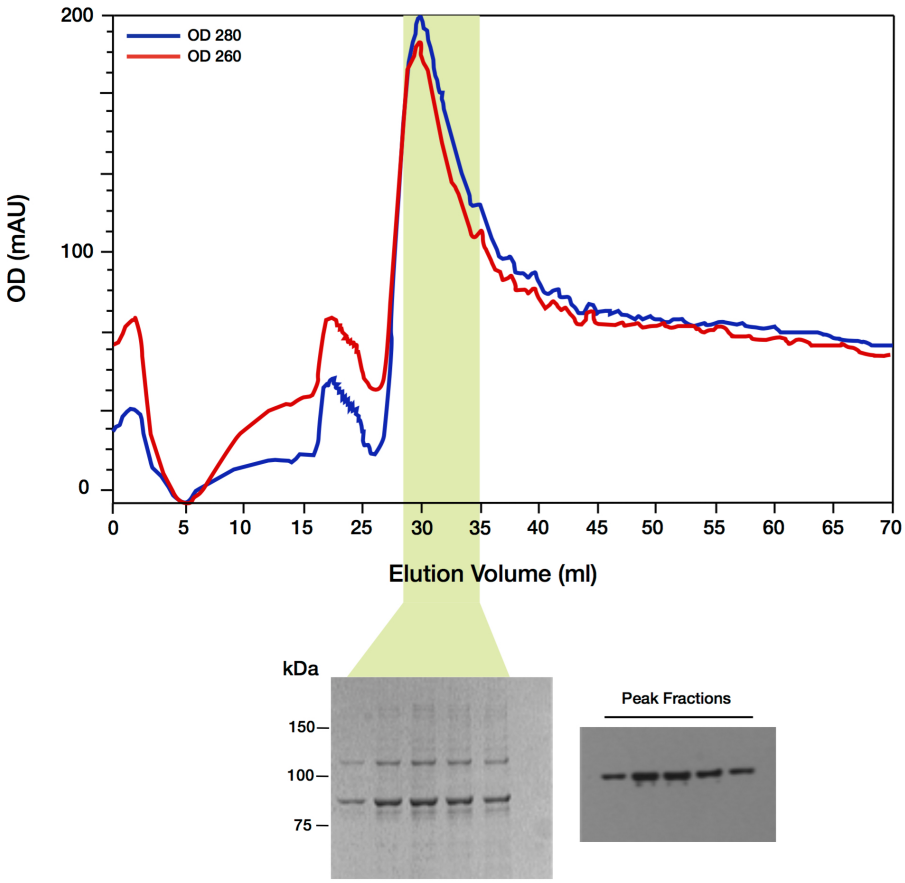


Figure 4.6: Elution profile from Avidin column at OD_{260} and OD_{280} under normal elution conditions. Peak fractions are stained with Coomasie blue and Western blotted against the biotin antibody.

Analysis of the OD peaks from the anion-exchange column showed indications of DNA contamination. This may have caused Alp14 to bind non-specifically to the Avidin matrix and cause the slow, diluted elution profile (Fig 4.7).

Benzonase was therefore added to the lysis buffer to digest DNA and reduce viscosity to improve column flow characteristics. 0.1% Tween-20 was added to all buffers to reduce possible non-specific interactions alongside 0.1% Zwittergent, a non-ionic detergent to further improve solubilisation.

Following the procedures as discussed, and with improved buffers, fresh Alp14 lysate was bound to a 5 ml HP SP column. The protein was eluted over a continuous gradient up to

100% high-salt buffer. The duration of the salt gradient was reduced in order to broaden the peak and prevent the protein from eluting at high salt in a very tight peak.

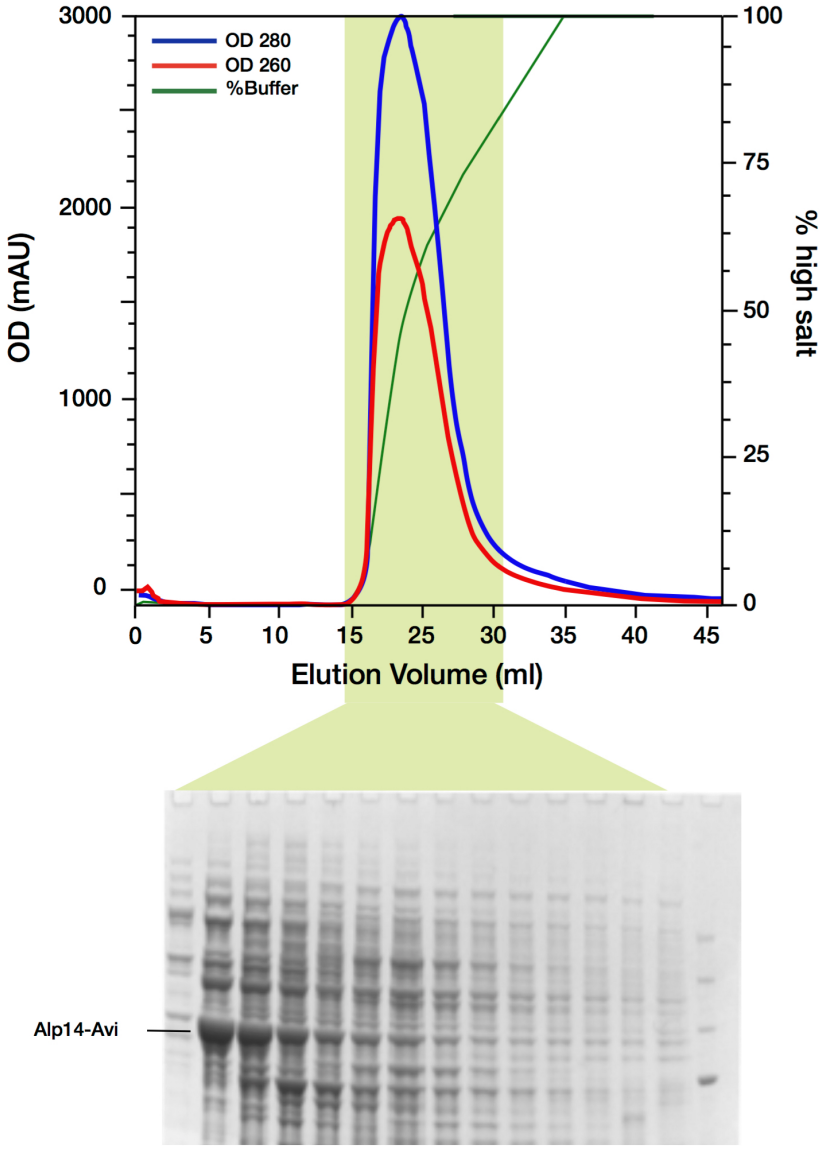


Figure 4.7: Second elution profile of anion-exchange chromatography. Peak fractions after staining with SimplyBlue.

The buffer improvements showed a considerable reduction in the overall level of DNA, as evident by Fig. 4.7, and an improved elution profile.

Fractions containing protein were loaded onto an avidin column and eluted with 10 mM D-Biotin + 2M Urea. This time the protein eluted in a single peak and fractions were collected and dialysed over night (Fig 4.8) and flash frozen for storage in LN_2.

Subsequent testing of the purified protein, before and after buffer exchange, on dynamic MTs showed no affect on growth. The protein was therefore deemed non-functional.

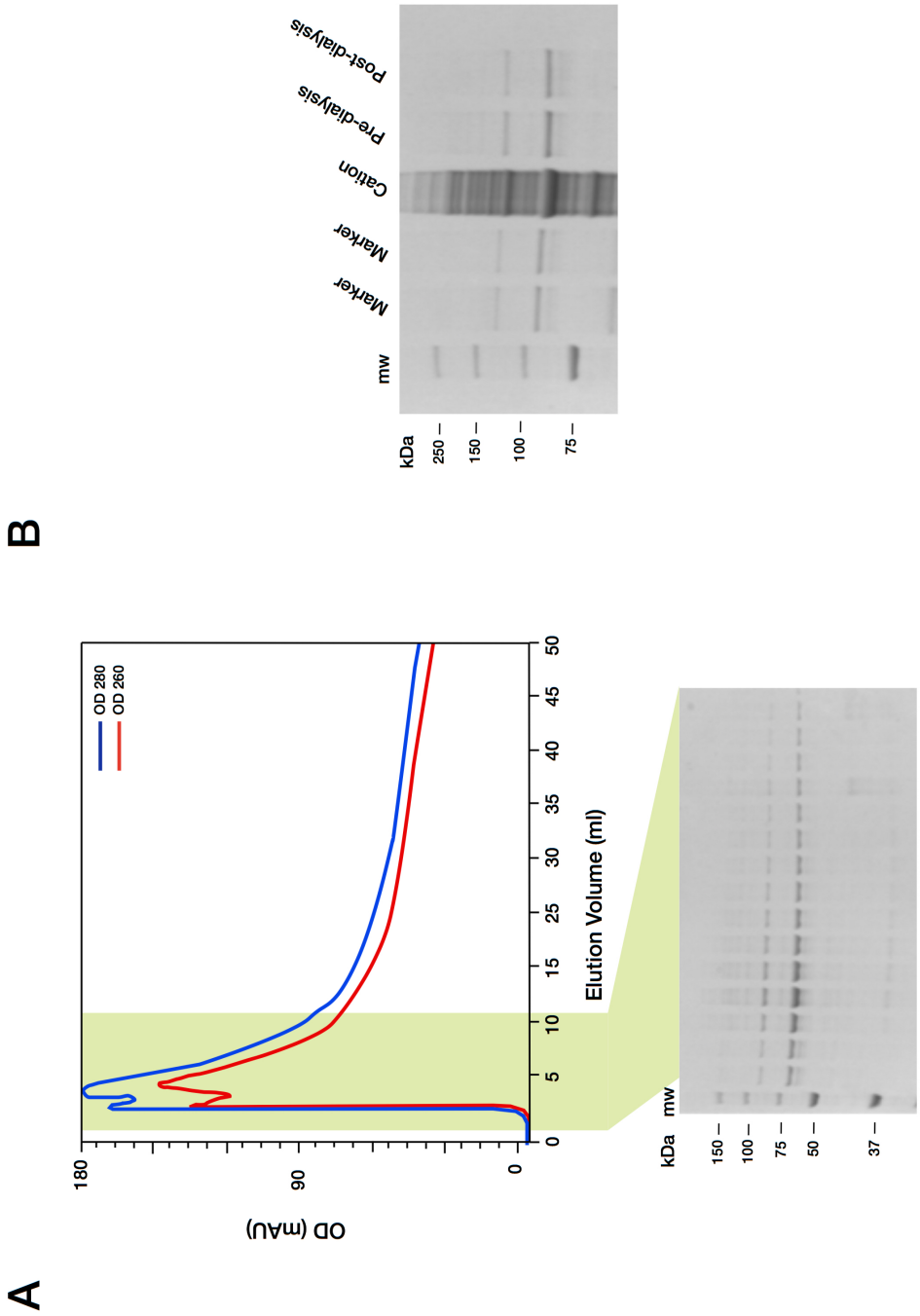


Figure 4.8: Elution profile from avidin column under enhanced elution conditions. (**A**) Peak fractions after staining with Simply Blue. (**B**) Gel of Alp14-Avi pre- and post-dialysis.

4.4 Materials & Methods

4.4.1 Tubulin Expression & Purification

S. pombe tubulin was expressed from a single-isoform strain created by Douglas Drummond (Cross Lab) where the atb2(α) gene was replaced with the nda2 gene (α1). An 80 L culture of *S. pombe* cells was grown externally in a commercial fermenter and collected by centrifugation.

Lysis

The pellet of *S. pombe* cells was re-suspended in lysis buffer (100 mM K-PIPES pH 6.9, 10 mM MgSO4, 2 mM EGTA, 50 mM NaCl, 1 mM GTP, 5 mM DTT, 1 mM TAME, 1 µg/ml aprotinin, 1 µg/ml AESSF, 1 µg/ml E-64, 1 µg/ml Pepstatin, 1 µg/ml Chymostatin, 1 µg/ml Antipain, 2 µg/ml Leupeptin) and the cells were disrupted using a bead mill at 4°C. The cell lysate was clarified by at 12,000 rpm for 30 mins at 4°C.

Weak Anion Exchange

The soluble protein extract was batch bound to equilibrated DE-52 resin (100 mM K-PIPES, 1 mM MgSO4, 2 mM EGTA pH 6.9) for 30 mins at 4°C by gentle stirring. This settled ion exchange resin was packed into a X50/30 chromatography column and using an Akta purifier washed with 3-4 column volumes of DE-52 low salt buffer (100 mM K-PIPES pH 6.9, 10 mM $MgSO_4$, 2 mM EGTA, 50 mM NaCl, 10% (v/v) glycerol, 0.1 mM GTP, 5 mM DTT, 1 µg/ml aproptinin, 1 µg/ml AEBSF, 1 µg/ml E-64, 1 µg/ml Pepstatin A, 1 µg/ml Chymostatin, 1 µg/ml Antipain, 2 µg/ml Leupeptin). The tubulin was eluted using a gradient from 0 to 60% of high salt DE-52 buffer (100 mM K-PIPES ph 6.9, 10 mM $MgSO_4$, 2 mM EGTA, 1M NaCl, 10% (v/v) glycerol, 0.1 mM GTP, 5 mM DTT, 1 µg/ml aproptinin, 1 µg/ml AEBSF, 1 µg/ml E-64, 1 µg/ml Pepstatin A, 1 µg/ml Chymostatin, 1 µg/ml Antipain, 2 µg/ml Leupeptin).

Precipitation and Desalting

Tubulin containing fractions were precipitated slowly by adding 375 mg of ammonium sulphate for each 1 ml of solution. Once the salt was dissolved it was left overnight at 4°C. The precipitated protein was pelleted at 9,143 rpm for 30 mins at 4°C and then resuspended in 12 ml of resuspension buffer (100 mM K-PIPES pH 6.9, 1 M $MgSO_4$, 2 mM EGTA, 0.1 mM GTP, 1 µg/ml aproptinin, 1 µg/ml AEBSF, 1 µg/ml E-64, 1 µg/ml Pepstatin A, 1 µg/ml Chymostatin, 1 µg/ml Antipain, 2 µg/ml Leupeptin) at 4°C for 45 min on a roller mixer. The

re-suspended pellets were centrifuged at 13,000 rpm for 2 min at 4°C and desalted into resuspension buffer with a HiPrep 26/10 desalting column.

Weak Anion Exchange

The peak fractions were pooled from desalting and loading onto a 5 ml HiTrap Q HP column and washed with Q low-salt buffer (100 mM K-PIPES pH 6.9, 1 mM $MgSO_4$, 2 mM EGTA, 50 mM KCl, 50 µM GDP) and then eluted with a linear gradient (0-70%) of Q high-salt buffer (100 mM K-PIPES pH 6.9, 1 mM $MgSO_4$, 2 mM EGTA, 1 M KCl, 50 µM GDP).

Tubulin Cycling

Tubulin containing fractions were polymerised by adding GMPCPP to 0.6 mM, 0.1% glycerol and 1% DMSO and incubated for 1 hr at 32°C. The polymerised MTs were pelleted by centrifuging for 10 min at 68,000 rpm at 32°C. The MT pellet was re-suspended in 5 ml of warm depolymerisation buffer (20 mM Na-PIPES pH 6.9, 1 mM $MgSO_4$, 5 mM $CaCl_2$, 1 mm GDP) and then incubated at 4°C to depolymerise. The depolymerised tubulin was clarified by centrifugation at 68,000 rpm for 10 mins at 4°C.

Gel filtration chromatography

The supernatant was loaded on a Superdex 200 XK16 60 gel filtration column equilibrated with Superdex buffer (100 mM K-PIPES pH 6.9, 1 mM $MgSO_4$, 2 mM EGTA, 150 mM NaCl, 50 µM GTP) and eluted in the same buffer.

Desalt

The tubulin was desalted into storage buffer (100 mM K-PIPES pH 6.9, 1 mM $MgSO_4$, 2 mM EGTA, 150 mM NaCl, 50 µM GTP) using a HiPrep 26/10 rapid desalting column. The eluted tubulin was concentrated to 30-40 µM using a Vivaspin 15R 10,000 MWCO Spin Concentrator at 4,100 rpm for 10 min at 4°C.

4.4.2 Mal3 Expression & Purification

The Mal3-308-GFP vector was provided by Miho Katsuki, formerly of the Cross Lab. The DNA fragment encoding full-length Mal3 was PCR amplified and cloned into a pET17b vector (Novagen) with the addition of an N-terminal 6xHis tag.

Bacterial Transformation

BL21Star (DE3) component cells were thawed on ice. 50 µl of cells were mixed with ~100pg of plasmid DNA, gently agitated and left to incubate on ice for 30 mins. The cells were then heat-shocked at 42°C for 30 secs and immediately placed back on ice for a further 5 mins. 950 µl of room temperature SOC media was added to the transformation mixture and incubated at 37°C for 1 hour in a rotary shaker. 50-100 µl of the transformation mixture was pipetted onto pre-warmed selection plates and dispersed with a sterilized spreader under open flame. Plates were incubated overnight at 37 °C and monitored for the presence of colonies.

Bacterial Expression

E. coli BL21Star (DE3) cells containing the pET vector expression constructs were grown in 6x 2L flasks containing 400 ml of LB media supplemented with 100 µm/ml ampicillin at 37°C, 250 rpm. Cells were grown to a density of OD_{600} = 0.6 and the temperature reduced to 20°C for induction with 0.5 mM IPTG and incubated for a further 6 hrs. Cultures were rapidly cooled on ice and harvested at 6,000 rpm for 10 mins. The pellets were flash frozen in liquid nitrogen and stored at -80°C.

Protein Purification

Cell pellets were re-suspended in lysis buffer (15 mM Bicine pH 8.0, 2 mM Mg Acetate, 0.4 M NaCl, 0.1 mM ATP, 1 mM DTT, 1 mM PMSF, Roche cOmplete EDTA-free protease inhibitors). Cells were lysed by sonication at 4°C for 4x 15 secs and the homogenate centrifuged at 45,000 rpm for 20 mins at 4°C. The collected supernant was batch bound to 4 ml of Ni-NTA pre-washed (2x ddH_20, 2x lysis buffer) at placed on a roller at 4°C for 30 mins. The Ni-NTA resin was packed into a Poly-Prep Chromatography Column with glass wool and washed with 15 mM Bicine (pH 8.0), 2 mM MgAcetate, 0.4 M NaCl, 10 mM ATP, 25 mM imidazole. Mal3 was eluted using 15 mM Bicine (pH 8.0), 2 mM Mg Acetate, 0.4 mM NaCl, 1 mM ATP, 100 mM imidazole. The pool fractions of His protein were then loaded on a Superdex 200 16/60 gel filtration column equilibrated with 100 mM K-PEM (pH 6.9), 1 mM $MgSO_4$, 2 mM EGTA, 100 mM NaCl and eluted in the same buffer. Aliquots were flash-frozen in liquid nitrogen.

4.4.3 Alp14 Expression & Purification

Transfection of Sf9 Cells

Escort IV reagent was homogenised by gentle shaking. 12 µl of Escort IV was diluted into 400 µl of insect cell media for each transfection reaction. Separately 5 µl of BacMagic DNA (100 ng) + 10 µl of transfer vector (500 ng) DNA was diluted into 400 µl of insect cell media. The Escort IV mixture was combined with the DNA mixture and left to incubate for 15-45 mins at room temperature. 5ml of Sf9 cells at a density of $1x10^6$ cells/ml were transferred into a 15 ml falcon tube and pelleted at 1,200 rpm for 5 mins. The supernatant was removed and the remaining cells were resuspended in 1 ml of insect cell media. 200 µl of the resuspended cells were gently pipette into each well of a 24 well-plate that contains the previously incubated DNA-Escort complex. The well-plate was covered with BreathEasy tape and sealed at its edges with parafilm before being incubated overnight at 200 rpm. After 24 hours, 1 ml of serum-free insect cell media was added to each well. The cells were incubated for at least a further 5-7 days or until the virus had taken hold. The passage 1 (P1) virus was collected by pelleting the cells at 1,300 rpm for 5 mins. The viral supernatant was filtered with a 0.45 µm syringe filter and stored at 4°C and protected from light.

Production of Baculovirus Infected Cell (BIIC) Stocks

50 ml of Sf9 cells at a density of $0.5x10^6$ cells/ml were infected with 50 µl of P1 virus plus 2% fetal bovine serum (FBS). The cells were monitored at 24, 30 and 36 hours. Once the cells reached a mean diameter of 22-23 µm they were harvested at 1,500 rpm for 5 mins. The cell pellet was re-suspended in 5 ml of insect cell medium supplemented with 10% DMSO. 200 µl of the cells containing DMSO were aliquot into Nunc cryotubes and frozen down to -80°C using a 1°C Nalgene cooling rack. The samples were transferred to LN_2.

Expression

One frozen BIIC stock was quickly thawed at 37°C and diluted 1:100 in insect cell media. 1ml of diluted stock was added to flasks containing 100 ml at a density of $1x10^6$ cells/ml. The infected cultures were incubated at 28°C, 120 rpm. Cell density, viability, and diameter were monitored. Cells were harvested by centrifugation.

Purification

Cell pellets were resuspended in lysis buffer on a roller mixer at 4°C for 40 mins and further lysed with a Dounce homogeniser. The lysate was clarified by ultracentrifugation with a TLA 100.3 rotor at 4°C, 38K rpm for 40 mins. Supernant was loaded onto a 5 ml HiTrap SP HP column pre-equilibrated with low-salt HiTrap buffer (6.7 mM HEPES pH 7.5, 6.7 mM MES, 6.7 mM Na-Acetate, 150 mM NaCl) and high-salt Hi Trap buffer (6.7 mM HEPES pH 7.5, 6.7 mM MES, 6.7 mM Na-Acetate, 1 M NaCl). Protein containing fractions were pooled and

loaded onto a self-packed 2 ml monomeric avidin column pre-equilibrated with low-salt avi buffer (100 mM $NaPO_4$, 150 mM NaCl) and high-salt avi buffer (100 mM $NaPO_4$, 450 mM NaCl). The Avi tagged protein was eluted with 10 mM D-Biotin + 2M Urea. Peak fractions were dialysed overnight into storage buffer (100 mM PIPES pH 6.9, 1 mM $MgSO_4$, 2 mM EGTA) and stored in LN_2.

4.4.4 Protein Detection Methods

SDS-Page

Protein samples were prepared in a mixture of 4x LDS sample buffer and 10x reducing agent and denatured at 70°C for 10 min. Samples were resolved by SDS-Page (NuPage 10% Bis-Tris Gels) run in MOPS buffer inside of a XLock SureLock Mini Cell. Gels were run at a constant 200V for 55 mins. Upon completion, samples were stained with SimplyBlue SafeStain (Invitrogen) and imaged with a Hamamatsu camera.

Western Blot

The iBlot dry blotting system (Invitrogen) was used to transfer proteins to nitrocellulose using cycle P3 for 7 mins at 200V. The membrane was blocked with blocking buffer (5% dried skimmed milk, 0.25% Tween-20, PBS) for 1 hour on an orbital shaker. Primary antibody was added to the membrane and left overnight at 4°C. The blot was then washed 3x with PBST buffer (PBS, 0.1% Tween-20) at and incubated with second antibody for 1 hour at room temperature on an orbital shaker. The membrane was rinsed 3x with PBST buffer for 10 mins on an orbital shaker. Proteins were developed by SuperSignal West Femto Maximum Sensitivity Substrate (Pierce) for 5 mins and exposed to Kodak BioMax film.

4.4.5 Flow Cells

Coverslips and slide glass were immersed in acetone for 55 min, ethanol for 10 min, rinsed in ddH_20 for 1min. Glass was oven dried at 60°C and plasma cleaned for 2x 150 sec cycles. Glass was then incubated in 0.1 M KOH, bathed twice in ddH_20 for 1 min and over dried at 100°C before silanisation in 0.05% (v/v) dichlorodimeythysilane (DDS) in trichloroethylene (TCE) for 1 hour. Silanised glass was sonicated in fresh methanol for 5, 15 and 30 mins. Glass was spun dried and stored in Ross Lens Tissue inside of a desiccator cabinet. Flow cells were assembled by attaching a silanised coverslip to the slide using double-side tape.

4.4.6 Microtubule Dynamics Assays

MT seeds were assembled by incubating 5 µM 10% Alexa488-labeled subunits in 1mM GMPCPP, 1mM $MgCl_2$, K-PEM (100 mM K-PIPES (pH 6.9), 1 mM $MgSO_4$, 2 mM EGTA) at 37°C for 35 min. Assembled seeds were pelleted at room temperature by an Air-Driven Ultracentrifuge for 5 min at 25 psi and re-suspended in warm K-PEM buffer.

A flow cell was coated with Anti-Alexa-488 antibodies diluted 1:10 in K-PEM buffer. MT seeds were added to the cell and allowed to adhere for 5 min. Flow cells were flushed with K-PEM buffer to remove unbound seeds. Dynamic MTs were assembled from stabilised seeds with tubulin, +TIP proteins, 1 mM GTP, 1% (v/v) 2-mercaptoethanol and GCO oxygen scavengers (8 µg/ml catalyse, 4.5 mg/ml glucose, 38 U/m glucose oxidase) in K-PEM.

4.4.7 Microscopy

Darkfield

MT dynamics were imaged using a Nikon E800 fluorescence microscopy with a customized dark-field illumination system for image capture. The microscope was mounted in a custom box with heater to maintain a constant temperature of 25$\pm$0.5°C. A Plan Fluor 10x 0.5-3 NA objective lens was fitted to the microscope along with an EM CDD camera for image capture. Dark-field illumination was provided by a 100 W mercury lamp connected via a fibre-optic light scrambler, cold mirror and a band-pass filter to remove both UV and IR wavelengths. Epifluoresence illumination used a second mercury lamp coupled to the microscope via a light pipe. Fluorescence excitation and emission filters were mounted in motorised filter wheels to permit rapid switching between fluorescence and dark-field imaging modalities. Metamorph software was used to control the microscope and camera.

TIRF

Fluorescent proteins were observed on an Olympus TIRF system using a 100x NA 1.49 objective, 1.6x magnification with 488 nm and 561 nm laser diode. Images were captured on a back-illuminated EM-CCD camera (Hamamatsu) under the control of xcellence software (Olympus).

4.5 Summary

Mal3 and *S. pombe* tubulin were successfully purified and established in a simple *in vitro* reconstituted system. The Mal3 conditions for tip-tracking on dynamic *S. pombe* MTs were determined and this will be important for any future work that looks at the interaction of these proteins.

The Alp14-Avi protein was purified, but further development of the downstream processing steps are required before protein of high-yield and purity is achievable. A number of possible alternative strategies could be taken to improve on the purification work:

1. The expression construct could be altered to include a maltose-binding protein tag which is known to improve both expression and solubility.

2. Protein could be expressed in alternative insect cell lines and tested for improved expression levels.

3. A two-step affinity purification protocol could be developed. This strategy has been successfully applied to the purification of a number of TOG-proteins expressed in Baculovirus.

5 DISCUSSION & FUTURE WORK

TOG family polymerases are enzymes that lower the activation barrier between their substrate, unpolymerised GTP-tubulin, and their product, polymerised GTP-tubulin. The scientific question I have addressed here is, what is the reaction pathway for the formation of the collisional complex between tubulin and TOG domains? TOG-family polymerases are proposed to catalyse microtubule assembly by enriching an otherwise-transient exchange intermediate of GTP-tubulin. The dynamics of formation and dissociation of this exchange intermediate will govern the catalytic action of TOG-polymerases, and the exchange intermediate is likely to be similar to the collisional complex studied here. As discussed, in most binding reactions the initial, collisional complex subsequently evolves into a more stable state. However high catalytic activity requires that GTP tubulin engage quickly with the TOG domain, and then be quickly donated into the growing GTP-polymer. Experiment indicates that this route is at least 10-fold faster than that for direct incorporation of GTP-tubulin into the lattice. Accordingly, the mechanism by which the initial collision complex of TOGs with GTP tubulin is formed and enriched is likely to be central to the catalytic mechanism of TOG-polymerases.

To analyse the dynamics of TOG-tubulin association and to understand how the structural features of TOGs and tubulin determine the reaction pathway, I have applied structural and electrostatic analysis, combined with Brownian Dynamics simulations. The results of this work offer new insights into the TOG-domain binding mechanism.

Chapter 2 presents the results of an analysis of the effects of electrostatics on TOG-domains, and their role in TOG-tubulin interactions. This analysis identified conserved features of the overall electrostatic field across homologues. It was found that the tubulin binding loops exert a dominant influence on the global electrostatic field. Mutations of charged residues to alanine further emphasised the importance of these binding loops for configuring the global electrostatic field across the entire TOG-array. In particular, residues outside of the binding loops had little effect on perturbing the electrostatic field and all those within the turns caused a similar, high degree of perturbation. Electrostatic similarity analysis and clustering showed that domains with similar electrostatics similarities showed analogous simulated association rates that reflect individual differences in binding kinetics.

Normal mode analysis and examination of residue coupling highlighted features of TOG domains that vary in relation to their array position. How this relates to the collective mechanism by which the TOG array binds and exchanges tubulin is unclear. Exploring collective behavior would make an obvious extension of this work in the future.

Normal mode analysis identified correlated motions all along the TOG-tubulin binding interface, with the more significant correlated motions occurring in the loops that engage with β-tubulin, suggesting that the interface with β-tubulin represents a core functional element of each TOG.

TOG domains possess exceptional sequence conservation in and around the central turn of each HEAT-repeat (A-E), despite considerable variation in the overall sequence identity between domains. This was already recognised as suggesting a common mode of interaction (Slep and Vale, 2007). Many of the residues around this region are conserved and solvent exposed, with the latter known for their critical role in protein-protein interactions (Al-Bassam et al., 2007). For example, conserved residues in the intra HEAT-repeat loops are also present in the binding surface of protein phosphatase 2A (Andrade et al., 2001). Similarly, the conservation of charge found in the face of the TOG-domain has been consistently implicated by mutational studies as being critical for mediating tubulin binding (Al-Bassam et al., 2010)(Ayaz et al., 2012). The variable effects of these mutations on TOG-tubulin interactions are linked to their effects on MT polymerase activity in full-length TOG proteins (Widlund et al., 2011).

The greater flexibility seen in TOG-loop regions that interact with β-tubulin suggests a possible mechanism of conformational selection for polymerisation-favourable tubulin states from free solution. Tubulin is thought to exist in at least two distinct conformations within and outside of the MT lattice. The conformations are distinguished largely by shifts in the β-tubulin due to compaction at the α-site as a result of nucleotide hydrolysis. Striking evidence for the conformational selection of tubulin by TOG-domains has been highlighted in recent CLASP crystal structures (Leano et al., 2013). TOG domains from these proteins, which share similar sequence determinants of tubulin binding, show a highly convex structure which contrasts with the flat tubulin-binding surface of XMAP215 family TOG domains. The CLASP TOG domains are thought to have a putative role in associating to tubulin dimers in curved, depolymerising protofilaments in order to promote rescue.

In future work, a number of improvements might be applied to the existing analysis to improve the resolution and accuracy of the results.

The evaluation of TOG-tubulin binding energies, with particular emphasis on mutation of charged residues within the TOG-domains, could be improved by using the molecular-mechanics Poisson-Boltzmann surface area (MM-PBSA) method. This technique uses the same thermodynamic scheme as the implicit electrostatic calculations, but has improved accuracy due to the addition of the binding free energy contributions from molecular mechanics and entropic changes. MM-PPBSA is more computationally efficient than more rigorous methods such as free energy perturbation (FEP) and thermodynamic integration methods and would make a more natural extension to the current work.

Further, the implicit electrostatic calculations might be improved by the addition of explicit water molecules at the protein surface. In an implicit solvent model, electrostatic calculations treat bulk water and water that is structurally or chemically important in the system as the same, so that both types of water are described by a single dielectric constant. However, studies have shown that bulk water is distinct from water that interacts with the protein surface, which can stabilise protein-protein interactions. In some cases, treating these functionally important waters as if they are part of the bulk solution may be detrimental to the quality of the calculation of a proteins electrostatic potential.

The ideal method of measuring (as opposed to predicting) the strength of binding between TOG domains and tubulin, plus the extent of their electrostatic fields, is to use isothermal calorimetry (ITC) on real proteins in solution. This technique is capable of determining the magnitude of the binding affinity by measuring the temperature change on protein binding and has the capability to measure different energetic contributions to the affinity. The electrostatic field of each TOG domain could be assessed by vibrational Stark spectroscopy which determines the interactions between a probe on the protein surface and the electric field due to the surrounding environment.

Chapter 3 directly explores pathways for the formation of diffusional encounter complexes between TOG domains and tubulin heterodimers. BD simulations were used to show that electrostatics alone drives TOG domains to form a consistent encounter complex at a site very close to the crystallographically-identified TOG-binding site on tubulin. These simulations, combined with protein docking, allowed comparison of the propensity of TOG domains to interact with tubulin, and in particular allowed a correlation of TOG-array position with binding propensity. For a native-complex end-state, TOG domain association rates showed species-specific and array-position-dependent conservation. In general, for cognate tubulin binding partners the fastest rates were found for TOG1 and TOG2, with the association rate decreasing for more distal TOGs.

For TOGS from a specific species, clear differences were identified between association rates for cognate tubulins and for porcine tubulin (Fig 3.4).

Association rates to a native binding position with porcine tubulin were consistently more variable across species, with TOG1-2 exhibiting the greatest difference. Intriguingly Alp14 TOG1-2 showed slower rates for porcine tubulin and faster rates for Dis1. This is consistent with unpublished data from the Cross lab showing that porcine tubulin forms an overly-stable complex with yeast TOGs, causing it to act as a competitive inhibitor for the interaction of yeast tubulin with yeast TOGs.

Mutations of charged residues in the regions near to the tubulin binding loops led to a reduction in the association rate for diffusion to either the native binding position or the transient complex, showing that the maximum rate of electrostatically-steered diffusion to capture required not only appropriate electrostatics of the binding loops themselves, but also supplementary electrostatic features of the host TOG.

Occupancy maps of rotational and translational diffusion for trajectories that reached a bound state showed translational steering over large distances and rotational steering at distances close to the final bound state (Fig 3.7 & 3.8). This pattern allows TOGs to be recruited initially at large contact distances with a wide cone angle of acceptance, but then to approach the final collision complex along a progressively more similar translational and rotational trajectories as the TOG-tubulin radial separation distance reduces.

The crystal structures of Stu2 TOG1 and TOG2 bound to $\alpha\beta$-tubulin revealed that these domains bind across both tubulin subunits, refuting earlier proposals of the binding of a second TOG domain. Our simulations thus support results from gel filtration, site-directed mutagenesis and crystallography, all of which point to a single site mode. This is also consistent with the results of analytical ultracentrifugation of a Stu2 TOG1-TOG2 construct with $\alpha\beta$-tubulin showing fast exchange between 1:1 and 1:2 complexes (Ayaz et al., 2014). The binding site we identify by simulated steered diffusion-to-capture matches closely, but not perfectly, to the crystallographic site Fig 2.1A.

Previous crystallographic studies suggest that TOG domains bind preferentially to curved tubulin in free solution, while experimental studies show they are unable to bind to straight, lattice-incorporated tubulin. The density maps calculated in Fig 3.3 show that straight tubulin dimers do possess the required electrostatic field to allow binding of TOG domains, although it is possible that the electrostatic field is altered to prevent binding when the tubulin is incorporated into the MT lattice. This would explain the observation that TOG-domains bind preferentially to curved tubulin rather than to MTs.

Individual TOG domains within the TOG array are known to differentially contribute to the protein's ability to accelerate tubulin incorporation (Widlund et al., 2011). Inactivating mutations of the individual TOG domains studied so far have shown that TOG1 and TOG2 contribute most strongly to polymerase activity whilst TOG3, TOG4 and TOG5 have marginal effects on activity. Specifically, constructs for Msps TOG1-2 and TOG3-4 show differential tubulin binding and MT polymerisation activities (Fox et al., 2014). Gel filtration showed that Msps TOG1-2 produced a peak elution indicative of a stable complex, while Msps TOG 3-4 failed to detectably bind tubulin. However, the TOG3-4 complex showed some polymerisation activity suggesting that Msps TOG3-4 can bind tubulin, at least transiently. Further evidence for position-dependent function of TOG-domains is derived from the alignment of multiple TOG sequences. This shows that TOG-domains in a particular array position have a high degree of cross-species conservation that is not seen between the TOG domains of each individual species. Although each TOG-domain has distinct structural and functional properties it is known that these domains function cooperatively. In future simulation work, it should be possible to explore the specific role of each TOG within an array.

The simulated rates for TOG domains associating to their cognate tubulin in general are an order of magnitude higher than to porcine tubulin, with less variability across species. Association rates to a transient complex with cognant tubulin were also consistently faster across species, with TOG1-2 exhibiting the greatest rates. Stu2p has also been shown to accelerate MT growth to a greater extent with yeast tubulin than porcine tubulin (Podolski et al., 2014). This suggests that porcine tubulin may not be the correct substrate to test the activity of all MT polymerases, given the divergence in sequence between these species.

TOG-domain proteins are thought to act as catalysts to promote the acceleration of MT growth. Experiments on GMPCPP-tubulin, a slowly hydrolysable form of GTP, showed that XMAP215 depolymerises MTs even in the absence of an energy source (Widlund et al., 2011). This suggests that XMAP215 accelerates growth by catalytically increasing the exchange rate for GTP-tubulin by at least five-fold.

The growth of MTs by the incorporation of tubulin subunits can be represented by the following kinetic description:

$$T + T_n \underset{k_d}{\overset{k_a[T]}{\rightleftharpoons}} T_{n+1} \tag{5.1}$$

The growth reaction can be further subdivided into an intermediate state or collision complex ($T{\cdot}T_n$) as follows:

$$T + T_n \underset{k_d}{\overset{k_a[T]}{\rightleftharpoons}} T_{n+1} \tag{5.2}$$

In this schema, the formation and dissociation of the collision complex is thought to be fast, while the tubulin incorporation step into the MT is slow.

$$T + XT_n \underset{k'_{-1}}{\overset{k'_1[T]}{\rightleftharpoons}} XT \cdot T_n \underset{k'_{-2}}{\overset{k'_2}{\rightleftharpoons}} XT_{n+1} \tag{5.3}$$

The presence of the catalyst (X) bound to the end of the MT works to stabilise the intermediate state in the polymerisation reaction by slowing the dissociation of the collision complex. Many tubulin dimers collide with protofilament ends and dissociate rather than becoming incorporated into the MT lattice. TOG-polymerases are proposed to aid in the incorporation of weakly-bound tubulin that would otherwise dissociate from the MT.

The best mechanistic model proposed to date, which incorporates catalyst-like properties and the preferential binding of TOG domains to curved tubulin, is a simple tethering mechanism (Ayaz et al., 2014). First the polymerase localises to the MT tip by binding to a curved $\alpha\beta$-tubulin that has no lateral neighbours. These longitudinal-bound $\alpha\beta$-tubulins, in absence of the polymerase, tend to dissociate more readily than tubulins associated with neighbouring contacts. A second $\alpha\beta$-tubulin is then tethered to another TOG-domain. Over-shortening the TOG1-2 linker has been shown to reduce polymerisation activity, consistent with a requirement for simultaneous tubulin interaction. The trapping of singleton tubulins is likely to fray the MT tips and potentially create additional more favourable sites of interactions. Maximal polymerase activity is likely to result from saturation of binding sites at the plus-end, rather than a range of tethered association rates.

For TOG polymerases to function correctly, TOGs must bind tubulin rapidly but also release it rapidly once the tubulin has been incorporated. The measured dissociation constant for the TOG:$\alpha\beta$-tubulin complex falls into the range of 70 nM to 160 nM for TOG1 and TOG2. These values gave an association rate of between 0.6 – 1.4 μM^{-1} (1/Kd) which falls into the lower limit for a diffusion limited reaction.

The simulations in this thesis describe the transient complex between the TOG-domain and tubulin as follows:

$$T + X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} T \cdot X \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} TX \tag{5.4}$$

This scheme, and our simulations, refer to the binding of free tubulin to TOG domains and may be different from tubulin capture by TOG domains within a native TOG array, with the TOG protein tip-tracking on the MT. In future work, this more complex situation should be accessible. The TransComp method (Fig. 3.4) gives additional confidence in the fundamental conclusion of this work which is that electrostatic steering plays a predominant role in the mechanism of TOG-catalysed microtubule assembly. TransComp-predicted, electrostatically-driven association constants are significantly higher (10^9) than for diffusion-limited protein association.

In equation 5.4, there is no conformational change upon TOG-tubulin binding. We have not explicitly simulated TOG-tubulin dissociation, but in this case the K_d would need to be in the order of 10^{-6} to be consistent with observed data. TOG-tubulin interactions would the be characterized both by a high, electrostatically-steered rate of tubulin capture, and by a high dissociation rate.

We have already discussed evidence for conformational selection, in which TOG binding selects a particular tubulin conformation from solution. It is possible nonetheless that tubulin undergoes a conformational change following initial capture of the TOG domain. This transition to a more stable state would create an additional population of TOG-bound tubulin in equilibrium with the rapidly-exchangeable encounter complex. With a sufficiently fast rate constant for reversion to the encounter complex, this putative tightly bound state might usefully increase the total pool of TOG-bound tubulin and thereby enhance catalysis.

Evidence for the existence of such a conformational change of the initial encounter complex is suggested from the basal association rate, with no external force, which has a calculated association rate of 10^5 $M^{-1}s^{-1}$. This is slower than the observed rate of 10^6 $M^{-1}s^{-1}$, calculated from gel filtration experiments, as expected where electrostatics plays an important role in binding. However, the results from BD simulations put the association rate constant closer to a value of 10^8 $M^{-1}s^{-1}$ – higher than experimentally observed values. This suggests that an additional factor could be in play in reducing the rate to experimentally observed values. The most likely mechanism would be a conformational change upon TOG-tubulin binding. Experimentally, an excess of the TOG polymerase Alp14 had the effect of sequestering

tubulin (pig tubulin) in *in vitro* assays . Tubulin purification by TOG binding (Al-Bassam et al., 2012) demonstrates that the TOG-tubulin complex is moderately stable and will only unbind rapidly at high ionic strengths. Both these observations suggest that strong (stable) binding, distinct from the encounter complex, may occur. One possible scenario is that following initial capture, the TOG-tubulin complex does transition into a relatively stable state, which then requires a conformational straightening of tubulin that is driven by incorporation into the MT lattice. The MT would then be the driver for fast TOG dissociation.

In summary, from the simulated association rates these observations could present two extreme scenarios for the catalysis of tubulin by TOG domains. One, rapid tubulin exchange on and off of the TOG occurs in solution with conformational selection and catalysis due to boosting the location tubulin concentration around the MT tip through fast interactions with a transient complex. Alternatively, the tubulin undergoes a stable transition following TOG binding and this creates a stable long-term complex that requires the MT to dissociate. Here the MT acts as the catalysis and the TOG facilitates activation of the mechanism-pathway.

Key to the further understanding of the mechanism of the TOG domain MT polymerases will be characterisation of the TOG-tubulin dissociation process. This could be calculated by coarse-grained (CG) MD simulations of protein-protein binding with the MARTINI force field (Marrink et al., 2007). CG would be required to allow efficient sampling of the conformational space, especially for fluctuations in the tubulin dimers. Further MD methods could include steered MD, which could explore the force response to TOG unbinding, and enhanced sampling techniques such as Markov models to explore the detailed binding pathway.

This thesis has found that electrostatically-biased diffusion is an important mechanism for modulating differential TOG-tubulin binding within the TOG array. For the lower eukaryotic TOG family members, a simple tethered exchange model can account for many aspects of the catalytic behavior. For higher family members such as ch-TOG, mechanisms of TOG-tubulin interaction are still completely unclear and this presents an exciting challenge to researchers in the future.

Bibliography

Aiken, Jayne, David Sept, Michael Costanzo, et al. (2014). Genome-wide analysis reveals novel and discrete functions for tubulin carboxy-terminal tails. *Current biology : CB* 24.12, pp. 1295–1303.

Al-Bassam, Jawdat, Mark van Breugel, Stephen C Harrison, and Anthony Hyman (2006). Stu2p binds tubulin and undergoes an open-to-closed conformational change. *Journal of Cell Biology* 172.7, pp. 1009–1022.

Al-Bassam, Jawdat, Nicholas A Larsen, Anthony A Hyman, and Stephen C Harrison (2007). Crystal structure of a TOG domain: conserved features of XMAP215/Dis1-family TOG domains and implications for tubulin binding. *Structure/Folding and Design* 15.3, pp. 355–362.

Al-Bassam, Jawdat, Hwajin Kim, Gary Brouhard, et al. (2010). CLASP promotes microtubule rescue by recruiting tubulin dimers to the microtubule. *Developmental Cell* 19.2, pp. 245–258.

Al-Bassam, Jawdat, Hwajin Kim, Ignacio Flor-Parra, et al. (2012). Fission yeast Alp14 is a dose-dependent plus end-tracking microtubule polymerase. *Molecular biology of the cell* 23.15, pp. 2878–2890.

Alberts, Bruce, Alexander Johnson, Julian Lewis, et al. (2014). *Molecular Biology of the Cell, Sixth Edition*. Garland Science.

Alsallaq, Ramzi and Huan-Xiang Zhou (2007a). Energy landscape and transition state of protein-protein association. *Biophysical Journal* 92.5, pp. 1486–1502.

– (2007b). Prediction of protein-protein association rates from a transition-state theory. *Structure/Folding and Design* 15.2, pp. 215–224.

– (2008). Electrostatic rate enhancement and transient complex of protein-protein association. *Proteins: Structure, Function, and Bioinformatics* 71.1, pp. 320–335.

Alushin, Gregory M, Gabriel C Lander, Elizabeth H Kellogg, et al. (2014). High-resolution microtubule structures reveal the structural transitions in $\alpha\beta$-tubulin upon GTP hydrolysis. *Cell* 157.5, pp. 1117–1129.

Amos, L and A Klug (1974). Arrangement of subunits in flagellar microtubules. *Journal of cell science* 14.3, pp. 523–549.

Andrade, M A and P Bork (1995). HEAT repeats in the Huntington's disease protein. *Nature genetics* 11.2, pp. 115–116.

Andrade, M A, C Petosa, S I O'Donoghue, C W Müller, and P Bork (2001). Comparison of ARM and HEAT protein repeats. *Journal of molecular biology* 309.1, pp. 1–18.

Asbury, Charles L, Jerry F Tien, and Trisha N Davis (2011). Kinetochores' gripping feat: conformational wave or biased diffusion? *Trends in Cell Biology* 21.1, pp. 38–46.

Ayaz, Pelin, Xuecheng Ye, Patrick Huddleston, Chad A Brautigam, and Luke M Rice (2012). A TOG:$\alpha\beta$-tubulin complex structure reveals conformation-based mechanisms for a microtubule polymerase. 337.6096, pp. 857–860.

Ayaz, Pelin, Sarah Munyoki, Elisabeth A Geyer, et al. (2014). A tethered delivery mechanism explains the catalytic action of a microtubule polymerase. *eLife* 3, e03069.

Baas, P W (1999). Microtubules and neuronal polarity: lessons from mitosis. *Neuron* 22.1, pp. 23–31.

Baker, N A, D Sept, S Joseph, M J Holst, and J A McCammon (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* 98.18, pp. 10037–10041.

Bieling, Peter, Liedewij Laan, Henry Schek, et al. (2007). Reconstitution of a microtubule plus-end tracking system in vitro. *Nature* 450.7172, pp. 1100–1105.

Blomberg, N, R R Gabdoulline, M Nilges, and R C Wade (1999). Classification of protein sequences by homology modeling and quantitative analysis of electrostatic similarity. *Proteins: Structure, Function, and Bioinformatics* 37.3, pp. 379–387.

Bowne-Anderson, Hugo, Marija Zanic, Monika Kauer, and Jonathon Howard (2013). Microtubule dynamic instability: a new model with coupled GTP hydrolysis and multistep catastrophe. *BioEssays* 35.5, pp. 452–461.

Bremer, A and U Aebi (1992). The structure of the F-actin filament and the actin molecule. *Current Opinion in Cell Biology* 4.1, pp. 20–26.

Brouhard, Gary J, Jeffrey H Stear, Tim L Noetzel, et al. (2008). XMAP215 is a processive microtubule polymerase. *Cell* 132.1, pp. 79–88.

Burns, R G and C Surridge (1990). Analysis of beta-tubulin sequences reveals highly conserved, coordinated amino acid substitutions. Evidence that these 'hot spots' are directly involved in the conformational change required for dynamic instability. *FEBS letters* 271.1-2, pp. 1–8.

Caplow, M and J Shanks (1996). Evidence that a single monolayer tubulin-GTP cap is both necessary and sufficient to stabilize microtubules. *Molecular biology of the cell* 7.4, pp. 663–675.

Caplow, Michael and Lanette Fee (2002). Dissociation of the tubulin dimer is extremely slow, thermodynamically very unfavorable, and reversible in the absence of an energy source. *Molecular biology of the cell* 13.6, pp. 2120–2131.

Carlier, M F and D Pantaloni (1981). Kinetic analysis of guanosine 5'-triphosphate hydrolysis associated with tubulin polymerization. *Biochemistry* 20.7, pp. 1918–1924.

Carlier, M F, T L Hill, and Y Chen (1984). Interference of GTP hydrolysis in the mechanism of microtubule assembly: an experimental study. *Proceedings of the National Academy of Sciences of the United States of America* 81.3, pp. 771–775.

Cassimeris, L, N K Pryer, and E D Salmon (1988). Real-time observations of microtubule dynamic instability in living cells. *Journal of Cell Biology* 107.6 Pt 1, pp. 2223–2231.

Cassimeris, L, D Gard, P T Tran, and H P Erickson (2001). XMAP215 is a long thin molecule that does not increase microtubule stiffness. *Journal of cell science* 114.Pt 16, pp. 3025–3033.

Charrasse, S, M Schroeder, C Gauthier-Rouviere, et al. (1998). The TOGp protein is a new human microtubule-associated protein homologous to the Xenopus XMAP215. *Journal of cell science* 111 (Pt 10), pp. 1371–1383.

Chrétien, D and S D Fuller (2000). Microtubules switch occasionally into unfavorable configurations during elongation. *Journal of molecular biology* 298.4, pp. 663–676.

Chrétien, D, S D Fuller, and E Karsenti (1995). Structure of growing microtubule ends: two-dimensional sheets close into tubes at variable rates. *Journal of Cell Biology* 129.5, pp. 1311–1328.

Coombes, Courtney E, Ami Yamamoto, Madeline R Kenzie, David J Odde, and Melissa K Gardner (2013). Evolving tip structures can explain age-dependent microtubule catastrophe. *Current biology : CB* 23.14, pp. 1342–1348.

Cooper, Christopher D, Jaydeep P Bardhan, and L A Barba (2013). A biomolecular electrostatics solver using Python, GPUs and boundary elements that can handle solvent-filled cavities and Stern layers. *arXiv.org* 3, pp. 720–729. arXiv: `1309.4018v1 [physics.comp-ph]`.

Dai, K, A Mukherjee, Y Xu, and J Lutkenhaus (1994). Mutations in ftsZ that confer resistance to SulA affect the interaction of FtsZ with GTP. *Journal of bacteriology* 176.1, pp. 130–136.

Dehouck, Yves and Alexander S Mikhailov (2013). Effective harmonic potentials: insights into the internal cooperativity and sequence-specificity of protein dynamics. *arXiv.org* 8, e1003209. arXiv: `1304.1385v1 [q-bio.BM]`.

Dolinsky, Todd J, Paul Czodrowski, Hui Li, et al. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Research* 35.Web Server issue, W522–5.

Ermak, Donald L and J A McCammon (1978). Brownian dynamics with hydrodynamic interactions. *The Journal of Chemical Physics* 69.4, p. 1352.

Fan, J, A D Griffiths, A Lockhart, R A Cross, and L A Amos (1996). Microtubule minus ends can be labelled with a phage display antibody specific to alpha-tubulin. *Journal of molecular biology* 259.3, pp. 325–330.

Fox, Jaime C, Amy E Howard, Joshua D Currie, Stephen L Rogers, and Kevin C Slep (2014). The XMAP215 family drives microtubule polymerization using a structurally diverse TOG array. *Molecular biology of the cell* 25.16, pp. 2375–2392.

Garcia, Miguel Angel, Nirada Koonrugsa, and Takashi Toda (2002). Spindle-kinetochore attachment requires the combined action of Kin I-like Klp5/6 and Alp14/Dis1-MAPs in fission yeast. *The EMBO journal* 21.22, pp. 6015–6024.

Gard, David L, Bret E Becker, and S Josh Romney (2004). MAPping the eukaryotic tree of life: structure, function, and evolution of the MAP215/Dis1 family of microtubule-associated proteins. *International review of cytology* 239, pp. 179–272.

Gardner, Melissa K, Alan J Hunt, Holly V Goodson, and David J Odde (2008). Microtubule assembly dynamics: new insights at the nanoscale. *Current Opinion in Cell Biology* 20.1, pp. 64–70.

Gardner, Melissa K, Blake D Charlebois, Imre M Jánosi, et al. (2011). Rapid microtubule self-assembly kinetics. *Cell* 146.4, pp. 582–592.

Gardner, Melissa K, Marija Zanic, and Jonathon Howard (2013). Microtubule catastrophe and rescue. *Current Opinion in Cell Biology* 25.1, pp. 14–22.

Gebremichael, Yeshitila, Jhih-Wei Chu, and Gregory A Voth (2008). Intrinsic bending and structural rearrangement of tubulin dimer: molecular dynamics simulations and coarse-grained analysis. *Biophysical Journal* 95.5, pp. 2487–2499.

Georges, Amédée des, Miho Katsuki, Douglas R Drummond, et al. (2008). Mal3, the Schizosaccharomyces pombe homolog of EB1, changes the microtubule lattice. *Nature Structural & Molecular Biology* 15.10, pp. 1102–1108.

Gorham, Ronald D, Chris A Kieslich, and Dimitrios Morikis (2011). Electrostatic clustering and free energy calculations provide a foundation for protein design and optimization. *Annals of Biomedical Engineering* 39.4, pp. 1252–1263.

Grafmüller, Andrea and Gregory A Voth (2011). Intrinsic Bending of Microtubule Protofilaments. *Structure/Folding and Design* 19.3, pp. 409–417.

Grant, Barry J, Ana P C Rodrigues, Karim M ElSawy, J Andrew McCammon, and Leo S D Caves (2006). Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* 22.21, pp. 2695–2696.

Grishchuk, Ekaterina L and J Richard McIntosh (2006). Microtubule depolymerization can drive poleward chromosome motion in fission yeast. *The EMBO journal* 25.20, pp. 4888–4896.

Grishchuk, Ekaterina L, Maxim I Molodtsov, Fazly I Ataullakhanov, and J Richard McIntosh (2005). Force production by disassembling microtubules. *Nature* 438.7066, pp. 384–388.

Hammond, Jennetta W, Dawen Cai, and Kristen J Verhey (2008). Tubulin modifications and their cellular functions. *Current Opinion in Cell Biology* 20.1, pp. 71–76.

Hemsath, Lars, Radovan Dvorsky, Dennis Fiegen, Marie-France Carlier, and Mohammad Reza Ahmadian (2005). An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. *Molecular Cell* 20.2, pp. 313–324.

Hill, T L (1985). Theoretical problems related to the attachment of microtubules to kinetochores. *Proceedings of the National Academy of Sciences* 82.13, pp. 4404–4408.

Hotani, H and T Horio (1988). Dynamics of microtubules visualized by darkfield microscopy: treadmilling and dynamic instability. *Cell motility and the cytoskeleton* 10.1-2, pp. 229–236.

Howard, Amy E, Jaime C Fox, and Kevin C Slep (2015). Drosophila melanogaster mini spindles TOG3 utilizes unique structural elements to promote domain stability and maintain a TOG1- and TOG2-like tubulin-binding surface. *Journal of Biological Chemistry* 290.16, pp. 10149–10162.

Huber, Gary A and J Andrew McCammon (2010). Browndye: A Software Package for Brownian Dynamics. *Computer physics communications* 181.11, pp. 1896–1905.

Hyman, A A, S Salser, D N Drechsel, N Unwin, and T J Mitchison (1992). Role of GTP hydrolysis in microtubule dynamics: information from a slowly hydrolyzable analogue, GMPCPP. *Molecular biology of the cell* 3.10, pp. 1155–1167.

Jánosi, I M, D Chrétien, and H Flyvbjerg (1998). Modeling elastic properties of microtubule tips and walls. *European biophysics journal : EBJ* 27.5, pp. 501–513.

Kappel, Christian, Ulrich Zachariae, Nicole Dölker, and Helmut Grubmüller (2010). An unusual hydrophobic core confers extreme flexibility to HEAT repeat proteins. *Biophysical Journal* 99.5, pp. 1596–1603.

Katsuki, Miho, Douglas R Drummond, Michael Osei, and Robert A Cross (2009). Mal3 masks catastrophe events in Schizosaccharomyces pombe microtubules by inhibiting shrinkage and promoting rescue. *Journal of Biological Chemistry* 284.43, pp. 29246–29250.

Katsuki, Miho, Douglas R Drummond, and Robert A Cross (2014). Ectopic A-lattice seams destabilize microtubules. *Nature Communications* 5, p. 3094.

Kinoshita, K, I Arnal, A Desai, D N Drechsel, and A A Hyman (2001). Reconstitution of physiological microtubule dynamics using purified components. *Science* 294.5545, pp. 1340–1343.

Kinoshita, Kazuhisa, Bianca Habermann, and Anthony A Hyman (2002). XMAP215: a key component of the dynamic microtubule cytoskeleton. *Trends in Cell Biology* 12.6, pp. 267–273.

Kippert, Fred and Dietlind L Gerloff (2009). Highly sensitive detection of individual HEAT and ARM repeats with HHpred and COACH. *PloS one* 4.9, e7148.

Kononova, Olga, Yaroslav Kholodov, Kelly E Theisen, et al. (2014). Tubulin bond energies and microtubule biomechanics determined from nanoindentation in silico. *Journal of the American Chemical Society* 136.49, pp. 17036–17045.

Krauhs, E, M Little, T Kempf, et al. (1981). Complete amino acid sequence of beta-tubulin from porcine brain. *Proceedings of the National Academy of Sciences of the United States of America* 78.7, pp. 4156–4160.

Krebs, Angelika, Kenneth N Goldie, and Andreas Hoenger (2005). Structural rearrangements in tubulin following microtubule formation. *EMBO reports* 6.3, pp. 227–232.

Leano, Jonathan B, Stephen L Rogers, and Kevin C Slep (2013). A cryptic TOG domain with a distinct architecture underlies CLASP-dependent bipolar spindle formation. *Structure (London, England : 1993)* 21.6, pp. 939–950.

Lowe, J, H Li, K H Downing, and E Nogales (2001). Refined structure of alpha beta-tubulin at 3.5 A resolution. *Journal of molecular biology* 313.5, pp. 1045–1057.

Luty, Brock A, J Andrew McCammon, and Huan-Xiang Zhou (1992). Diffusive reaction rates from Brownian dynamics simulations: Replacing the outer cutoff surface by an analytical treatment. *The Journal of Chemical Physics* 97.8, p. 5682.

MacNeal, R K and D L Purich (1978). Stoichiometry and role of GTP hydrolysis in bovine neurotubule assembly. *The Journal of biological chemistry* 253.13, pp. 4683–4687.

Maddox, Paul, Arshad Desai, Karen Oegema, Timothy J Mitchison, and E D Salmon (2002). Poleward microtubule flux is a major component of spindle dynamics and anaphase a in mitotic Drosophila embryos. *Current Biology* 12.19, pp. 1670–1674.

Mandelkow, E M, R Schultheiss, R Rapp, M Müller, and E Mandelkow (1986). On the surface lattice of microtubules: helix starts, protofilament number, seam, and handedness. *The Journal of Cell Biology* 102.3, pp. 1067–1073.

Mandelkow, E M, E Mandelkow, and R A Milligan (1991). Microtubule dynamics and microtubule caps: a time-resolved cryo-electron microscopy study. *Journal of Cell Biology* 114.5, pp. 977–991.

Margolin, Gennady, Ivan V Gregoretti, Trevor M Cickovski, et al. (2012). The mechanisms of microtubule catastrophe and rescue: implications from analysis of a dimer-scale computational model. *Molecular biology of the cell* 23.4, pp. 642–656.

Marrink, Siewert J, H Jelger Risselada, Serge Yefimov, D Peter Tieleman, and Alex H de Vries (2007). The MARTINI force field: coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B* 111.27, pp. 7812–7824.

Maurer, Sebastian P, Peter Bieling, Julia Cope, Andreas Hoenger, and Thomas Surrey (2011). GTPgammaS microtubules mimic the growing microtubule end structure recognized by end-binding proteins (EBs). *Proceedings of the National Academy of Sciences* 108.10, pp. 3988–3993.

Maurer, Sebastian P, Franck J Fourniol, Gergő Bohner, Carolyn A Moores, and Thomas Surrey (2012). EBs recognize a nucleotide-dependent structural cap at growing microtubule ends. *Cell* 149.2, pp. 371–382.

McIntosh, J Richard, Mary K Morphew, Paula M Grissom, Susan P Gilbert, and Andreas Hoenger (2009). Lattice structure of cytoplasmic microtubules in a cultured Mammalian cell. *Journal of molecular biology* 394.2, pp. 177–182.

Melki, R, S Fievez, and M F Carlier (1996). Continuous monitoring of Pi release following nucleotide hydrolysis in actin or tubulin assembly using 2-amino-6-mercapto-7-methylpurine ribonucleoside and purine-nucleoside phosphorylase as an enzyme-linked assay. *Biochemistry* 35.37, pp. 12038–12045.

Meltzer, Robert H, Errol Thompson, Kizhake V Soman, et al. (2006). Electrostatic steering at acetylcholine binding sites. *Biophysj* 91.4, pp. 1302–1314.

Mickey, B and J Howard (1995). Rigidity of microtubules is increased by stabilizing agents. *Journal of Cell Biology* 130.4, pp. 909–917.

Minoura, Itsushi, You Hachikubo, Yoshihiko Yamakita, et al. (2013). Overexpression, purification, and functional analysis of recombinant human tubulin dimer. *FEBS letters* 587.21, pp. 3450–3455.

Mitchison, T and M Kirschner (1984). Dynamic instability of microtubule growth. *Nature* 312.5991, pp. 237–242.

Mitchison, T J (1993). Localization of an exchangeable GTP binding site at the plus end of microtubules. *Science* 261.5124, pp. 1044–1047.

Mücke, N, L Kreplak, R Kirmse, et al. (2004). Assessing the Flexibility of Intermediate Filaments by Atomic Force Microscopy. *Journal of molecular biology* 335.5, pp. 1241–1250.

Müller-Reichert, T, D Chrétien, F Severin, and A A Hyman (1998). Structural changes at microtubule ends accompanying GTP hydrolysis: information from a slowly hydrolyzable analogue of GTP, guanylyl (alpha,beta)methylenediphosphonate. *Proceedings of the National Academy of Sciences of the United States of America* 95.7, pp. 3661–3666.

Nakamura, Satoko, Ilya Grigoriev, Taisaku Nogi, et al. (2012). Dissecting the Nanoscale Distributions and Functions of Microtubule-End-Binding Proteins EB1 and ch-TOG in Interphase HeLa Cells. *PloS one* 7.12, e51442.

Nogales, E, S G Wolf, and K H Downing (1998a). Structure of the alpha beta tubulin dimer by electron crystallography. *Nature* 391.6663, pp. 199–203.

Nogales, Eva, Kenneth H Downing, Linda A Amos, and Jan L we (1998b). Tubulin and FtsZ form a distinct family of GTPases. *Nature structural biology* 5.6, pp. 451–458.

Northrup, S H and H P Erickson (1992). Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proceedings of the National Academy of Sciences of the United States of America* 89.8, pp. 3338–3342.

Northrup, Scott H, Stuart A Allison, and J Andrew McCammon (1984). Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *The Journal of Chemical Physics* 80.4, pp. 1517–9.

O'Brien, E T, W A Voter, and H P Erickson (1987). GTP hydrolysis during microtubule assembly. *Biochemistry* 26.13, pp. 4148–4156.

Odde, D J, L Cassimeris, and H M Buettner (1995). Kinetics of microtubule catastrophe assessed by probabilistic analysis. *Biophysj* 69.3, pp. 796–802.

Odde, D J, H M Buettner, and L Cassimeris (1996). Spectral analysis of microtubule assembly dynamics. *AIChE journal* 42.5, pp. 1434–1442.

Oosawa, F and M Kasai (1962). A theory of linear and helical aggregations of macromolecules. *Journal of molecular biology* 4, pp. 10–21.

Pablo, P J de, I A T Schaap, F C MacKintosh, and C F Schmidt (2003). Deformation and Collapse of Microtubules on the Nanometer Scale. *Physical review letters* 91.9.

Peng, Lili X, Monica T Hsu, Massimiliano Bonomi, David A Agard, and Matthew 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, e1003464–13.

Pettersen, Eric F, Thomas D Goddard, Conrad C Huang, et al. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25.13, pp. 1605–1612.

Podolski, Marija, Mohammed Mahamdeh, and Jonathon Howard (2014). Stu2, the budding yeast XMAP215/Dis1 homolog, promotes assembly of yeast microtubules by increasing growth rate and decreasing catastrophe frequency. *Journal of Biological Chemistry* 289.41, pp. 28087–28093.

Pryer, N K, R A Walker, V P Skeen, et al. (1992). Brain microtubule-associated proteins modulate microtubule dynamic instability in vitro. Real-time observations using video microscopy. *Journal of cell science* 103 (Pt 4), pp. 965–976.

Romanowska, J, D B Kokh, and J C Fuller (2015). „Computational Approaches for Studying Drug Binding Kinetics". Wiley-VCH Verlag GmbH & Co. KGaA, pp. 211–235.

Sali, A and T L Blundell (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology* 234.3, pp. 779–815.

Schaap, Iwan A T, Carolina Carrasco, Pedro J de Pablo, Frederick C MacKintosh, and Christoph F Schmidt (2006). Elastic response, buckling, and instability of microtubules under radial indentation. *Biophysj* 91.4, pp. 1521–1531.

Schek, Henry T, Melissa K Gardner, Jun Cheng, David J Odde, and Alan J Hunt (2007). Microtubule assembly dynamics at the nanoscale. *Current Biology* 17.17, pp. 1445–1455.

Schlosshauer, Maximilian and David Baker (2004). Realistic protein-protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. *Protein Science* 13.6, pp. 1660–1669.

Sept, D, A H Elcock, and J A McCammon (1999). Computer simulations of actin polymerization can explain the barbed-pointed end asymmetry. *Journal of molecular biology* 294.5, pp. 1181–1189.

Sept, David, Nathan A Baker, and J Andrew McCammon (2003). The physical basis of microtubule structure and stability. *Protein Science* 12.10, pp. 2257–2261.

Slep, Kevin C and Ronald D Vale (2007). Structural basis of microtubule plus end tracking by XMAP215, CLIP-170, and EB1. *Molecular Cell* 27.6, pp. 976–991.

Spiegelman, B M, S M Penningroth, and M W Kirschner (1977). Turnover of tubulin and the N site GTP in Chinese hamster ovary cells. *Cell* 12.3, pp. 587–600.

Tanaka, Kozo, Etsushi Kitamura, Yoko Kitamura, and Tomoyuki U Tanaka (2007). Molecular mechanisms of microtubule-dependent kinetochore transport toward spindle poles. *Journal of Cell Biology* 178.2, pp. 269–281.

Tilney, L G, J Bryan, D J Bush, et al. (1973). Microtubules: evidence for 13 protofilaments. *The Journal of Cell Biology* 59.2 Pt 1, pp. 267–275.

Tran, P T, R A Walker, and E D Salmon (1997). A metastable intermediate state of microtubule dynamic instability that differs significantly between plus and minus ends. *Journal of Cell Biology* 138.1, pp. 105–117.

Tuszynski, Jack A, Eric J Carpenter, J T Huzil, et al. (2006). The evolution of the structure of tubulin and its potential consequences for the role and function of microtubules in cells and embryos. *The International journal of developmental biology* 50.2-3, pp. 341–358.

VanBuren, Vincent, David J Odde, and Lynne Cassimeris (2002). Estimates of lateral and longitudinal bond energies within the microtubule lattice. *Proceedings of the National Academy of Sciences of the United States of America* 99.9, pp. 6035–6040.

VanBuren, Vincent, Lynne Cassimeris, and David J Odde (2005). Mechanochemical model of microtubule structure and self-assembly kinetics. *Biophysj* 89.5, pp. 2911–2926.

Vandecandelaere, A, M Brune, M R Webb, S R Martin, and P M Bayley (1999). Phosphate release during microtubule assembly: what stabilizes growing microtubules? *Biochemistry* 38.25, pp. 8179–8188.

Vasquez, R J, D L Gard, and L Cassimeris (1994). XMAP from Xenopus eggs promotes rapid plus end assembly of microtubules and rapid microtubule polymer turnover. *Journal of Cell Biology* 127.4, pp. 985–993.

Venier, P, A C Maggs, M F Carlier, and D Pantaloni (1994). Analysis of microtubule rigidity using hydrodynamic flow and thermal fluctuations. *The Journal of biological chemistry* 269.18, pp. 13353–13360.

Walker, R A, E T O'Brien, N K Pryer, et al. (1988). Dynamic instability of individual microtubules analyzed by video light microscopy: rate constants and transition frequencies. *The Journal of Cell Biology* 107.4, pp. 1437–1448.

Walker, R A, S Inoué, and E D Salmon (1989). Asymmetric behavior of severed microtubule ends after ultraviolet-microbeam irradiation of individual microtubules in vitro. *Journal of Cell Biology* 108.3, pp. 931–937.

Wang, Zhe and Gunnar F Schröder (2012). Real-space refinement with DireX: From global fitting to side-chain improvements. *Biopolymers* 97.9, pp. 687–697.

Wasilko, David J, S Edward Lee, Kim J Stutzman-Engwall, et al. (2009). „The titerless infected-cells preservation and scale-up (TIPS) method for large-scale production of NO-sensitive human soluble guanylate cyclase (sGC) from insect cells infected with recombinant baculovirus.“ *Protein expression and purification*. Pfizer Inc., Global Research and Development Groton/New London Laboratories, Eastern Point Road, Groton, CT 06340, USA., pp. 122–132.

Weisenberg, R C, G G Borisy, and E W Taylor (1968). The colchicine-binding protein of mammalian brain and its relation to microtubules. *Biochemistry* 7.12, pp. 4466–4479.

Widlund, Per O, Jeffrey H Stear, Andrei Pozniakovsky, et al. (2011). XMAP215 polymerase activity is built by combining multiple tubulin-binding TOG domains and a basic lattice-binding region. *Proceedings of the National Academy of Sciences* 108.7, pp. 2741–2746.

Widlund, Per O, Marija Podolski, Simone Reber, et al. (2012). One-step purification of assembly-competent tubulin from diverse eukaryotic sources. *Molecular biology of the cell* 23.22, pp. 4393–4401.

Yokota, R, J P Bardhan, M G Knepley, and L A Barba (2011). Biomolecular electrostatics using a fast multipole BEM on up to 512 GPUs and a billion unknowns. *Computer physics communications* 182.6, pp. 1272–1283.

Zanic, M, J H Stear, A A Hyman, and J Howard (2009). EB1 recognizes the nucleotide state of tubulin in the microtubule lattice. *PloS one*.

Zanic, Marija, Per O Widlund, Anthony A Hyman, and Jonathon Howard (2013). Synergy between XMAP215 and EB1 increases microtubule growth rates to physiological levels. *Nature Cell Biology* 15.6, pp. 688–693.

Zhou, Huan-Xiang (2013). Influence of crowded cellular environments on protein folding, binding, and oligomerization: biological consequences and potentials of atomistic modeling. *FEBS letters* 587.8, pp. 1053–1061.