AN ABSTRACT OF THE DISSERTATION OF

Jing Jie for the degree of Doctor of Philosophy in Biochemistry and Biophysics presented on August 23, 2016.

Title: Protein Disorder in the Evolution of Dynein Regulation.

Abstract approved: _____

Elisar J. Barbar

Intrinsically disordered proteins (IDP) are a class of proteins that lack a three-dimensional structure and their prevalence and diverse functions in the cell have only been discovered relatively recently. The intermediate chain (IC) subunit of the microtubule motor protein complex dynein contains an N-terminal disordered region, N-IC, which is central to both dynein assembly and regulation. The disordered N-IC contains binding sites for dynein light chain subunits, and dynein regulatory proteins---dynactin and NudE/Nudel, whose binding sites on IC are overlapping and interactions are mutually exclusive. Light chains binding and IC phosphorylation are suggested to affect dynein-regulator interactions, yet the mechanism and evolution of dynein regulation by dynactin and NudE/Nudel are unknown. This dissertation is aimed at characterizing the interactions between N-IC, light chains, and regulatory proteins in different species to gain insights into how the interplay of protein disorder, protein phosphorylation, and alternative splicing regulates protein-protein interactions using thermodynamical and structural biophysics methods.

Chapter 1 gives an introduction to IDPs including their biophysical properties, the basis for their diverse functions, general mechanisms for interactions, and methods

to study IDPs with the emphasis on NMR spectroscopy. Chapter 2 provides background on the dynein complex including structural and assembly properties, subunit composition. Background information on the two dynein regulators of interest---dynactin and NudE/Nudel are also presented.

Chapter 3 presents studies on the structure and dynamics of yeast dynein N-IC (Pac11) and its interactions with light chain Dyn2 and dynactin subunit Nip100. We show that the N-Pac11 is primarily disordered except for an N-terminal single alpha helix (SAH) and another shorter nascent helix (LH) in the linker that separates the two Dyn2 binding sites. Both helices (SAH and LH) are required for Nip100 binding. Dyn2 binding induces structural changes in LH and alters the conformational ensemble of disordered Pac11 so that it promotes allosteric regulation of Nip100 binding. These results elucidate the role of light chains in dynein-dynactin interactions in yeast.

Chapter 4 presents studies on the structure and dynamics of mammalian (rat) dynein N-IC and the role of a single phosphorylation on different IC isoforms in dynactin subunit $p150^{Glued}$ and NudE/Nudel binding to IC. We show that rat N-IC also contains a SAH domain and a second shorter helix but with much stronger helicity comparing to Pac11. The SAH domain is necessary and sufficient for both $p150^{Glued}$ and NudE/Nudel binding, although not optimal for $p150^{Glued}$. The phosphomimetic S84D mutation, which is not required in either binding event, decreases the affinity with $p150^{Glued}$ but not NudE/Nudel. Furthermore, we show that the $p150^{Glued}$-unfavorable conformation induced by S84D requires optimal coordination between the amino acid composition and length of disordered linkers, which contributes to the specificity of phosphorylation effect on IC isoforms. Together, these results reveal a multilayered dynein regulation in mammalian species.

Chapter 5 presents studies on the evolution of differential regulation of IC by $p150^{Glued}$ and NudE/Nudel among three species. We show that folding coupled to binding and long range effects on IC are hallmarks of $p150^{Glued}$ binding; $p150^{Glued}$ and NudE/Nudel bind to non-coinciding surfaces on the SAH domain of N-IC; the differential dynein regulation mechanisms across species involve the second helix on

IC; phosphorylation and alternative splicing are evolved in mammals as factors controlling the fine-tuning of dynein-regulator interactions.

Chapter 6 presents a summary of main results and conclusions in this dissertation and proposed future work to complement our on-going efforts to elucidate the mechanism and structural basis of dynein regulation. The work presented here characterizes structural and dynamic of the intrinsically disordered region of dynein intermediate chain (N-IC) in detail, and its interactions with two regulatory proteins. Together, the results provide answers to long-standing questions in the dynein field, such as what's the effect of IC phosphorylation on dynein regulation, and how does IC bind to $p150^{Glued}$ and NudE/Nudel, and the first systematic studies on the evolution of dynein-regulator interactions.

Protein Disorder in the Evolution of Dynein Regulation

by

Jing Jie

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented August 23, 2016
Commencement June 2017

Doctor of Philosophy dissertation of Jing Jie presented on August 23, 2016.

APPROVED:

Major Professor, representing Biochemistry and Biophysics

Chair of the Department of Biochemistry and Biophysics

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Jing Jie, Author

ACKNOWLEDGEMENTS

I would like to thank my mentor Dr. Elisar Barbar sincerely and gratefully for her guidance, patience, support and understanding during the past five years. I thank her for the detailed explanations and instructions when I first started in the lab, for her encouragement and help in scientific thinking and writing, and for the independence she provides when I gradually grow into a PhD candidate. I would also like to thank everyone of the Barbar group, for all the support and help they've provided me every day, all the valuable discussions in lab meetings, and for making the lab such a comfortable place to work in. I'd like to thank Yujuan Song and Dr. Afua Nyarko, who taught me every technique for the benchwork I've done and helped me in preparing the manuscripts and proposals I've written. I'd like to thank Dr. Jessica Morgan and Ariam Kidane for the help in NMR and MALS data collection and analysis. I'd like to thank my collaborator, Dr. Frank Löhr for his outstanding expertise and generosity, and all the beautiful NMR data he collected for my experiments.

I would also like to thank my committee members, past and present, Dr. Andrew Karplus, Dr. Gary Merrill, Dr. Weihong Qiu, Dr. Barbara Taylor and Dr. Dahong Zhang, for their help, advice and encouragement. I'd like to thank Dr. Theresa Filtz and Dr. Michael Freitag for all the help when I was doing rotations in their labs. I'd like to thank the staff in the NMR and mass spectrometry facilities at Oregon State University.

I'm grateful to my parents, Yifang and Changhong, who have supported every decision I've made along the way unconditionally and always believed in me.

I thank my dearest friend, Lifang Zhang, who's always there for me through the ups and downs in graduate school. Her support and faith were the source of my courage and perseverance. I thank her for making me laugh, listening to my complaints and giving me advice in school and in life.

I'd like to thank all my friends especially Yuyu Gu for always caring for me even when I'm not being a dearest friend. I thank her for her kindness, understanding, patience and encouragement since 9 years ago when we became roommates in college.

Finally I'd like to thank the Oregon State University Graduate School Student Travel Award, the College of Science Student Travel Award and the Department of Biochemistry and Biophysics for financial assistance.

CONTRIBUTION OF AUTHORS

Dr. Frank Löhr contributed to NMR data collection in NMR data collection in Chapters 3, 4, and 5. Dr. Elisar Barbar contributed to experimental design, data analysis, and writing in all chapters.

TABLE OF CONTENTS

Page

Chapter 1: Introduction of Intrinsically Disordered Proteins 1

Definition, Biophysical Properties and Functions 2

Molecular Basis for the Diverse Functions of IDPs 6

Regulation and Evolution of IDPs 10

Methods to Study IDPs 10

Chapter 2: Introduction of Cytoplasmic Dynein 15

Functions and Structures of Cytoplasmic Dynein 16

Regulation of Cytoplasmic Dynein.....21

Chapter 3: Interactions of Yeast Dynein with Dynein Light Chain and Dynactin: General Implications for Intrinsically Disordered Duplex Scaffolds in Multi-protein Assemblies.....27

Abstract 28

Introduction.....28

Results.....30

Discussion 45

Experimental Procedures …… 54

TABLE OF CONTENTS (Continued)

Page

Acknowledgements 59

Chapter 4: Regulation of Dynein IC Involves a Novel Interplay between Phosphorylation and Disordered Spliced Linkers 60

Abstract 61

Introduction 61

Results 64

Discussion 84

Experimental Procedures 90

Acknowledgements 93

Chapter 5: Evolution of Dynein IC Regulation by Dynactin among Mammals, Drosophila, and Yeast 94

Abstract 95

Introduction 95

Results 97

Discussion 105

Experimental Procedures 112

Acknowledgements 115

Chapter 6: Summary and Future Work 116

TABLE OF CONTENTS (Continued)

Page

Summary 117

Future Work 120

Bibliography 122

Appendices 142

Appendix 1: Regulation of Dynein IC Involves a Novel Interplay between Phosphorylation and Disordered Spliced Linkers - supplemental material 143

Appendix 2: Crystallization of the Mammalian Dynein Intermediate Chain (IC) in Complex with Nudel 146

Appendix 3: Purification of Human and Drosophila Huntingtin and Its Interaction with the Dynein Intermediate Chain 148

LIST OF FIGURES

Figure Page

1.1 Typical energy landscapes of ordered proteins and IDPs 4

1.2 Spectra of two spins/conformations (A, B) in exchange 14

2.1 Schematic representation of the cytoplasmic dynein complex..... 17

2.2 Dynein regulatory protein composition and dynein recruitment pathways 23

2.3 Cryo-EM structure of the dynactin complex 24

3.1 Sequence and solution analyses of apo Pac11 1-87..... 31

3.2 Thermodynamic measurements of interactions of Pac11 1-87 and Pac11 1-87Δ66-73 with Dyn2 and with Nip100 CC1B 34

3.3 NMR titrations of Pac11 1-87 with Dyn2..... 36

3.4 NMR titrations of Pac11 1-87 with Nip100 CC1B..... 40

3.5 Elution profiles of Pac11 1-87 and Pac11 1-87Δ66-73 in ternary complexes with Nip100 and Dyn2 44

3.6 Dynamics of apo Pac11, Pac11/Dyn2 and Pac11/Nip100 46

3.7 Structural models of apo Pac11 1-87, Pac11/Dyn2 and Pac11/Nip100 50

4.1 Dynein intermediate chain from two higher eukaryotes and yeast 65

4.2 IC-2C$_{96}$ structure 66

4.3 Interactions of IC isoforms with p150* 70

4.4 Interactions of IC isoforms with nNudel..... 71

4.5 Comparison of WT and phosphomimetic IC-2C$_{96}$ interactions 73

4.6 Helix packing inferred from field-dependence of NH peak intensities and from temperature-dependence as measured by circular dichroism spectra..... 75

4.7 Conformational effects of phosphomimetic S84D mutation on IC-2C 79

4.8 Dynamics of isoform 2C in WT, phosphomimetic and truncated constructs 82

LIST OF FIGURES (Continued)

Figure Page

4.9 Model for effect of phosphomimetic mutation on dynamic structure and interaction of IC-2C_{96} 85

5.1 Dynein intermediate chain, dynactin $p150^{Glued}$ and NudE/Nudel from two higher eukaryotes and yeast 98

5.2 CD-detected secondary structure and stability of $p150^{Glued}$ and NudE/Nudel from human 101

5.3 Thermodynamic measurements of interactions of IC_{143} cysteine mutants with p150CC1 and nNudE 102

5.4 Thermodynamic measurements of interactions of $Pac11_{87}$ and $Pac11_{87}\Delta$66-73 with nNdl1 and Nip100CC1B 103

5.5 NMR titrations and dynamics of $Pac11_{87}$ with nNdl1 and comparison to those of Nip100CC1B 106

5.6 Different interacting mechanisms of the IC-p150 and IC-Nudel complexes 107

LIST OF TABLES

Table Page

2.1 Constructs used in this dissertation…… 26

3.1 Thermodynamics parameters for association of Pac11 1-87 and Pac11 1-87Δ66-73 with Dyn2 and Nip100 CC1B …… 35

4.1 Thermodynamics parameters for the IC-p150*and IC-nNudel interactions recorded at 25 °C …… 72

5.1 Thermodynamics parameters for the IC-p150 and IC-NudE interactions recorded at 25 °C …… 104

LIST OF APPENDIX FIGURES

Figure Page

A1.1 IC-2C_{44} structure 144

A1.2 SAH and H2 pack in multiple conformations in intermediate exchange in apo IC-2C_{96} 145

A2.1 Crystal formation of the IC-2C_{44}-nNudel protein complex (8 mg/ml) 147

A3.1 Purification of human Htt 536-698..... 149

A3.2 Purification of Drosophila Htt 380-481 and its interaction with dynein N-IC 151

Chapter 1

Introduction of Intrinsically Disordered Proteins

Definition, Biophysical Properties and Functions

Definition. The well-established, dominant view in the field of structural biology is that a unique three-dimensional structure is necessary for a protein to exert its biological functions. During the past two decades, the numbers of proteins and protein regions that do not possess a fixed structure and yet participate in manifold biological processes have increased greatly, which calls for re-assessment of the protein structure-function paradigm (Wright and Dyson, 1999). This new class of proteins has been called intrinsically disordered proteins (IDPs), because they lack a well-defined, folded tertiary structure under physiological conditions as a result of their amino acid sequences. The intrinsic disorder can be either global or local, and 33% of eukaryotic proteins are predicted to have long (>30 residues) intrinsically disordered regions (IDRs) (Ward et al., 2004). Instead of a single thermodynamic entity, IDPs exist in conformational ensembles that undergo rapid, dynamic exchange or interconversion among the conformers, and this disorder is critical for their functions (Tompa, 2002).

Biophysical Properties. IDPs are obviously different from globular proteins and have 'unusual' biophysical properties (Uversky, 2013a). The clue of unstructured state as an IDP's native state is embedded in its sequences, i.e. the amino acid sequence of an IDP dictates it not to fold (Williams et al., 2001). Analysis on experimentally verified IDP sequences show that combination of low mean hydrophobicity and relatively high net charge results in the intrinsic disorder (Uversky et al., 2000). In fact, IDPs often have diminished number of hydrophobic residues such as Leu, Ile, and Val and aromatic residues Phe, Tyr and Trp, which all would be frequently found in the hydrophobic core of a globular protein (Uversky, 2013a). In addition, IDPs are enriched in polar residues Gln, Asp, Ser and Thr, and charged residues Arg, Lys, Asp and Glu, as well as secondary structure breakers: Gly and Pro (Williams et al., 2001). IDPs also tend to have low number of cysteine residues that usually stabilize the protein tertiary or quaternary structures via disulfide bonds. The sequences determine the structural feature of IDPs—behaving like random coils, which includes depletion of secondary structures, high

structural flexibility, and display of extended structural characteristics such as larger hydrodynamics in gel filtration (Uversky, 2002).

Unlike a single (or a small number of) conformation (s) that defines a folded protein, IDPs sample a large number of conformations in solution and exist as conformational ensembles. This is reflection of their unique energy landscapes. Energy landscapes are used to describe the folding path of an unfolded polypeptide towards the native folded state (Jahn and Radford, 2005). For a folded protein, its folding energy landscape is like a 'funnel' (Fig. 1.1A) meaning that the unfolded polypeptide searches for the native fold on energy surfaces until the global energy minimum (the native state) is reached (Jahn and Radford, 2005). On the other hand, IDPs possess an energy landscape with no global energy minimum but many local energy minima (Fig. 1.1B) (Uversky, 2013a). This type of energy landscape is particularly sensitive to environment thus results in the conformational plasticity of IDPs (Uversky, 2013a).

Functions of IDPs in normal physiology and Disease. IDPs play important roles in a number of cellular processes, including molecular recognition, molecular assembly, protein modification, and entropic chains (Dunker et al., 2002). These functionalities can be further divided into six different functional classes of IDPs: entropic chains, display sites, chaperones, effectors, assemblers, and scavengers (Tompa, 2002). These functional classes of IDPs are frequently involved in regulation, i.e. cell-cycle regulation, transcription and translation regulation etc., and signaling pathways in the cell (Iakoucheva et al., 2002). Well studied examples include cyclin-dependent kinase regulators (CKRs) p27 and p21, the p65 subunit of the transcription factor NF-κB, and PTEN (phosphatase and tensin homolog).

Cyclin-dependent kinases (Cdks) and their regulatory proteins are essential in controlling cell cycle and division in eukaryotes (Galea et al., 2008). CKRs such as p27 and p21, which directly interact with and exert opposite effects on different Cdk-cyclin complexes via the intrinsically disordered N-terminal kinase inhibitory domain (KID), contribute to the regulation of cell cycle progression (Galea et al., 2008). Another example is the transcription factor NF-κB, which participates in the expression of

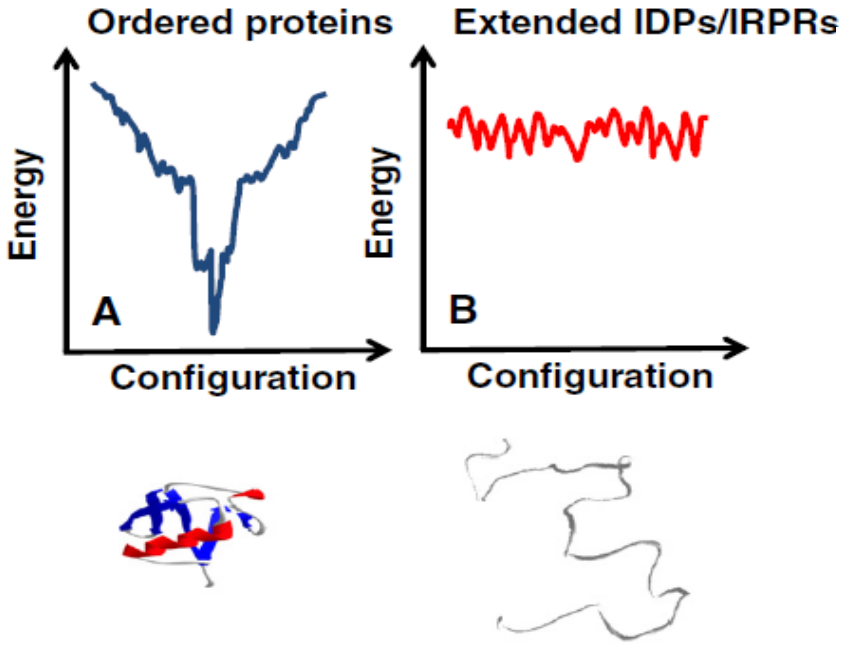


Figure 1.1: Typical energy landscapes of ordered proteins and IDPs. Shown are diagrams depicting schematically energy landscapes in one-dimensional cross-section of a globular protein (A) and of an intrinsically disordered protein (B). Shown below are the illustrative examples of corresponding protein structures. Figure adapted from Figure 5 of Uversky *et. al.* 2013a (Uversky, 2013a).

various genes involved in the cellular response to stresses and inflammation (Schmitz et al., 1994). The p65 subunit of NF-κB contains a disordered C-terminal transcription activation domain (TA_1) and studies have shown that the repeats of Asp/Glu-Phe sequence flanked by additional acidic and hydrophobic residues within the TA_1 domain are important in regulation of the transactivation by acting synergistically (Schmitz et al., 1994). IDPs are tightly involved in cell signaling pathways by their prevalence in kinases. Kinases conduct protein phosphorylation in the cell and play an important role in signal transduction, regulation of cell proliferation and growth, and apoptosis (Manning et al., 2002). Studies show that 83% of human kinases contain IDRs, which could potentially expand the kinase-kinase interaction network (Kathiriya et al., 2014). The tumor suppressor PTEN, which dephosphorylates phosphatidylinositol (3, 4, 5)-triphosphate (PIP3) to phosphatidylinositol (4,5)-bisphosphate (PIP2) (Maehama and Dixon, 1998), contains a C-terminal IDR that modulates its membrane association, activity, and stability (Malaney et al., 2013).

Intrinsic disorder is also an important feature that many scaffold proteins have. Scaffold proteins bring together specific proteins that participate in common molecular pathways and serve as a nexus thus modulate the spatial and temporal aspects of interactions among signaling proteins and pathways (Cortese et al., 2008). The tumor suppressor axis inhibition protein 1 (axin) functions as a scaffold protein for several cellular signaling pathways including the Wnt pathway (Hart et al., 1998), which is essential in modulating cell division, differentiation, motility and other developmental processes (Wodarz and Nusse, 1998). Conformational changes in the disordered region on axin are required to facilitate the interactions among several proteins in the Wnt pathway, fulfilling the scaffold functionality (Cortese et al., 2008).

As intrinsic disorder is involved in such diverse protein functions in the cell, dysfunction of IDPs leads to a number of diseases. Breast cancer type 1 susceptibility protein (BRCA1) contains a substantial amount of disorder (1480 amino acids) (Uversky et al., 2008), which is responsible for binding tumor suppressor p53, oncogenes c-Myc and JunB, and the DNA damage repair protein Rad50 (Mark et al.,

2005; Uversky et al., 2008). Another example is α-synuclein, the single protein component that aggregates to form massive insoluble fibrils under pathological conditions in the cell termed as Lewy bodies in Parkinson's disease. α-synuclein is intrinsically disordered in mammalian neurons (Theillet et al., 2016) and its misfolding eventually leads to aggregation clearly accounts for the pathology in several neurodegenerative diseases.

Molecular Basis for the Diverse Functions of IDPs

It is fascinating that IDPs without a rigid three-dimensional structure have such an extended functional repertoire. The diverse functionality stems from the intrinsic biophysical properties of IDPs, which is further expanded and regulated in the interactome involved with IDPs. Sequence and structure heterogeneity, conformational variability and adaptability, post-translational modifications, and alternative splicing all contribute to this matter.

Sequence and Structure Heterogeneity. Since the amino acid sequences of IDPs do not dictate them to fold into an ordered structure spontaneously, the sequence space of IDPs is suggested to be greater than that of globular proteins (Uversky, 2013a). The sequence of a single-domain globular protein is considered rather homogeneous in that it drives the formation of a unique well-defined structure (Uversky, 2013a). On the contrary, the sequence of an IDP is rather heterogeneous, meaning that each amino acid can be in different states in terms of rigidity and susceptibility to conformational/structural changes upon environment (pH, temperature, cleavage) and binding events (see below). Thus each amino acid could be important in the functions of IDPs and undergoing different evolution events. Structure heterogeneity, dictated by the sequence heterogeneity, is another feature of IDPs. It would be completely wrong to assume that all IDPs are structure-less. IDPs can possess different amounts of disorder, secondary structures/tertiary contacts and compactness (Dunker et al., 2001; Uversky, 2013a). Altogether, sequence and structure heterogeneity and the conformational plasticity (due to shallow energy landscapes, see above) give rise to an ultra-high degree

of structural flexibility of IDPs, referring to 1) the possibilities of adopted structures and dynamics upon environmental changes and binding events (see below); 2) the speed of interconversion among the numerous conformers in the structural ensemble (Dunker et al., 2001). How this structural flexibility contributes to the diverse functions of IDPs is discussed below.

IDPs in Protein-Protein Interaction Networks. Intrinsic disorder is shown to be a common feature of hub proteins in eukaryotes interactome networks (Haynes et al., 2006). IDPs as hub proteins can interact with a large number of cellular partners, for example, the transcriptional regulator p300 is thought to interact with 400 different partners (Gsponer and Babu, 2009). This binding promiscuity behavior, which enables IDPs to interact with various partners simultaneously to fulfill the needs of cells, is due to the structural flexibility mentioned above. The flexible nature of IDPs is also proposed to increase the recognition and association rate with partners explained by a fly-casting model (Shoemaker et al., 2000). Lack of structure and compactness is assumed to facilitate relatively rapid molecular recognition during binding events and provide larger interaction surfaces (Gsponer and Babu, 2009).

Conformational Variability and Adaptability. The interactions between an IDP and its partners cannot be considered as the traditional ‘lock and key’ process that requires fixed binding surfaces as for globular proteins (Dyson and Wright, 2002). Many IDPs undergo disorder-to-order transition or coupled folding and binding (Wright and Dyson, 2009). The short regions that undergo coupled binding and folding within a longer region of disorder are called ‘molecular recognition elements’ (MoREs) (Oldfield et al., 2005). The disorder-to-order transition can be either induced by the folded partner or selected from the IDP’s conformational ensemble (Uversky, 2013a). The MoREs can fold into different structures upon binding to different partner proteins and the conformational fluctuation of IDPs can either expose or bury the MoREs and thus modulate binding of IDPs to partner proteins (Babu et al., 2012). This conformational adaptability can contribute significantly to the interaction specificity involving IDPs (Wright and Dyson, 1999) and is also a result of the structural flexibility

mentioned above. Examples include the IDR in p53, which can adopt helical, strand, and even extend structures upon binding to different protein partners (Wright and Dyson, 2009). And the disordered GTPase-binding domain (GBD) of the Wiskott-Aldrich syndrome protein (WASP) adopts extended conformation when binding to its physiological target the Cdc42 GTPsase. However, the C-terminal VCA domain of WASP autoinhibits GBD's binding to Cdc42 GTPase by interacting with GBD and causes it to fold into a compact structure, thus forming an unfavorable conformation for the GTPase to bind (Dyson and Wright, 2002).

While many IDPs adopt ordered structure when interacting with partners, there are other IDPs and IDRs that remain disordered in the complex; these are termed as fuzzy complexes (Tompa and Fuxreiter, 2008). The preserved disorder could be 'random fuzziness' where the IDP is entirely disordered or 'flanking fuzziness' where the binding interface of the IDP is folded but the rest is still disordered (Uversky, 2013b). The interaction of fuzzy complexes is often staccato, namely that the IDP and/or its partner contains multiple weak binding sites—a 'binding cloud'(Uversky, 2013b). Those multiple weak binding sites can interact with the partner protein with equal probability and disassociate easily, therefore the fuzzy complex is a highly dynamic ensemble (Uversky, 2013b) with a number of transient interactions, and can modulate and lead to different biological functions (Sharma et al., 2015). The weak binding sites often present low complexity in sequence, thus the flexibility may allow interactions with different partner proteins (Coletta et al., 2010). Fuzzy complexes have been shown to be involved in regulation of gene expression and cell cycle, viral protein complexes and some enzymes (Sharma et al., 2015). For example, a ubiquitin ligase substrate---the N-terminal disordered region of the Hedgehog transcriptional regulator Gli3 has multiple weak binding sites for the ubiquitin ligase $CRL3^{SPOP}$. Studies show that the presence of the multiple weak binding sites in the substrate facilitates ligase recruitment and stimulates Gli3 ubiquitination in *in vitro* assays potentially through avidity effects (Pierce et al., 2016).

Post-Translational Modifications. Post-translational modifications (PTMs) such as acetylation, phosphorylation, and ubiquitination, contribute to numerous aspects of protein function regulation (Xie et al., 2007) including protein activity, localization in the cell, stability, dynamics, kinetics and degradation rate (Nishi et al., 2013). Among different types of PTMs, protein phosphorylation is one of the most abundant types in eukaryotic cells (Schlessinger, 2000). Phosphorylation sites have been shown to be located preferably in IDRs compared to ordered regions (Iakoucheva et al., 2004). Phosphorylation occurring within, near and even distant from the interaction site on an IDP often causes conformational changes that regulate the binding affinity of the IDP to its partners. Those conformational changes include disorder-order, order-disorder (Hauer et al., 1999; Vetter and Leclerc, 2001; Xu et al., 1999), altered secondary structures (Bah et al., 2015), and altered tertiary contacts that either expose or block the interaction site (He et al., 2015).

Alternative Splicing. The post-transcriptional regulation via alternative splicing (AS) of mRNAs from a single gene (Babu et al., 2012) is proposed to generate cellular complexity in eukaryotes by controlling protein expression and expanding the proteome (Lareau et al., 2004) in a time- and tissue-specific manner, facilitating cell proliferation, differentiation, and development (Romero et al., 2006). There are 92–97% of multi-exonic human genes that undergo AS (Wang et al., 2008). AS in general and tissue-specific AS in particular (Buljan et al., 2012) have been found to occur in IDR more frequently than in ordered regions to avoid disrupting protein structures, and they contribute to the functional and regulatory diversity of IDPs (Romero et al., 2006). AS could result in the inclusion or exclusion of the IDRs that have MoREs, multiple weak binding sites, and PTM sites, thus has the potential to rewire protein interaction networks (Buljan et al., 2012) and increase the conformational and functional plasticity of IDPs (Buljan et al., 2013). For example, an AS event in the transcription factor Ulthrabithorax *(UBX)* results in the splice variants that have different lengths of IDRs near the DNA binding domain and different DNA binding affinity (Buljan et al., 2013).

Regulation and Evolution of IDPs

In addition to increased protein functional versatility, disordered sequences have been shown to affect protein turnover rate without altering protein function (van der Lee et al., 2014). The lifetime, synthesis rates and abundances (Gsponer and Babu, 2012) of IDPs are tightly regulated in cell (Latysheva et al., 2015). IDPs tend to be less abundant in the cell and have lower protein synthesis rate comparing to ordered proteins (Gsponer et al., 2008). IDPs also have high proteolytic sensitivity that permits modulation by cleavage (Tompa, 2002). There are also certain IDPs that are required at high concentrations and long lifetime (Gsponer et al., 2008) because of their critical roles in signaling and protein complex assembly. Those IDPs can avoid degradation by interacting with specific partners and being part of a stable protein complex (Babu et al., 2011).

Generally, IDPs evolve more rapidly than ordered proteins (Brown et al., 2002) due to the lack of structural constraints and less restrictive residue replacements (Romero et al., 2006). Function-determining residues such as those in the MoREs, multiple weak binding sites, and PTM sites have more evolutionary constraints, like residues in ordered proteins (Pajkos et al., 2012). This rapid rate allows protein evolution to occur over few generations which is beneficial in particular for multi-cellular organisms that have long life cycles in terms of fulfilling their need of adapting to environments (Romero et al., 2006). In addition, IDPs have greater tolerance of mutations than structured proteins, and this could be favorable for a new motif to emerge through convergent evolution (Babu et al., 2012). Studies on cancer-associated proteins confirm that IDPs can allow mutations without affecting the function or causing cancer (Pajkos et al., 2012), thus are more resistant to mutation-induced phenotypes (Babu et al., 2012).

Methods to Study IDPs

Many studies of IDPs often involve computational methods. Amino acid composition bias in IDPs make disorder prediction algorithms possible (Oldfield and

Dunker, 2014). The prediction algorithms have been refined to have accuracy >80% (Varadi et al., 2015), including PONDR (Li et al., 1999; Romero et al., 2001), IUPred (Dosztányi et al., 2005), and ESpritz (Walsh et al., 2012). And there are IDP databases for IDPs detected experimentally---Disprot (Sickmeier et al., 2007) and identified using prediction algorithms from Uniprot---D^2P^2 (Oates et al., 2013).

Structural studies on IDPs are often challenging since the IDPs exist in solution in dynamic conformational ensembles instead of a single well-defined structure. This precludes X-ray crystallography as an experimental method (Varadi et al., 2015). On the other hand, circular dichroism (CD), nuclear magnetic resonance (NMR) spectroscopy and electron paramagnetic resonance (EPR) spectroscopy, ensemble Forster resonance energy transfer (FRET) (Haas, 2012), small-angle X-ray scattering (SAXS), and mass spectrometry-based techniques have been applied to study IDPs. Solution NMR is a versatile tool to study protein structures and is uniquely powerful in characterizing the conformational heterogeneity, flexibility and dynamics of IDPs (Jensen et al., 2013; Mittag and Forman-Kay, 2007). EPR has been used to investigate the induced folding of the disordered domain of measles virus nucleoprotein upon binding to the phosphoprotein (Belle et al., 2008). SAXS has been used to provide the overall shape information of proteins or protein complexes and in combination with other methods that provide local structural information such as NMR, good ensemble models can be generated (Receveur-Brechot and Durand, 2012). Mass spectrometry-based techniques including hydrogen-deuterium exchange could be a useful tool when characterizing a single conformer of interest instead of the average conformation adopted by an IDP (Beveridge et al., 2013). Both unrestrained (*de novo*) and restrained (by NMR experiments-derived ensemble) molecular dynamics (MD) simulations are used to elucidate the ensemble modeling and conformations sampled (Mittag and Forman-Kay, 2007).

NMR Spectroscopy of IDPs. In this dissertation, the major method used to study IDPs is solution NMR spectroscopy, which provides atomic-resolution information on secondary structures, tertiary contacts, transient/low populated states, solvent

accessibility, conformational exchange properties, and motions on a variety of timescales (Mittag and Forman-Kay, 2007). In addition, NMR also provide insights on the residue-specific interactions of IDPs with their partners. Different structural observables and constraints obtained by NMR such as NOEs, PREs and RDCs can be used to generate the ensemble, thus determining the structures of IDPs.

The multi-dimensional NMR experiments yielding structural constraints are built on two-dimensional heteronuclear correlation experiments, among which the ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) is the most widely used in protein NMR. HSQC spectrum provides the backbone N and H resonance-residue correlation 'fingerprint' for every amino acid except for prolines in a protein, and the assignment of the majority of HSQC peaks is the prerequisite for further advanced NMR experiments and analysis. Due to similar chemical environments for almost all residues in an IDP, its HSQC spectrum features poorly dispersed peaks and often large spectral overlap especially in the proton dimension. Increase in protein size and number of residues worsens this problem, thus creating the molecule size barrier for suitable targets for protein NMR. To overcome this, transverse relaxation optimized spectroscopy (TROSY) HSQC has been introduced (Pervushin et al., 1997), which takes advantage of the interference between dipole-dipole (DD) coupling and chemical shift anisotropy (CSA) to efficiently decrease the transverse relaxation rate of large molecules that give poor quality normal HSQC spectra (Pervushin et al., 1997). TROSY is also extremely useful in studying the behavior of IDPs in large protein complexes.

The chemical shifts of NMR signals contain inherent local structural information and are exceptionally sensitive to secondary structure and chemical environment changes, thus are considered as structure constraints to define IDP ensembles (Showalter, 2007). Chemical shift perturbation (CSP) is used to monitor binding events including the residue-specific binding site, affinity and potentially the structure of the complex (Williamson, 2013).

In NMR spectra, peak or signal intensities often provide useful information on the dynamics properties and conformational heterogeneity of IDPs. Decreased peak

intensity resulted from line broadening could be caused by several reasons, including direct binding, transverse (T_2) relaxation, and conformational exchange. When residues are at the binding interface and upon interaction they tumble together with its partner, their T_2 relaxation increases and result in peak broadening. Alternatively, due to the sequence and structure heterogeneity that many IDPs manifest, peak intensity heterogeneity is often seen in the spectrum of IDPs due to different T_2 relaxation rates. Last, in the case of conformational exchanges occurring within one region, the signal intensity is correlated to the exchange rate (K_{ex}) between conformations and the frequencies separating the two chemical shifts (Δf) corresponding to the two conformations which are correlated to the magnetic field strength (Fig. 1.2).

NMR is the only technique that can characterize protein motions at different timescales (pico-second to milli-second). T_1 relaxation, T_2 relaxation and steady-state heteronuclear NOE (nuclear Overhauser effect) measurements are the workhorse NMR dynamics experiments conducted to characterize residue-specific motion on the pico-second to nano-second timescale. Thus they provide the atomic-resolution of backbone flexibility and mobility changes upon ligand binding. Relaxation dispersion experiments monitor protein dynamics on the micro-second to milli-second timescale, and provide information on kinetics, thermodynamics, population (even as low as 0.5% populated states) as well as the structure of the protein (Neudecker et al., 2009). In order to determine the solvent accessibility of a protein in free form and in the bound form, CLEANEX-PM (Phase-Modulated chemical EXchange) experiments are used to measure the exchange rates of amide protons with water (Hwang et al., 1998).

For structure determination, the nuclear Overhauser enhancement (NOE) measurements provide short distance constraints (up to 5 Å), and the paramagnetic relaxation enhancement (PRE) measurements provide long-range distance constraints (up to 25 Å). The residual dipolar coupling (RDC) measurements provide information on the orientation of a bond vector with respect to the alignment frame (Mittag and Forman-Kay, 2007), thus together with the distance constraints above they can be used to determine the structure of IDP ensemble via computational approaches.

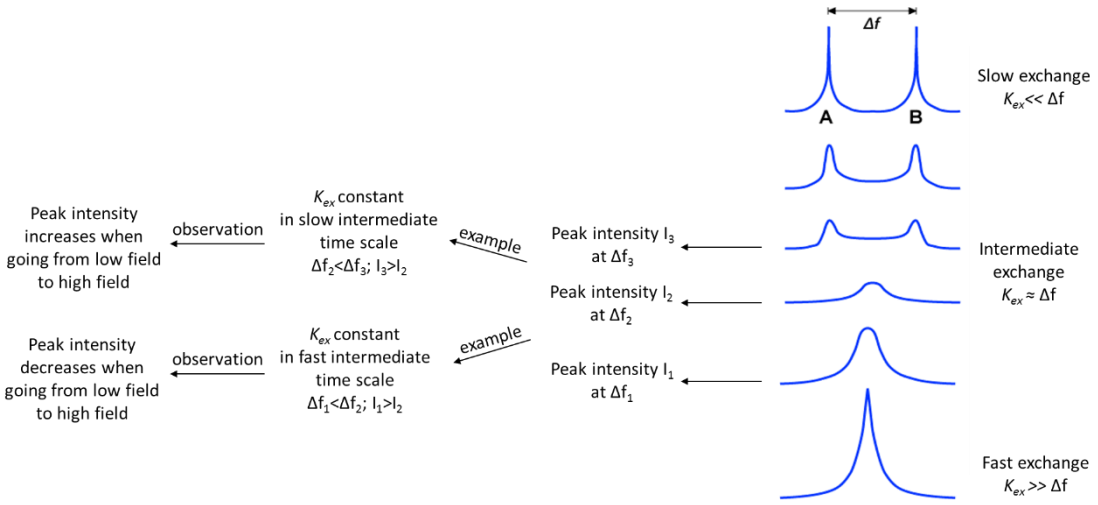


Figure 1.2: Spectra of two spins/conformations (A, B) in exchange. The exchange process is shown as a function of exchange rate K_{ex} and the differences in frequencies/chemical shifts (Δf) between A and B. In this dissertation (chapters 3 and 4), we show two examples of probing conformational exchanges by analyzing peak intensities collected at different magnetic fields. Figure adapted from Allen D Elster, MRIquestions.com.

Chapter 2

Introduction of Cytoplasmic Dynein

Intracellular trafficking defines the spatial complexity of eukaryotic cells (Fu and Holzbaur, 2014) and it depends on cytoskeletal motor proteins: myosin uses actin filaments to fulfill its tasks, and kinesin and dynein transport cellular cargos along microtubules. Unlike myosin and kinesin which have 14-17 classes of proteins to carry out different functions, dynein only falls into two classes: cytoplasmic dynein and axonemal dynein (Hook and Vallee, 2006). Axonemal dynein is essential in cilia/flagella for cell motility in lower eukaryotes and for sperm motility and airway maintenance in trachea (Ichikawa et al., 2015). This dissertation will focus only on cytoplasmic dynein.

Functions and Structures of Cytoplasmic Dynein

Cytoplasmic dynein is the major minus end-directed microtubule motor protein complex in the cell and is responsible for the retrograde transport (i.e. from cell periphery to cell center) of a wide variety of cargos, including membranous organelles such as late endosomes (Bananis et al., 2004), lysosomes and Golgi apparatus (Harada et al., 1998), vesicles, proteins, chromosomes, mRNAs, and viruses (Vale, 2003). Dynein also participates in other cellular processes including organelle positioning, mitosis, microtubule organization (Raaijmakers and Medema, 2014), and cell migration. During mitosis, dynein is involved in the separation of centrosomes, mitotic spindle assembly and orientation, and chromosome segregation (Holzbaur and Vallee, 1994). The transport functions of dynein are of the most interest in the nervous system since dynein is the single motor complex that is responsible for retrograde axonal transport of organelles as well as synaptic vesicles in the neuronal cells (Miller, 2016). Previous studies have shown that mutated dynein causes activity loss of motor neurons and eventually neurodegenerative diseases such as ALS (amyotrophic lateral sclerosis) (Ori-McKenney et al., 2010).

Dynein is a large (~1.6 MDa) multisubunit protein complex that functions as a homodimer (i.e. two copies of each subunit). Vertebrate dynein consists of six different

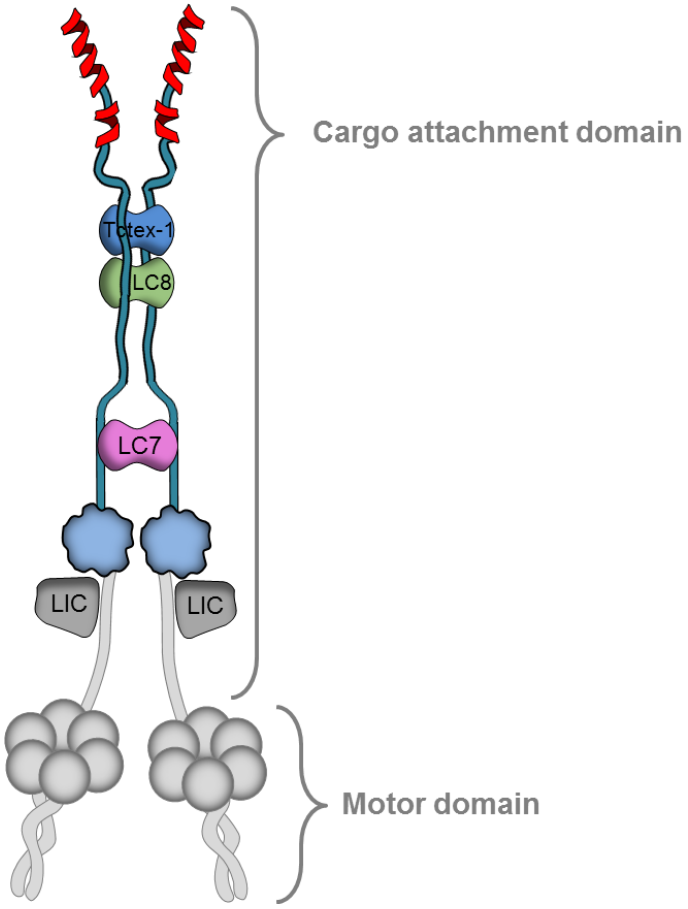


Figure 2.1: Schematic representation of the cytoplasmic dynein complex. The heavy chains are shown in light grey with their C-terminal two thirds (head domain) making up the motor domain and the N-terminal one third (tail domain) associating with the other subunits making up the cargo attachment domain. N-HC binds to LIC (grey) and C-terminal IC (blue bulb) directly. N-terminal IC (blue chain) is predominately disordered except for a single α helix (red helix) at the beginning and it provides binding sites for all 3 light chains: Tctex-1 (blue), LC8 (green) and LC7 (pink). Figure adapted from Figure 2 from Clark *et al.* 2015 (Clark et al., 2015).

subunits: heavy chain (HC), light intermediate chain (LIC), intermediate chain (IC) and 3 light chains (LC)---Tctex, LC8 and LC7 (Fig. 2.1). The HC is responsible for motor activity and has an ATPase domain and a microtubule binding domain at C-terminal two thirds in its sequence, which makes up the motor domain of dynein. The N-terminal one third of HC that contains binding sites for both LIC and IC, and the rest dynein subunits make up the cargo attachment domain.

The Dynein Heavy Chain. The heavy chain dimer is considered the core of dynein and each heavy chain (HC) protein is 470-540 kDa (Schiavo et al., 2013; Straube et al., 2001). The HC contains an N-terminal extended tail (or stem) domain that contains the dimerization domain and binding sites for cargos as well as other dynein subunits. The C-terminus of HC, the motor (or head) domain, folds into a globular ring-like domain that consists of six tandemly linked ATPase domains (AAA+) and a stalk region that contains the microtubule binding domain (MTBD) located in between AAA4 and AAA5 (Qiu et al., 2013). The motor domain of HC uses the chemical energy from ATP hydrolysis to produce the force required to power the translocation along microtubules (Asai and Koonce, 2001). The tail domain of HC is responsible for regulation of the motor activity and specification of the intracellular location through interactions with the other dynein subunits and adaptor proteins (Asai and Koonce, 2001). Recent mutational studies done on mouse and human dynein heavy chain have identified a number of single amino acid substitutions that cause malfunctions of the nervous system and development, including “Legs at odd angles” (Loa), “Cramping 1” (Cra1), and “Sprawling” (Swl) (Chen et al., 2007; Hafezparast, 2003). The structure of dynein HC motor domain dimer has been solved (Carter et al., 2011) and several studies have provided insights into the mechanism of motor activity (Burgess et al., 2003; Kon et al., 2012; Reck-Peterson et al., 2006; Roberts et al., 2009). But little is known about the structure and interactions of the HC tail domain.

The Dynein Light Intermediate Chain. In lower eukaryotes, LIC is encoded by a single gene whereas in mammals there are two LICs (LIC1 and LIC2) (Zhang et al., 2009) that undergo post-translational modifications to generate a number of LIC species

(Tynan et al., 2000). LIC is unique to cytoplasmic dynein and is essential in many aspects of dynein functions during mitosis (Yoder and Han, 2001). LIC has been shown to interact with centrosome proteins (Purohit et al., 1999) and mediate cargo binding. In addition, LIC is required for the assembly and stability of dynein complex in lower eukaryotes (Mische et al., 2008; Zhang et al., 2009). The N-terminal conserved G-protein domain of LIC adopts a Ras-like fold and interacts with the dynein heavy chain and GDP (only metazoan LICs) (Schroeder et al., 2014). The less conserved C-terminal domain binds to cargo adaptor proteins that allows membrane transport (Schroeder et al., 2014).

The Dynein Light Chains. Dynein has been found to contain three distinct light chains that are 8 kDa, 14 kDa and 22 kDa (King et al., 1996a), and all of them interact with the intermediate chain subunit in dynein assembly. The 14 kDa one was identified to be Tctex-1 (t complex-testis-expressed-1) which was originally shown to be a member of the t-complex in mouse (King et al., 1996b; Mok et al., 2001). Tctex-1 also interacts with a number of non-dynein proteins in various cellular pathways, including vesicle-binding protein Doc2 in cargo transport (Nagano et al., 1998), tyrosine kinase p50fyn (Fyn) in signaling and cytokinesis (Campbell et al., 1998), neurogenesis inhibitor Lfc (Gauthier-Fisher et al., 2009), tumor suppressor REIC/Dkk-3 (Ochiai et al., 2011). Detailed biochemical and biophysical studies revealed that Tctex-1 is a homodimer of 25 kDa (Talbott et al., 2006) and each monomer is comprised of a short N-terminal β-strand, 2 α–helices and 4 β-strands (Williams et al., 2005) and uses its β2 strand and β2/β3-loop to interact with IC (Mok et al., 2001). The binding site of Tctex-1 was mapped to a 12-amino acid fragment in the N-terminal IC (Makokha et al., 2002; Mok et al., 2001; Williams et al., 2007). The stoichiometry of the Tctex-IC interaction is 1:1 (one Tctex dimer binds to two IC chains) and the affinity was determined to be 9.6 μM in Drosophila. (Hall et al., 2009).

The 8 kDa dynein light chains was named light chain 8 (LC8) when first identified in outer arm axonemal dynein in *Chlamydomonas reinhardtii* (Pazour et al., 1998). LC8 is highly conserved among eukaryotes; the mammalian LC8 shares 94%

and 72% sequence identity with those of *Drosophila melanogaster* and *Aspergillus nidulans*, respectively. LC8 is later found to be a 20 kDa homodimer that shares a similar fold with Tctex-1 despite low sequence identity. Monomeric LC8 consists of a central β-sheet of 4 β-strands that's flanked by 2 α–helices and another β-strand crossed over from the other monomer (Liang et al., 1999). The LC8 dimer provides two identical binding sites for its partners at its dimer interface and all known LC8 partners bind as a sixth β-strand of the β-sheet in LC8 (Barbar, 2008), making it impossible to bind different partners simultaneously. The recognition sequence of LC8 is an intrinsically disordered 7-10 amino acid linear motif that contains a TQT triplet (Clark et al., 2016) and undergoes disorder-to-order transition upon LC8 binding. The LC8 binding site was mapped to residues 126-138 next to the Tctex-1 binding site on IC and the affinity was 8 μM in Drosophila, similar to that of the IC-Tctex interaction (Hall et al., 2009; Williams et al., 2007). However, when either light chain is bound first to IC forming the duplex scaffold, the affinity to the second light chain is enhanced by 50-fold, suggesting a role of polybivalency in the assembled IC (Hall et al., 2009). Another effect of LC8 binding on IC is to cause IC dimerization via a self-association domain (residues 222-231) downstream LC8 site (Nyarko and Barbar, 2011). Promoting dimerization or increasing the dimerization efficiency of partners has also been observed in interactions of LC8 and Swallow (Kidane et al., 2013; Wang et al., 2004), myosin V (Wagner et al., 2006), Nup159 (Nyarko et al., 2013; Romes et al., 2012), and Ana2 (tetramerization) (Slevin et al., 2014), fostering and supporting the hypothesis that LC8 is a hub protein acting as a dimerization engine in diverse protein networks and cellular activities (Barbar, 2008).

The third light chain is LC7, first discovered in *C. reinhardtii*; its Drosophila orthologue is roadblock and mammalian orthologue is km23 (Bowman et al., 1999; Tang et al., 2002). Km23 is shown to interact with the growth factor TGFβ, indicating a role in cell signaling and growth (Tang et al., 2002). LC7 belongs to an ancient protein superfamily (RLC7-MglB superfamily) involved in regulation of NTPase activity (Koonin and Aravind, 2000). Km23 is also a homodimer but unrelated to Tctex-1 and

LC8 in structure. Each km23 monomer contains an N-terminal short α–helix, and a five-strand β-sheet flanked by another helix (Ilangovan et al., 2005; Liu et al., 2006). The roadblock/km23 binding site on IC has been mapped to a 40-residue segment, significantly downstream of the Tctex-1 and LC8 binding sites (Hall et al., 2010; Song et al., 2005; Susalka et al., 2002), which also undergoes disorder-to-order transition and formation of an α–helix upon LC7 binding (Hall et al., 2010). Interestingly, the LC8-induced IC self-association via a segment immediately upstream the LC7 binding site is reversed by LC7 binding (Nyarko and Barbar, 2011).

The Dynein Intermediate Chain. The intermediate chain (IC) subunit is ~74 kDa and central to both dynein assembly and regulation. The conserved C-terminal half of IC (C-IC) contains a WD repeat domain that adopts a β-propeller fold. C-IC is also responsible for interacting with the HC tail, whereas the N-terminal half of IC (N-IC) binds to all 3 LCs, together forming and stabilizing the dynein complex assembly. N-IC is an IDP that lacks a fixed 3D structure and is monomeric in solution. N-IC also directly interacts with some dynein cargos (Dhani et al., 2003; Ye et al., 2000) and a number of dynein regulatory proteins including dynactin, NudE/Nudel, and huntingtin, which makes N-IC a nexus for dynein regulation (Caviston et al., 2007; Karki and Holzbaur, 1995; Wang and Zheng, 2011).

IC is present in dynein from yeast to mammals and is encoded by a single gene in invertebrates and two genes in vertebrates (Pfister et al., 2006). Various tissue-specific isoforms produced by alternative splicing to generate the complexity have been shown in both Drosophila and mammals (Kuta et al., 2010; Nurminsky et al., 1998). Only in mammals, there are spliced sites located within the N-IC region that is responsible for interacting with dynein regulators, suggesting a role of alternative splicing in modulating dynein activity.

Regulation of Cytoplasmic Dynein

To achieve the diverse cellular functions, dynein's localization, motor activity and cargo loading/unloading are tightly regulated in a temporally and spatially manner

(Tai et al., 2001). One mechanism for dynein recruitment and cargo specificity is the heterogeneity of dynein subunits. In higher eukaryotes, each dynein subunit has several isoforms that are shown to be expressed in a tissue-specific manner (Criswell and Asai, 1998; Tai et al., 2001). Post-translational modifications such as phosphorylation of those isoforms is also suggested to contribute to the different dynein complex pools.

Another mechanism for dynein regulation is achieved through dynein interactions with numerous regulatory proteins and cargo adaptors. Among those regulators, some are interacting with the heavy chain to modulate the motor activity (Markus et al., 2009; Sasaki et al., 2000), and some are interacting with the intermediate chain (IC). This dissertation will be focusing on two of the best studied IC-interacting dynein regulators—dynactin and NudE/Nudel.

Dynactin. Dynactin was first discovered as a dynein activator in vesicle transport (Gill et al., 1991) and later found out to participate in most aspects of dynein functions (Schroer, 2004). Dynactin is required for dynein accumulation at the microtubule plus ends (Zhang et al., 2003), the initiation of dynein-mediated transport in neurons (Moughamian and Holzbaur, 2012), dynein cargo recruitment (Karki and Holzbaur, 1999), and increased dynein processivity along microtubules (Culver-Hanlon et al., 2006; King and Schroer, 2000). Dynactin is a large 1.2 MDa protein complex that contains multiple copies of 11 different subunits (Schroer, 2004). The interaction between dynactin and dynein is mediated by the largest subunit of dynactin $p150^{Glued}$ (King et al., 2003), which exists in extended dimeric forms and together with other subunits make up the projecting arm of dynactin complex. $P150^{Glued}$ contains an N-terminal CAP-Gly (cytoskeleton-associated protein, glycine-rich) domain responsible for binding microtubules, and two coiled-coil domains (CC1 and CC2) (Schroer, 2004). The C-terminal half of $p150^{Glued}$ CC1 (CC1B) has been shown to bind the N-terminal disordered region of dynein intermediate chain (McKenney et al., 2011).

NMR studies on Drosophila proteins identified a multiregional recognition site on IC (residues 1-41 and 46-75) for $p150^{Glued}$ (Morgan et al., 2011). A short stretch of residues 10-44 on mammalian N-IC was shown to be sufficient for binding $p150^{Glued}$

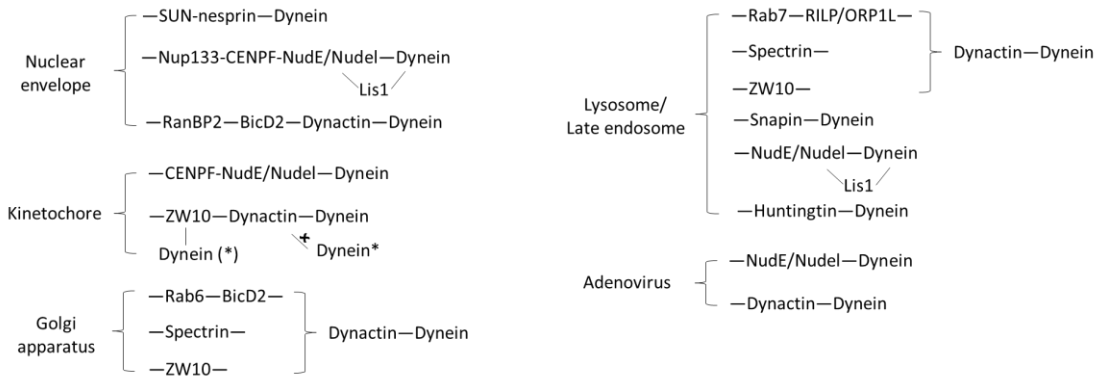


Figure 2.2. Dynein regulatory protein composition and dynein recruitment pathways. Different organization of dynein regulators are shown when recruiting dynein to different subcellular sites (Vallee et al., 2012). Both dynactin and NudE/Nudel participate in almost all of the recruitments and they are found to function in the same pathway as well. Huntingtin facilitates dynein-dynactin-mediated vesicle transport (Caviston et al., 2007) and is required for the dynein-dependent lysosome/late endosome positioning (Caviston et al., 2011). Dynein IC phosphorylation has been shown to regulate interactions with dynein regulators. For example, T89 phosphorylated IC (Dynein*) promotes binding to ZW10 at the kinetochores while decreases dynactin binding (Whyte et al., 2008).

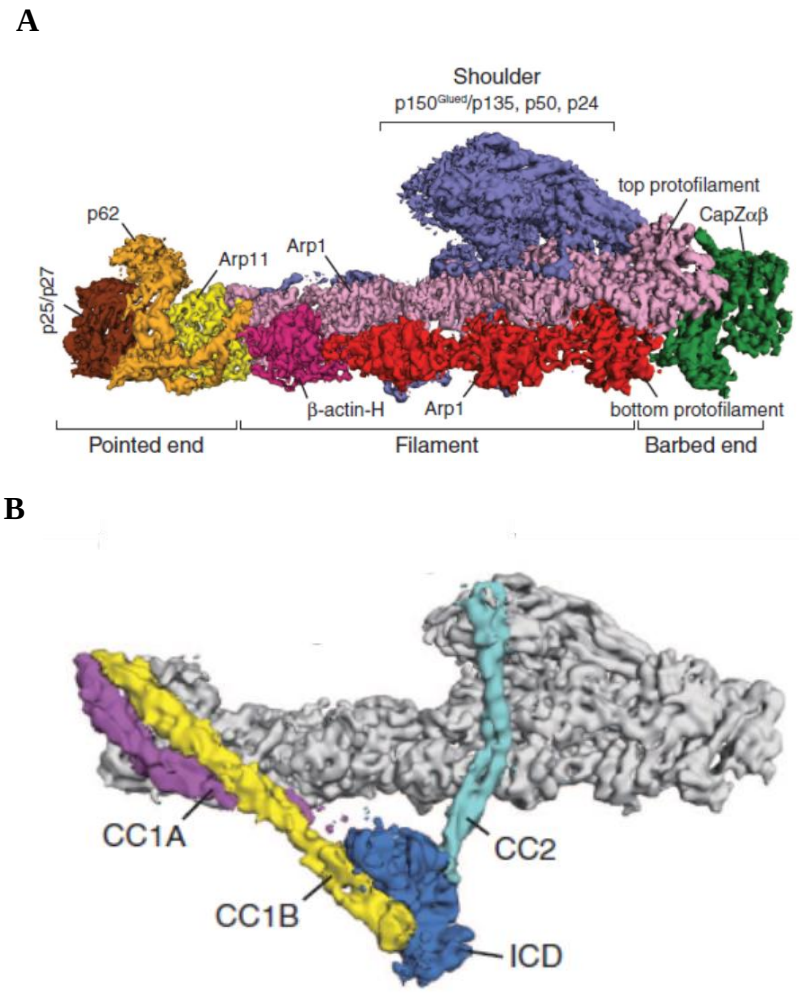


Figure 2.3: Cryo-EM structure of the dynactin complex. (A) A 4.0 Å cryo-EM map of dynactin showing the subunit composition. The overall architecture consists of a rod-like filament containing Arp1 and β-actin, a pointed end, a barbed end and a shoulder (projecting arm). (B) An 8.6 Å cryo-EM map of dynactin (grey) with a docked p150 ^{Glued} (colored) projection. Shown are the domains of p150^{Glued}: CC1A (purple), CC1B (yellow), intercoiled domain (ICD, dark blue) and CC2 (light blue). Figure adapted from Figures 1 and 3 of Urnavicius *et al.* 2015 (Urnavicius et al., 2015).

mainly through electrostatic interactions (Siglin et al., 2013). In the recent near-atomic resolution structure of dynein-dynactin complex, the p150^{Glued}-IC is missing due to the flexibility of the complex (Urnavicius et al., 2015); hence no detailed structural information on the p150^{Glued}-IC interaction is available.

NudE/Nudel. The gene for the nuclear distributive protein E (NudE) and its closely related vertebrate paralogue Nudel was first identified in *A. nidulans*, which encodes a protein that is essential for nuclear migration (Efimov and Morris, 2000). NudE is required in mitotic spindle assembly (Feng and Walsh, 2004) and Nudel is responsible for modulating cortical neuronal positioning (Shu et al., 2004). NudE/Nudel interacts with LIS1 (Lissencephaly 1), which is also a dynein regulator that binds to dynein HC (Sasaki et al., 2000). Together with LIS1, NudE/Nudel is essential in many aspects of dynein functions including microtubule organization, nuclear envelop invagination, spindle orientation (Chansard et al., 2011) and targeting dynein to the kinetochores (Lam et al., 2010) during mitosis. NudE/Nudel has a long N-terminal coiled-coil domain (Derewenda et al., 2007) and a C-terminal unstructured domain. The N-terminal coiled-coil domain of NudE/Nudel has been shown to interact with N-IC of dynein (Nyarko et al., 2012; Wang and Zheng, 2011) and the NudE binding site on N-IC has been mapped to the first 40 residues in Drosophila IC (Nyarko et al., 2012).

Constructs used in this dissertation. To investigate the mechanism for selective binding of dynein IC to dynactin p150^{Glued} and NudE/Nudel, we characterize in detail the interactions of IC-p150^{Glued} and IC-NudE/Nudel using recombinantly expressed and purified proteins from multiple species (Table 2.1). The coiled-coil domain of mammalian NudE and Nudel responsible for binding dynein IC share ~70% sequence identity, thus the interactions of mammalian IC-NudE and IC-Nudel are considered identical. For mammalian IC constructs, we used isoforms 1A and 2C from *Rattus norvegicus*; for mammalian p150^{Glued} and NudE/Nudel constructs, we used coiled-coil domains from human p150^{Glued} and human Nudel that are identical in sequence to those of the rat proteins, separately. The *Saccharomyces cerevisiae* orthologues are Pac11 (IC), Nip100 (p150^{Glued}), Ndl1 (Nudel) and Dyn2 (LC8).

Table 2.1. Constructs used in this dissertation.

Species / Protein	*Saccharomyces cerevisiae*	*D. melanogaster*	*R. norvegicus*
IC	Pac11	IC	IC-1A; IC-2C
$p150^{Glued}$	Nip100	$p150^{Glued}$	$p150^{Glued}$
NudE/Nudel	Ndl1	NudE	Nudel
LC8	Dyn2	LC8 (not used)	LC8 (not used)

Chapter 3

Interactions of Yeast Dynein with Dynein Light Chain and Dynactin: General Implications for Intrinsically Disordered Duplex Scaffolds in Multi-protein Assemblies

Jing Jie, Frank Löhr and Elisar Barbar

Published in The Journal of Biological Chemistry,
2015, 290 (39), 23863-23874

Abstract

Intrinsically disordered protein (IDP) duplexes composed of two IDP chains cross-linked by bivalent partner proteins form scaffolds for assembly of multi-protein complexes. The N-terminal domain of dynein intermediate chain (N-IC) is one such IDP that forms a bivalent scaffold with multiple dynein light chains including LC8, a hub protein that promotes duplex formation of diverse IDP partners. N-IC also binds a subunit of the dynein regulator, dynactin. Here we characterize interactions of a yeast ortholog of N-IC (N-Pac11) with yeast LC8 (Dyn2) or with the IC-binding subunit of yeast dynactin (Nip100). Residue level changes in Pac11 structure are monitored by NMR spectroscopy, and binding energetics are monitored by isothermal titration calorimetry (ITC). N-Pac11 is monomeric and primarily disordered except for a single alpha helix (SAH) at the N-terminus and a short nascent helix, LH, flanked by the two Dyn2 recognition motifs. Upon binding Dyn2, the only Pac11 residues making direct protein-protein interactions are in and immediately flanking the recognition motifs. Dyn2 binding also orders LH residues of Pac11. Upon binding Nip100, only Pac11 SAH residues make direct protein-protein interactions, but LH residues at a distant sequence position and L1 residues in an adjacent linker are also ordered. The long-distance, ligand-dependent ordering of residues reveals new elements of dynamic structure within IDP linker regions.

Introduction

Cytoplasmic dynein is an essential microtubule-based motor that controls diverse cellular processes including centrosome separation and movement, mitotic spindle assembly and orientation, and axonal transport (1, 2). In higher eukaryotes, cytoplasmic dynein is a ~1.6 MDa multi-subunit protein complex that moves along microtubules using its motor domain and transports cargos associated with its cargo attachment domain. The motor domain, part of the heavy chain subunit, includes the ATPase activity that generates force to drive dynein motion. The cargo attachment domain functions in loading of cargo, maintenance of stability and modulation of dynein

activity (Kini and Collins, 2001). In Drosophila the primary chains that compose the cargo domain are IC (intermediate chains), LC7, LC8 and Tctex1 (light chains), and the LICs (light intermediate chains).

Dynein processivity and activity require interactions with various regulatory proteins (Vallee et al., 2012), including dynactin, an essential multi-subunit complex (Schroer, 2004). Dynactin mediates cargo recruitment (Gill et al., 1991) and increases dynein processivity along microtubules (Kardon et al., 2009; King and Schroer, 2000). Since *Saccharomyces cerevisiae* dynein has structural and functional differences from Drosophila dynein, we sought in these studies to compare and contrast the consequences of dynactin and light chain binding on intermediate chain structure in the two species. In yeast, the single non-essential function of dynein is positioning the mitotic spindle into the mother/bud neck during mitosis (Adames and Cooper, 2000). Also, the yeast cargo attachment domain has a simpler subunit composition. While higher eukaryotes have one copy each of the three different light chains, yeast has two copies of one light chain, Dyn2, the LC8 ortholog. Disrupting Dyn2 binding impairs dynactin recruitment (Stuchell-Brereton et al., 2011).

In Drosophila, the specific binding of dynein to dynactin involves protein-protein interactions of IC and $p150^{Glued}$, a domain of the largest subunit of dynactin. A predicted coiled-coil domain of $p150^{Glued}$ associates with the N-terminal domain of dynein IC, also predicted to be a coiled-coil (McKenney et al., 2011). Specifically, $p150^{Glued}$ binds IC at a multi-region interface corresponding to residues 1-40 and 46-75 (Morgan et al., 2011). Tctex1 and LC8 bind the disordered IC segment 110-138, which is separated from the $p150^{Glued}$ binding residues by a flexible linker. Since the light chains are dimeric, their binding affinity for IC is thought to be enhanced by mutual bivalency effects (Stuchell-Brereton et al., 2011; Williams et al., 2007).

In yeast, Nip100 is the dynactin subunit that binds yeast intermediate chain, Pac 11. The proximity of the binding sites for Dyn2 and Nip100 in the Pac11 sequence allows measurement of Pac11 binding to each partner individually and both partners simultaneously.

A combined analysis of high resolution NMR data, binding energetics and gel filtration profiles delineate for the first time residue-level changes in Pac11 structure and dynamics upon binding Nip100 or Dyn2, show clear differences in the energetics of light chain binding to Pac11 versus Drosophila IC, and identify within Pac11 new structural domains formed when either Dyn2 or Nip100 is bound. Linkers are the residues between partner protein recognition motifs. Although they remain disordered in the Pac11 complex with Dyn2 or Nip100, our results show that linkers are not featureless 'random coils' but rather sequences containing distinctive elements of dynamic structure.

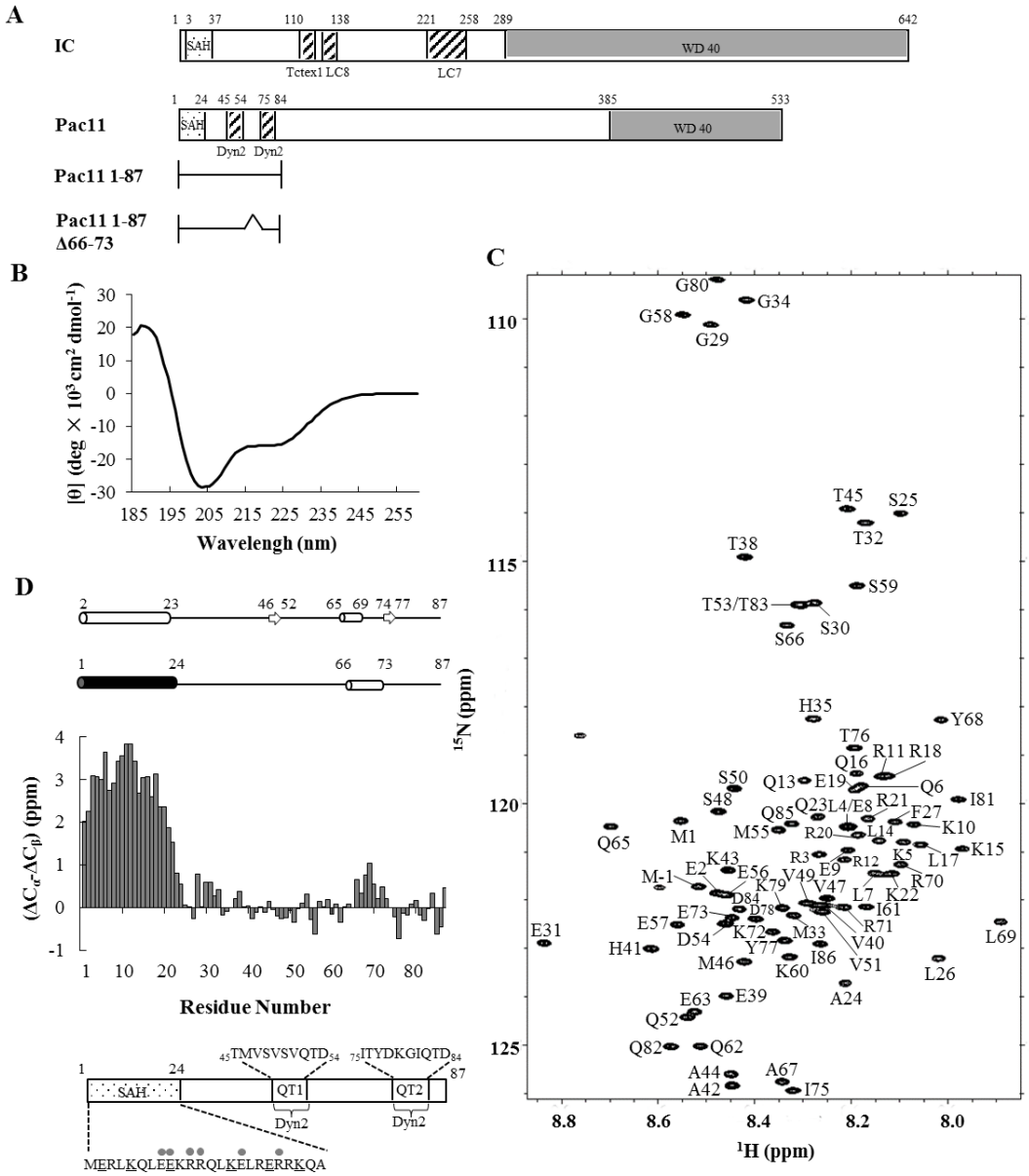
Sequence analysis of apo Pac11 indicates that, contrary to commonly expressed assumptions (McKenney et al., 2011; Siglin et al., 2013; Stuchell-Brereton et al., 2011), the N-terminal residues predicted to be coiled-coil are instead a single alpha helix (SAH). This conclusion is consistent with the strongly helical NMR-detected secondary structure of apo Pac11 residues 1-24. Our hypothesis that dynein intermediate chains, in general, have a SAH sequence at or near the N-terminus, is consistent with the IC sequences of other species.

Results

Construct Design and Characterization. The *Saccharomyces cerevisiae* dynein intermediate chain Pac11 1-87 construct includes the sequence-predicted coiled-coil domain (1-24), as well as two Dyn2 recognition motifs, QT1 (45-54) and QT2 (75-84) (Fig. 3.1A and D). Pac11 1-87Δ66-73 includes the same domains and is used to determine the effect of the 66-73 deletion on Dyn2 and Nip100 binding affinities. The Nip100 214-389, Nip100 CC1B, contains the coiled-coil domain identified in mammalian homolog $p150^{Glued}$ as sufficient to bind dynein intermediate chain (McKenney et al., 2011). Pac11 1-87 is primarily disordered with a small percentage of CD-detected helical structure indicated by negative ellipticity at 222 nm (Fig. 3.1B), and is monomeric with a MALS-determined molecular weight of 12.3 kDa (calculated MW is 10.6 kDa). Nip100 CC1B and Dyn2 are stable dimers as determined by MALS.

Figure 3.1. Sequence and solution analyses of apo Pac11 1-87. (A) Binding site organization of IC (from *Drosophila melanogaster*) and Pac11 (IC ortholog from yeast) showing predicted N-terminal single α helix (SAH), and C-terminal WD repeat (WD40) domains. Pac11 has two Dyn2 binding sites while Drosophila IC has binding sites for three different light chains Tctex1, LC8 and LC7. The N-terminal SAH domain identified in these studies replaces the coiled-coil designation commonly assumed from sequence-based structure prediction. Constructs used in this work: Pac11 1-87 and Pac11 1-87Δ66-73 are also shown. (B) Far UV CD spectrum of Pac11 1-87 recorded at 25 °C; negative ellipticity at 222 nm indicates partial helical structure. (C) [^{15}N, ^{1}H]-TROSY HSQC spectrum of Pac11 1-87 showing backbone assignments. The spectrum was acquired at 15 °C on a 900 MHz Bruker Avance NMR Spectrometer. (D) The secondary structure of apo Pac11 1-87 determined from the C_α and C_β ^{13}C chemical shifts in the bar graph is diagramed above the graph as a stable helix 1-24 (black) and a nascent helix 66-73 (white). The top diagram gives the algorithm-predicted secondary structure of the same sequence. Below the bar graph, a domain diagram of Pac11 1-87 gives the sequence of the two Dyn2 binding sites QT1 and QT2. The sequence of the N-terminal 24 residues fits the definition of a single α helix (SAH) as described in the text. Underlined are all possible K-E interactions and grey filled circles indicate all possible R-E interactions.

Figure 3.1. (*Continued*)



Resonance Assignments and Secondary Structure of Pac11 1-87. The ^{15}N-^{1}H HSQC spectrum of apo Pac11 1-87 (Fig. 3.1C) exhibits a limited range of amide proton chemical shifts (7.90 to 8.85 ppm), indicating a predominantly disordered protein. Backbone assignments were determined for 82 out of 84 non-proline residues. Secondary chemical shifts indicate high helix propensity for residues 1–24 and lower helix propensity for residues 66-73 (Fig. 3.1D). These propensities match the sequence-based algorithmic predictions of helix (Buchan et al., 2013) for apo Pac11, except the NMR-detected 66-73 helix is slightly longer, and considerably weaker, than the predicted 65–69 helix (Fig. 3.1D, top).

Although N-terminal residues of Pac11 1-87 are predicted to form a coiled-coil, they lack the hydrophobic seam found in coiled-coils. Instead, the 1-24 segment is rich in highly charged residues (65.2% R, K, E) (Fig. 3.1D), suggesting helix stabilizing electrostatic interactions between residues i and i+4 (eg. K5-E9, K15-E19) and i and i+3 (eg. E8-R11), characteristic of a single α-helix (SAH) domain (Peckham and Knight, 2009). Hereafter we refer to helix 1-24 as SAH, and nascent helix 66-73 as LH (for linker helix). The linker region (25-44) between SAH and QT1 is L1, and linker region (58-65) between QT1 and LH is L2.

Pac11-Dyn2 and Pac11-Nip100 CC1B Interactions. Pac11 1-87 binds dimeric Dyn2 with a stoichiometry of 1:2 (n value of 2) and apparent K_d value of 0.04 μM (Fig. 3.2A and Table 3.1). As the stoichiometry is computed for binding sites, 1:2 means 2 Pac11 monomers bind 2 Dyn2 dimers. In comparison, Dyn2 binding to a short Pac11 peptide containing only one Dyn2 binding motif (QT2) has a K_d of 0.62 μM and a stoichiometry of 1:1 (Rao et al., 2013). The deletion of LH residues 66-73 in Pac11 does not significantly affect Dyn2 binding affinity (Table 3.1).

Pac11 binding affinity for Nip100 is much lower than for Dyn2. The Pac11-Nip100 CC1B interaction measured by ITC has a stoichiometry of 1:1 (n value of 1, 2 Pac11 monomers bind 1 Nip100 dimer) and an apparent K_d value of 5.6 μM (Fig. 3.2C and Table 3.1). The absence of LH in Pac11 1-87Δ66-73 significantly reduces Nip100 binding affinity (Fig. 3.2C and D).

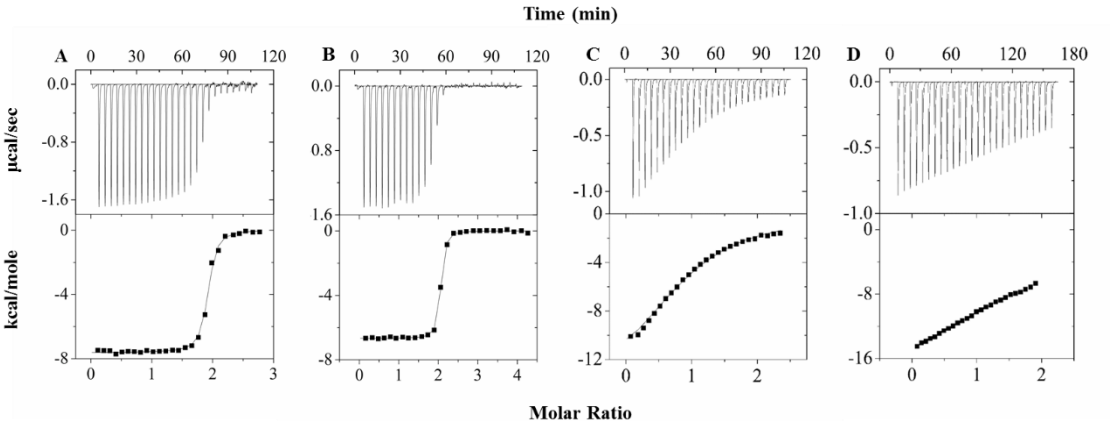


Figure 3.2. Thermodynamic measurements of interactions of Pac11 1-87 and Pac11 1-87Δ66-73 with Dyn2 and with Nip100 CC1B. In each panel representative results of ITC experiments are shown as thermograms (top) and binding isotherms (bottom) for the titrations of (A) Pac11 1-87 with added Dyn2; (B) Pac11 1-87Δ66-73 with added Dyn2; (C) Nip100 CC1B with added Pac11 1-87; and (D) Nip100 CC1B with added Pac11 1-87Δ66-73. Data were collected at 25 °C in 50 mM sodium phosphate, 50 mM sodium chloride, 0.5 mM sodium azide, pH 7.2, and fit to a single-site binding model using Origin 7.0 (Microcal) to compute the parameters listed in Table 3.1. The weak binding of Nip100 CC1B with Pac11 1-87Δ66-73 cannot be reliably fit at these concentrations.

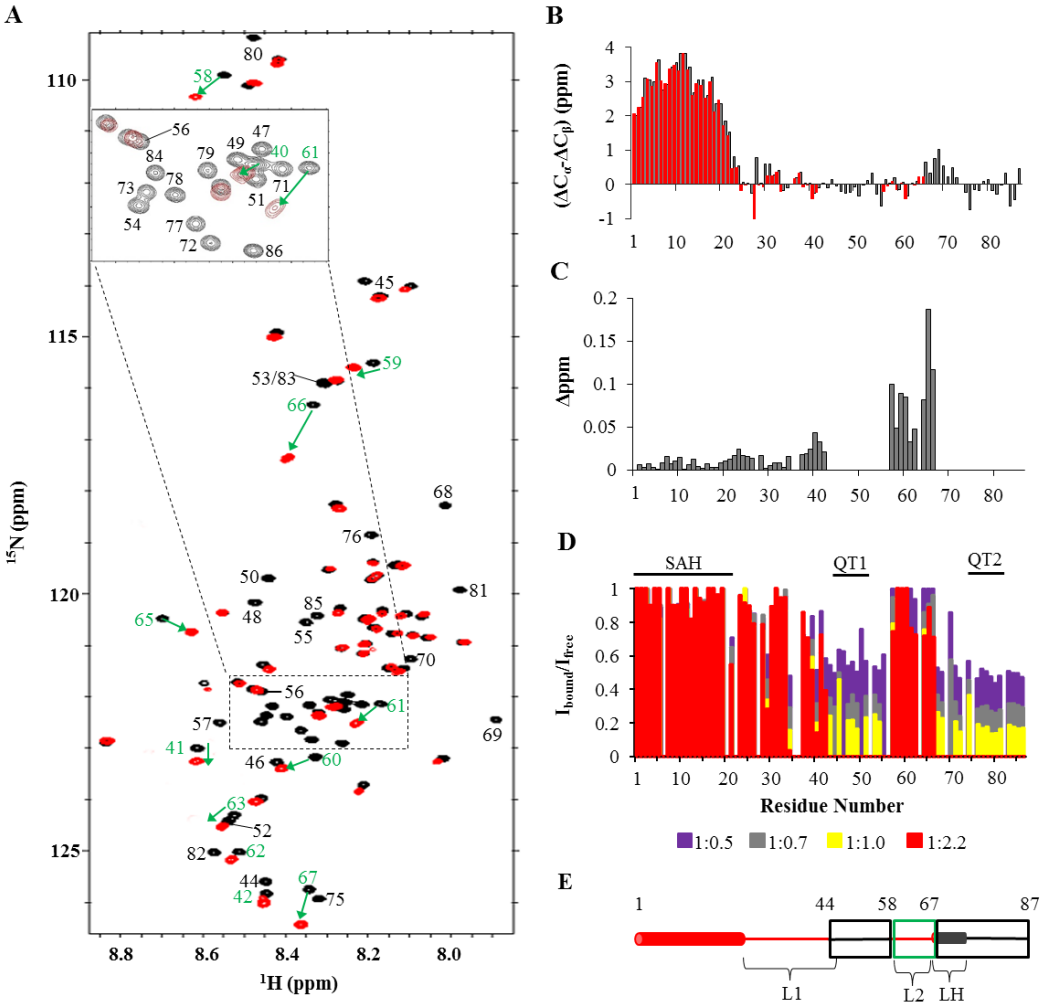
Table 3.1. Thermodynamic parameters for association of Pac11 1-87 and Pac11 1-87Δ66-73 with Dyn2 and Nip100 CC1B.

Protein	Ligand	n	K_d	ΔG^0	ΔH^0	$-T\Delta S$
			μM		*kcal/mol*	
Pac11[a]	Dyn2	2.0±0.1	0.04±0.01	-10.1±0.1	-7.2±0.6	-2.9±0.7
Pac11Δ66-73[a]	Dyn2	1.8±0.2	0.04±0.01	-10.1±0.1	-6.8 ±0.2	-3.3±0.04
Nip100[b]	Pac11	0.97±0.005	5.6±0.07	-7.1±0.02	-8.0±0.5	0.8±0.5

a. Data are the average of duplicate ITC runs. Pac11 denotes Pac11 1-87.
b. Data are the average of duplicate ITC runs. Nip100 denotes Nip100 CC1B.

Figure 3.3. NMR titrations of Pac11 1-87 with Dyn2. (A) Superposition of 900 MHz [^{15}N, ^{1}H]-TROSY HSQC spectra collected at 15 °C of Pac11 1-87 in the apo form (black) and with Dyn2 added at a Pac11:Dyn2 molar ratio of 1:2.2 (red). With Dyn2, some peaks disappear (black labels) and some peaks shift (green labels with green arrows between black and red peaks for the same NH in apo and Dyn2-bound forms). *Inset*, enlargement of the region enclosed by dotted lines. (B) Secondary chemical shifts of observable peaks in spectra of Dyn2-bound Pac11 1-87 (Pac11:Dyn2 molar ratio 1:2.2) (red) compared to those of free Pac11 1-87 (black of Fig. 1D). (C) Bar graph of residue number *versus* chemical shift change (Δppm) of the same NH peak in spectra of Pac11 without and with Dyn2 (ratio 1:2.2). (D) Changes in relative intensity of Pac11 1-87 NH peaks upon addition of Dyn2 in HSQC spectra collected at 700 MHz at 25 °C are shown as a bar graph of residue number *versus* relative peak intensity (I_{bound}/I_{free}). I_{bound} is peak intensity for samples at the indicated Pac11:Dyn2 ratio, except for residues 58-67 for which I_{bound} is the combined intensity of both peaks for the same NH. I_{free} is peak intensity of the same NH in spectra of apo Pac11 with no added Dyn2. Pac11:Dyn2 ratios are 1:0.5 (purple), 1:0.7 (grey), 1:1.0 (yellow) and 1:2.2 (red). Peaks with intensities near the noise level were not included in the final titration (molar ratio 1:2.2). (E) Various sequence regions of Pac11 1-87 are differently affected by Dyn2 binding (ratio 1:2.2). Some regions behave uniformly; their peaks either disappear (black boxes—QT1, QT2 and LH residues), undergo significant chemical shift changes (green box—L2 residues), or retain full intensity with no chemical shift change (red, no box). LH residues (dark grey cylinder) form a helix that is more stable in the Dyn2-bound form.

Figure 3.3. (*Continued*)



NMR Studies of Dyn2 Binding to Pac11. In NMR spectra of ^{15}N labeled Pac11 1-87, taken before and after addition of unlabeled Dyn2 at varying molar ratios, two titration effects are readily apparent for peaks in the amide NH spectral region (Fig. 3.3A) – some NH peaks are attenuated, but their lost intensity is observed in a new peak at a different chemical shift, while other NH peaks lose intensity that is not detected elsewhere in the spectrum. Resonance assignments of all detected peaks were determined using standard 3D experiments. Secondary chemical shifts of observable peaks in the spectrum of Dyn2-bound Pac11 (Fig. 3.3B, red) are the same as for apo-Pac11 (Fig. 3.3B, black), indicating that, for residues that can be detected in Dyn2-bound Pac11, there is no change in secondary structures compared to apo Pac11 (e.g. SAH peaks 1-24).

The first titration effect is illustrated in Fig. 3.3A which overlays HSQC spectra of Pac11 1-87 without Dyn2 (black) and with a Pac11:Dyn2 molar ratio of 1:2.2 (red). With Dyn2, peaks of Pac11 residues 58-67 in linker region L2 are observed at a new chemical shift as shown in Fig. 3.3C, where chemical shift changes are quantified. L2 residues 58-67 show significant shifts, as compared to much smaller shifts of any other residue. Although the 58-67 chemical shifts change with added Dyn2, the combined intensity of the 2 peaks for each NH does not change. A second NH peak at a different chemical shift, with no loss in combined intensity, indicates slow exchange between apo and Dyn2-bound ensembles in which L2 residues experience different local environments, and hence different average chemical shifts. For both apo and Dyn2-bound Pac11, L2 peaks are sharp and within the spectral region considered random coil. We conclude that, in Dyn2-bound Pac11, the L2 domain is an ensemble of conformations that still retain a high degree of local flexibility but are, on average, slightly more ordered than in apo Pac11, as evidenced from the small increase in steady-state NOEs (Fig. 3.6).

The second titration effect, attenuation of peak intensity as Dyn2 is added, is quantified in Fig. 3.3D, where bar heights indicate normalized peak intensity and bar colors code the Pac11:Dyn2 ratio. Peak intensity that is first diminished and then

completely lost as Dyn2 concentration is increased, is observed for residues 44-57 and 68-87, two segments that encompass residues in or immediately flanking binding motifs QT1 and QT2 (45-54 and 75-84) and LH (66-73) (Fig. 3.3D and E).

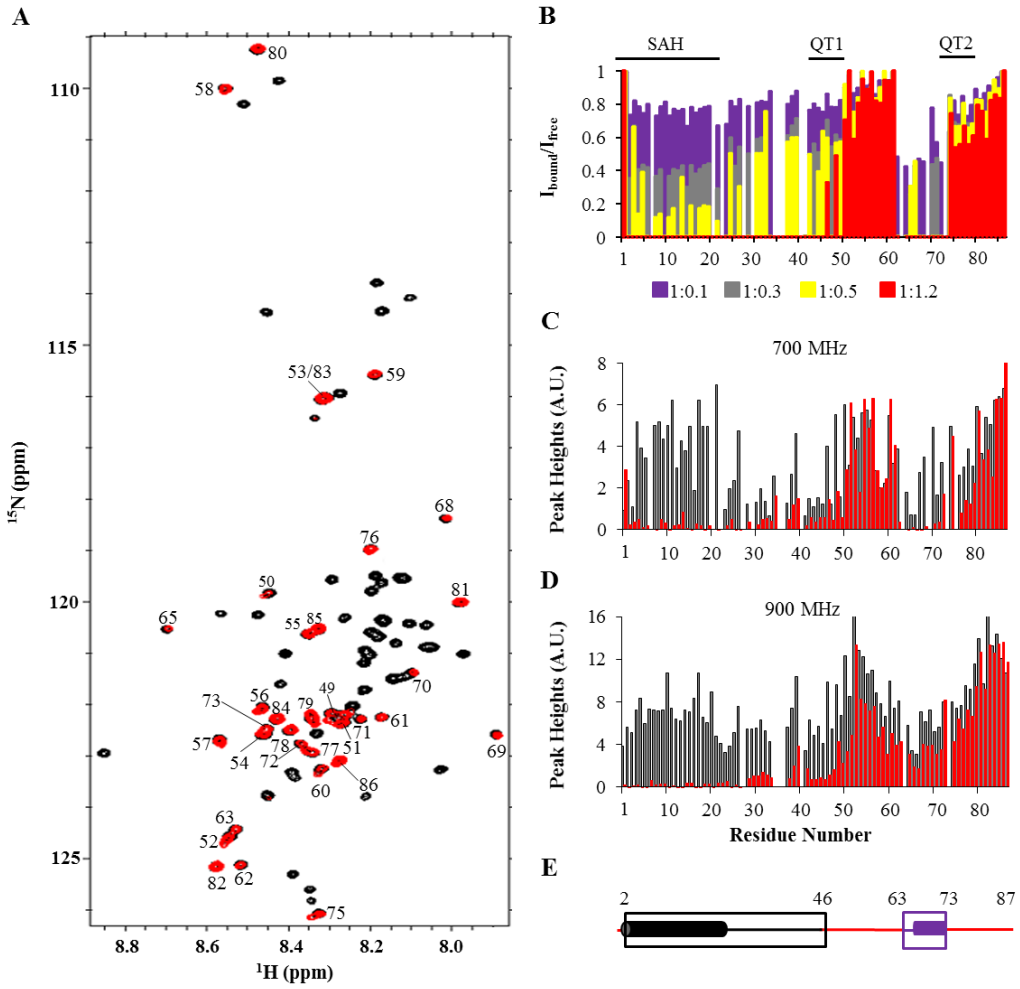
The titration of peak intensity as apo Pac11 is converted to Pac11/Dyn2 complex is explained by one or both of two scenarios. In the first, these residues interact with a large, collapsed domain and so tumble with that domain; the consequent increase in their rotational correlation time, and associated decrease in NMR relaxation time, renders their peaks undetectable. A second possible explanation for total loss of peak intensity at saturating Dyn2 concentration is intermediate exchange between two or more conformational ensembles of Dyn2-bound Pac11 forms. The QT1 and QT2 domains best fit the first explanation as their motifs are incorporated into the three-dimensional fold of the Dyn2 dimer (20) with which they would tumble. Although LH peaks disappear in a manner similar to residues in the QT domains, their tight integration into the nearby QT2 is not consistent with their NMR spectral behavior when Nip100 is added to apo Pac11 (Fig. 3.4) discussed below, nor with the unchanged binding energetics of Dyn2 to Pac11 *versus* Pac11Δ66-73 (Table 3.1).

NMR Studies of Nip100 CC1B Binding to Pac11. NMR spectra of ^{15}N labeled Pac11 1-87, taken before and after addition of unlabeled Nip100 CC1B (Fig. 3.4), identify large segments of contiguous Pac11 residues for which NH peaks are attenuated as Nip100 is added. Comparing the results of titration experiments in Fig. 3.3D and 3.4B, it is striking that the overall patterns of titrated peak intensity show a rough reciprocity; in one graph the two segments of tall red bars (relative peak intensity unchanged at highest ligand concentration) tend to be in regions that in the other graph have no red bars (all intensity lost at highest ligand concentration). Peaks that markedly decrease in intensity with added Nip100 are those of N-terminal residues 2-46 and residues 63-73.

Peak attenuation due to direct interactions (line broadening associated with slower tumbling time) *versus* attenuation due to conformational exchange (line broadening due to intermediate exchange) may be distinguished by comparing Pac11

Figure 3.4. NMR titrations of Pac11 1-87 with Nip100 CC1B. (A) Superposition of 900 MHz [^{15}N, ^{1}H]-TROSY HSQC spectra collected at 15 °C of Pac11 1-87 in the apo form (black) and with added Nip100 CC1B at a Pac11:Nip100 ratio of 1:1.2 (red). (B) Changes in relative intensity of Pac11 1-87 NH peaks upon addition of Nip100 CC1B in HSQC spectra collected at 700 MHz at 25 °C are shown as a bar graph of residue number *versus* relative peak intensity (I_{bound}/I_{free}). Pac11: Nip100 molar ratios are 1:0.1 (purple), 1:0.3 (grey), 1:0.5 (yellow) and 1:1.2 (red). Peak intensities near the noise level are not included for spectra at the highest molar ratio (1:1.2). (C) and (D) Peak heights for residues in free Pac11 1-87 (black) and Nip100-bound Pac11 (red, Pac11 1-87: Nip100 CC1B molar ratio of 1:1.2) in spectra collected at 700 MHz and at 900 MHz. (E) Regions of Pac11 1-87 sequence affected by Nip100 CC1B binding (ratio 1:1.2). Color coding is similar to Fig. 3E. LH residues are shown in purple to indicate higher helical preference (deduced from peak attenuation) than in apo Pac11.

Figure 3.4. (*Continued*)



spectra of free *versus* Nip100-bound samples collected at 700 MHz (Fig. 3.4C) and at 900 MHz (Fig. 3.4D). Broadening due to conformational exchange is sensitive to field strength. At higher field there is an increased $\Delta\omega$, the difference in Hz between two exchange peaks, but no change in k_{ex}, the molecular exchange rate between conformations. At 700 MHz, a conformational exchange rate on the order of $\Delta\omega$ would cause line broadening of an NH peak. At 900 MHz, the same conformational exchange rate may be smaller than $\Delta\omega$, and by definition in the slow exchange regime, with the result that 900 MHz spectra have less broadened peaks for the same NH.

The data in Fig. 3.4B-D, taken together, differentiate three structural ensembles affected by Nip100 binding: SAH (2-24), L1 (25-46) and LH (63-73). In Fig. 3.4B, SAH peak intensities have a distinctive attenuation pattern – a fairly uniform and greater loss of intensity at substoichoimetric Nip100 concentrations (e.g., grey and yellow bars are < 50%). L1 peaks differ in having relatively non-uniform and higher intensity at the same Nip100 concentrations, consistent with SAH being the Pac11 region that interacts directly with Nip100. Further, the intensities of SAH and L1 residues in apo Pac11 spectra have different field-strength dependences (black bars, Fig. 3.4C *vs* D); the intensities of L1 peaks relative to SAH peaks are lower at 700 MHz as compared to 900 MHz. This implies that SAH and L1 residues form different ensembles, each having distinctive conformational dynamics, consistent with the conclusion that in apo Pac11, only residues 2-24 form a single α-helix.

The dynamic ensemble structure of LH is clearly distinguishable from both SAH and L1. The significantly higher intensities of LH peaks observed in 900 MHz spectra of Nip100-bound Pac11 (residues 63-73 red bars, Fig. 3.4C *vs* D) implies slow exchange at the higher field, between LH conformations separated by a significant energy barrier. While peaks 63-73 in 700 MHz spectra are apparently exchange broadened beyond the detection limit, in 900 MHz spectra they are sharpened and more intense due to a shift from an intermediate to a slow exchange regime, with a possible contribution from increased sensitivity.

Effect of Dyn2 on Nip100 Binding to Pac11. Gel filtration elution profiles (Fig.

3.5A) indicate that the presence of Dyn2 increases the stability of Pac11 1-87/Nip100 CC1B complex. When a sample of Pac11 1-87 and Nip100 CC1B at equimolar concentrations (two Pac11 monomers to one Nip100 dimer) was loaded on an analytical column, the two elution peaks (red, Fig. 3.5A) correspond to Nip100 CC1B dimer (42.8 kDa) and to Pac11 1-87 monomer (10.6 kDa), which migrates earlier than expected for a globular protein of similar size consistent with it being disordered. The absence of detectable Pac11/Nip100 complex is consistent with their low binding affinity given in Table 3.1.

When a sample of a complex of Pac11 1-87, Nip100 CC1B and Dyn2, at a molar ratio of 1:1:2 was loaded, the major feature of the elution profile (blue, Fig. 3.5A) is a single intense peak with a MALS-determined molecular weight of 127 kDa, consistent with a Pac11/Dyn2/Nip100 ternary complex, which has an expected molecular weight of 110 kDa (two chains of Pac11 1-87, two Dyn2 dimers and one Nip100 CC1B dimer). The ternary complex is confirmed in lane 'a' of the SDS PAGE results in Fig. 3.5C.

Similar elution profiles are obtained for Pac11 1-87Δ66-73, a Pac11 variant lacking LH residues (Fig. 3.5B). Whether with Nip100 alone, or with both Nip100 and Dyn2, no difference in elution volume of the resultant complexes is observed. The ternary complex with Pac11 1-87Δ66-73 in the high molecular weight peak is confirmed in lane 'b', Fig 3.5C. Apparently LH residues do not contribute to the relative stability of Pac11/Nip100 *versus* Pac11/Nip100/Dyn2.

NMR-Detected Dynamics of Pac11 1-87. Backbone dynamics of Pac11 1-87 were determined from ^{15}N T_1, T_2 and steady-state heteronuclear NOE experiments (Fig. 3.6). The T_1 values are fairly uniform throughout the sequence with an overall average of 0.64 s (Fig. 3.6A). The T_2 values in contrast show highly heterogeneous dynamics. The heterogeneity in T_2 relaxation with residues in the SAH and LH domains evincing low T_2 values (average T_2 value of 0.11 s and 0.21 s, respectively) compared to the overall average value of 0.25 s, is indicative of nonrandom structure for these segments (Fig. 3.6B). Steady-state heteronuclear NOE measurements, which provide indications of ordered structure on the nanosecond to picosecond time scale, suggest order in

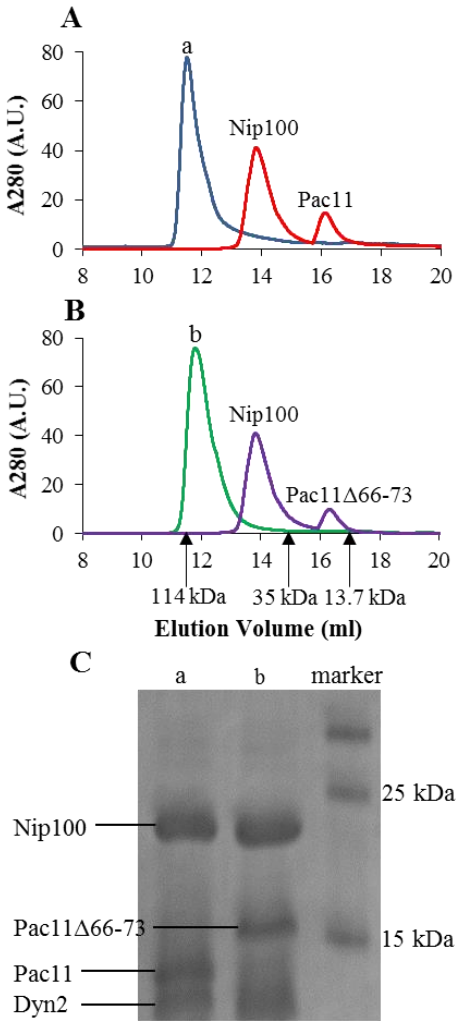


Figure 3.5. Elution profiles of Pac11 1-87 and Pac11 1-87Δ66-73 in ternary complexes with Nip100 and Dyn2. (A) Elution profiles of samples of Nip100 CC1B and Pac11 1-87, with and without Dyn2, are shown with a void volume of 8 ml. Samples containing Nip100 and Pac11 1-87 elute as two peaks (red). With added Dyn2, the complex elutes as a single peak (blue). (B) Elution profiles of samples of Nip100 CC1B and Pac11 1-87Δ66-73, with and without Dyn2, on the same column. Samples containing Nip100 and Pac11 1-87Δ66-73 elute as two peaks (purple). With added Dyn2, the same sample elutes as a single peak (green) for the ternary complex. Arrows indicate elution volumes of molecular weight markers: the 114 kDa complex of IC_{TL7}•Tctex1•LC8•LC7 (chosen because it forms a dynamic IDP complex) (Hall et al., 2010); 35 kDa β-lactoglobulin; 13.7 kDa ribonuclease A. (C) SDS PAGE of sample fractions from elution peaks a and b demonstrate the components of the ternary complexes.

residues 1–24, and to a lesser extent in residues 65–81, segments which have positive NOE values with an average of 0.45 and 0.17, respectively (Fig. 3.6C). The higher order is consistent with the helical propensities observed from secondary chemical shift analysis. The rest of the sequence is primarily disordered with residues exhibiting near zero or negative NOE values.

To examine the effect of Dyn2 binding, similar dynamics experiments were carried out for the Pac11/Dyn2 complex. Because peaks for residues 44-57 and 68-87 disappear in spectra of the bound form, comparison of the dynamics was assessed only for the remaining detectable peaks. Except for a more pronounced dip around residue 31 in T_1, and a significant drop in T_2 for the same region, indicating increase in ordered structure, the T_1 and T_2 values (Fig. 3.6D, E) are similar on average to those of the free protein. NOE values (Fig. 3.6F) indicate a small gain in ordered structure for some regions: for residues around 41, NOE values change from negative to near zero, and for residues 58-67 NOE values increase from 0.08 to 0.16. In sum, dynamics measurements indicate no change for residues within the SAH domain, and a small increase in ordered structure in L1 and L2.

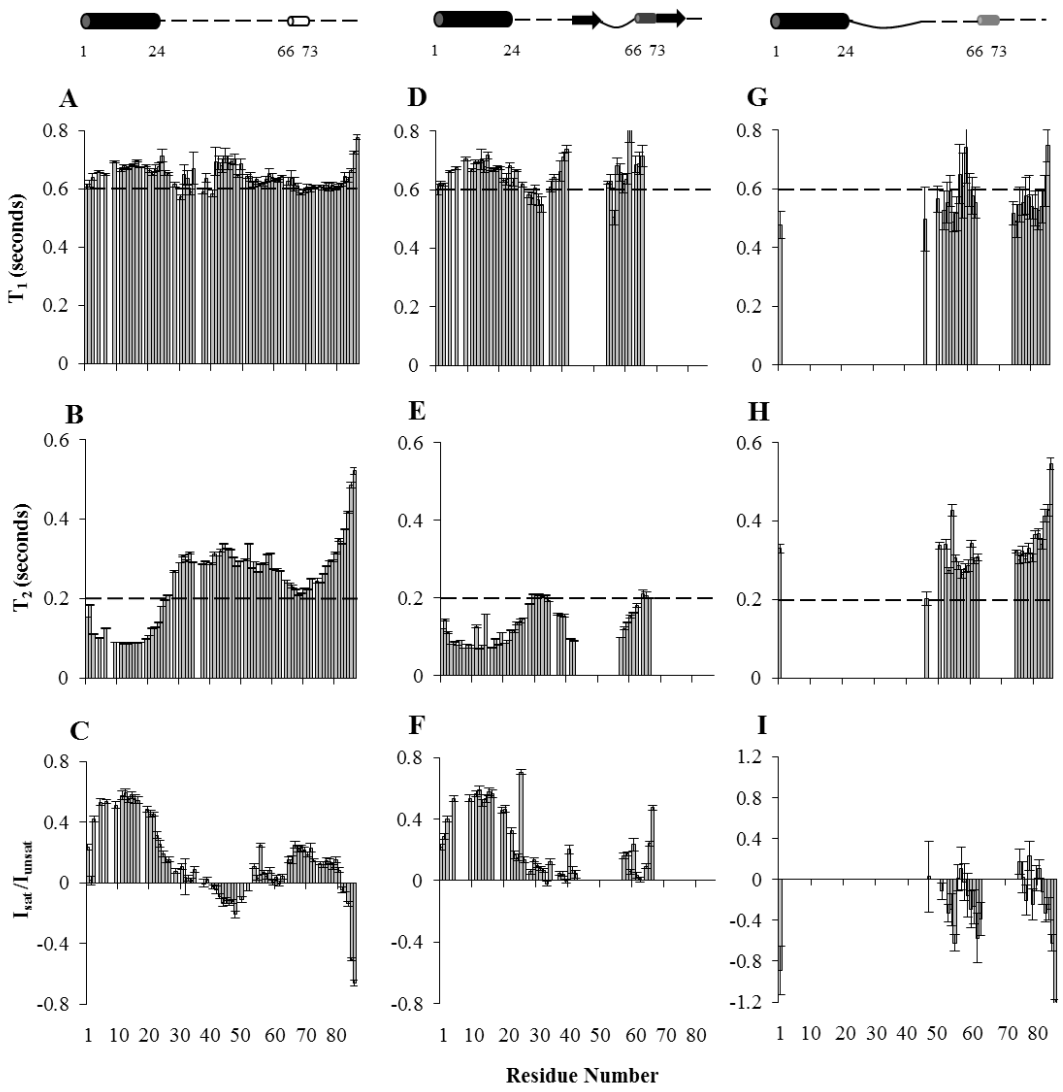
For the Pac11/Nip100 CC1B complex, while many peaks disappear, relaxation parameters for the remaining peaks are essentially the same as for free Pac11. Interesting to note are the more negative NOE values and slightly higher T_2 values for residues in L2 and at the C-terminal end distant from binding, suggesting increased flexibility in this region upon Nip100 binding. Increased flexibility on both sides of LH makes the peak attenuation observed in LH even more striking.

Discussion

High resolution methods for monitoring structural changes of a large, partially disordered, flexible, macromolecular assemblies are few. Our approach – production of appropriate model polypeptides amenable to NMR spectroscopy and titration calorimetry, and further development of NMR ligand titration techniques (Benison et al., 2006; Morgan et al., 2011; Nyarko and Barbar, 2011; Nyarko et al., 2012, 2013) –

Figure 3.6. Dynamics of apo Pac11, Pac11/Dyn2 and Pac11/Nip100. Longitudinal ^{15}N relaxation (T_1), transverse relaxation (T_2) and steady-state heteronuclear NOE values I_{sat}/I_{unsat}) are plotted in (A-C) for apo Pac11 1-87, (D-F) for Dyn2-bound Pac11 1-87, Pac11:Dyn2 molar ratio of 1:2.2, and (G-I) for Nip100-bound Pac11 1-87, Pac11: Nip100 molar ratio of 1:1.2. A dashed line is shown in each plot to facilitate visual comarison. Dynamics parameters of residues for which peaks disappear in bound forms are left blank in the bar graphs. The models (top) depict Pac11 1-87 without and with bound ligand, based on data in Fig. 3 and 4; apo Pac11 1-87 is primarily disordered (dotted lines) with secondary structure the same as Fig. 1D; ligand-bound Pac11 1-87 models show secondary structures when Dyn2 (middle, SAH domain as black cylinder, QT1 and QT2 as arrows, and LH as dark grey cylinder) and Nip100 (right, SAH domain as black cylinder and LH as light grey cylinder) are bound. The dark grey LH in Dyn2-bound Pac11 contains residues that lose intensity, while the light grey LH in Nip100-bound Pac11 contains residues for which intensity loss is field-strength dependent. The effect of Dyn2 binding on Pac11 is the same for data collected at 700 MHz and 900 MHz spectrometers. Solid lines correspond to residues that become more ordered in the complex, while dotted lines correspond to residues that remain disordered.

Figure 3.6. (*Continued*)



has been extremely useful in characterization of dynein intermediate chain complexes, and is generally applicable to characterization of other large dynamic complexes.

The N-terminal 87-residue segment of Pac11, containing two Dyn2 recognition motifs (QT1 and QT2) and the binding site for Nip100, is primarily monomeric and disordered except for a single alpha helix (SAH, residues 1-24) and a short nascent helix (LH, residues 66-73) in the 20-amino-acid segment separating the Dyn2 recognition motifs. Recognition sites and changes in structure upon binding to either Dyn2 or to Nip100 are inferred from NMR titration results, Fig. 3.3-3.4, based on peak attenuation patterns and/or changes in chemical shift with added Dyn2 or Nip100, and on the field strength dependence of peak attenuation. Increased ordered structure in Pac11 complexes is monitored at the residue-level by comparison of ^{15}N relaxation data for apo Pac11 and for Pac11/Dyn2 and Pac11/Nip100 complexes.

An N-terminal SAH Domain Is A Conserved Structural Feature of Dynein Intermediate Chain Responsible for Binding Dynactin. Three models of apo Pac11 1-87 corresponding to the lowest energy structures are presented as an overlay in Fig. 3.7A. The ordered SAH domain aligns well in all three conformations (long red ribbon); in one conformation LH is a short segment exhibiting some helical propensity (short red ribbon); in the rest of the chain, all three conformations have considerable disorder. Dynamics measurements (Fig. 3.6), in particular the significantly lower T_2 values for residues 1-24, support the conclusion that this segment is an ordered SAH domain of the type first identified in murine myosin 10 (Knight et al., 2005) where it serves as a stiff, processive extension of the myosin lever (Baboolal et al., 2009).

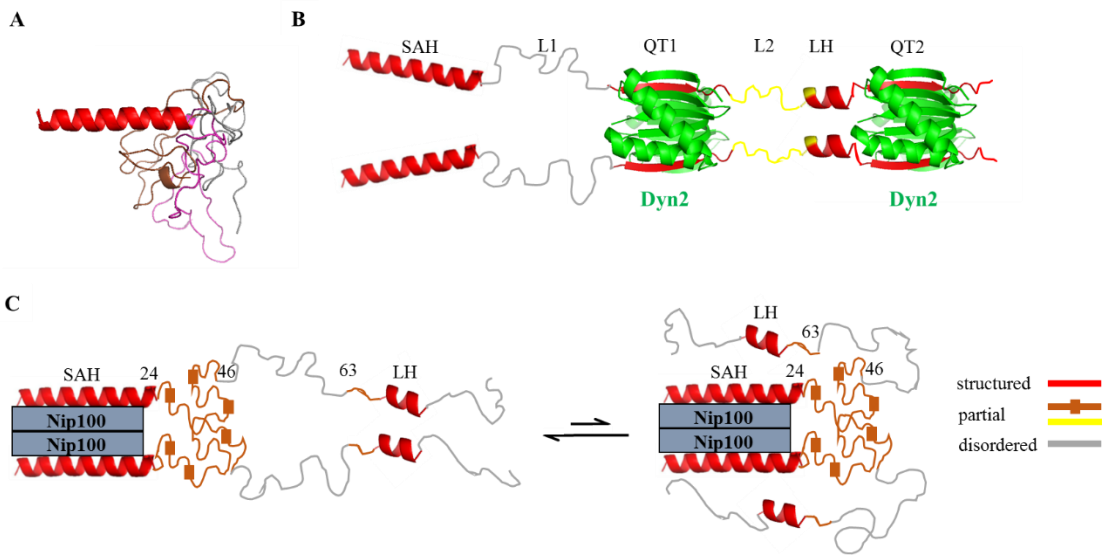
Several characteristics of Pac11 and IC in higher eukaryotes support the generalization that an N-terminal SAH domain is a conserved feature of dynein intermediate chain. The domain architecture is similar for apo forms of *Saccharomyces cerevisiae* Pac11 and Drosophila IC, as illustrated in Fig. 3.1A. At the N-terminus, each has a sequence predicted by standard algorithms to be coiled-coil. However, as noted here for IC and Pac11, and in (Knight et al., 2005) for murine myosin, the predicted coiled-coil sequence lacks coiled-coil-stabilizing hydrophobic residues and has residues

characteristic of a SAH domain conducive to helix-stabilizing electrostatic interactions. Similar to Pac11, in its first 40 residues Drosophila IC favors α-helical structure with restricted motions detected by NMR dynamics experiments (Morgan et al., 2011). Significantly, the yeast Pac11 N-terminal SAH domain binds dynactin (via Nip100) as demonstrated here, and Drosophila and mammalian IC N-terminal SAH domains bind dynactin (via $p150^{Glued}$) (McKenney et al., 2011; Morgan et al., 2011). An elongated SAH may favor binding of this relatively short sequence to the coiled-coil domain of Nip100 or $p150^{Glued}$. In summary, although the N-terminal 40-50 amino-acid segment in IC of higher eukaryotes is commonly assumed to be a coiled-coil based on prediction algorithms, we conclude instead that they form a single α-helix to which a subunit of dynactin binds.

Dyn2 Binding Orders LH Residues As Well As QT1 and QT2 Domains. The aligned IDP duplex formed by two Pac11 1-87 chains and two Dyn2 dimers is illustrated in Fig. 3.7B. Incorporation of Pac11 QT recognition motifs into Dyn2 β-sheets is evident in an X-ray structure for Dyn2 bound to a short peptide corresponding to QT2 (Rao et al., 2013), other LC8-partner peptide structures with similar motifs (Benison et al., 2007; Hall et al., 2009), and the NMR titration results in Fig. 3.3D. A complete loss of peak intensity of residues in and around QT1 and QT2 sequences 44-54 and 75-84 indicate the expected effect of protein-protein interactions at the Pac11-Dyn2 interfaces (Fig. 3.3D). However, along with residues in and near QT1 and QT2, also attenuated are residues 68-73 within LH, which in apo Pac11 forms a nascent helix 66-73 (see Fig. 3.3E). Although LH residues lose all peak intensity in the complex (Fig. 3.3D), they are likely not incorporated into the well-packed QT2 interface of Pac11/Dyn2 because deletion of LH affects neither Dyn2 binding affinity (Table 3.1) nor Dyn2 stabilization of the Pac11/Nip100/Dyn2 ternary complex (Fig. 3.5). An explanation for LH peaks attenuation is that ordering of LH helices accompanies Dyn2 binding. LH residues in the Pac11/Dyn2 complex, shown in Fig. 3.7B as separate red ribbons, may tend to self-associate. Such inter-chain self-association of LH sequences between QT recognition motifs is reminiscent of the Ana2 (the anastral spindle-2 centriole duplication factor)

Figure 3.7. **Structural models of apo Pac11 1-87, Pac11/Dyn2 and Pac11/Nip100.** (A) Apo Pac11 1-87 is shown as an overlay of three CS Rosetta-based models. In each low energy structure, SAH residues 1-24 (red) form a single alpha helix (long red ribbon on left), and, as they are aligned in this view, only 1 ribbon is apparent; for the rest of the sequence (25-87) all three backbone conformations are shown (grey, brown, pink) including one corresponding to LH residues 66-73. (B) Pac11 1-87 in a complex with two Dyn2 dimers. Domains SAH (1-24), L1 (25-43), QT1 (44-57), L2 (58-67), LH (68-73) and QT2 (75-87) are labeled. The Pac11/Dyn2 structures at QT1 and QT2 are from the crystal structure (Protein Data Bank, entry 4HT6), with Dyn2 dimers in green, and bound Pac11 β-sheet strands in red. Dyn2 binding aligns the two Pac11 chains, highly orders most of LH (red ribbon), and partially orders residues L2 (yellow backbone). (C) Pac11 1-87 in a complex with Nip100 CC1B. Nip100 directly binds the SAH domain and orders L1 residues 25-46 (brown with small boxes to indicate local ordered structure) as well as LH residues 63-73 (red ribbon), while the rest of the chain remains disordered (grey). The partially disordered brown region in (C) contains residues that lose intensity, while the partially disordered yellow region in (B) contains residues that undergo chemical shift changes. An alternative model consistent with the NMR peak attenuation data is shown on the right, with the LH domain making tertiary contacts with either the SAH or Nip100. All structures were generated using PyMOL (Delano, 2002).

Figure 3.7. (*Continued*)



system involving two LC8 recognition motifs flanking a 34-amino-acid predicted helix. The self-association of this helix was proposed to be involved in formation of four-helix bundle resulting in higher order oligomerization of Ana2 (Slevin et al., 2014). Similarly for Pac11, it is possible that binding of two Dyn2 at two somewhat separated motifs promotes formation of ordered self-associated helices as a new surface for binding other partners.

Dyn2 binding has no effect on the structure and dynamics of the distant SAH domain, which comprises the Nip100 site. This observation implies that the increased binding of Pac11 to Nip100 upon Dyn2 binding (Fig. 3.5) is due solely to bivalency.

Nip100 Binding to Pac11 SAH Domain Orders Adjacent L1 and Distal LH Residues. Dimeric Nip100 CC1B binds two chains of Pac11 1-87 (Table 3.1) as illustrated in Fig. 3.7C. As Nip100 concentration is increased, SAH peaks in Pac11 spectra diminish and then completely disappear (Fig. 3.4B). Peaks are also attenuated for L1 residues adjacent to SAH, and for LH residues, which are quite distant in sequence from SAH. The rest of the Pac11 backbone (53-62 and 75-87) retains considerable intensity (red bars, Fig. 3.4B) and so remains highly disordered in the complex. Although NH peaks of SAH (2-24), L1 (25-46) and LH (63-73) are all diminished with added Nip100, and therefore more ordered in the Pac11/Nip100 complex than in apo Pac11, their attenuation patterns and field-strength dependence, described in Results, imply that these three regions form local dynamic conformational ensembles having different responses to Nip100 binding.

We conclude that direct interactions of Nip100 coiled-coil domain with Pac11 are localized to residues in the SAH. Attenuation of NH peaks of proximal L1 and of the more distant LH is primarily due to self-organization of these segments into domains that are somewhat more ordered but still flexible. A dynamic L1 domain is visualized in Fig. 3.7C as brown backbone chains that are somewhat collapsed. A dynamic LH domain is visualized as favoring helical conformations, possibly self-associated helices. Residues within these domains could transiently interact with Nip100 or pack against the SAH domain (Fig. 3.7C, right), consistent with diminished binding of Nip100 to a

Pac11 variant having LH residues deleted (Fig. 3.2D).

The salient outcome of these conclusions is that Pac11 contains domains that form elements of local dynamic structure which respond differently to ligand binding and duplex formation. Functional implications include new conformations that may selectively bind additional ligands that are not necessarily bivalent. Such species could arise from an ensemble shift in conformational populations associated with complex formation. For example, dynein assembly of intermediate chains and light chains transiently bind other partners, such as ZW10 subunit of the RZZ complex in higher eukaryotes (Whyte et al., 2008), and Num1 (Farkasovsky and Küntzel, 2001), which might bind newly favored conformations. Further, in the IC sequences of higher eukaryotes the LH domain preceding QT2 is replaced by an extended motif for binding dynein light chain Tctex1, suggesting the possibility of a ligand that binds Tctex1 in higher eukaryotes and Pac11 LH in yeast.

IDP Duplex Scaffolds. This work identifies residue-specific binding effects of dimeric proteins Nip100 and Dyn2 on apo Pac11, a monomeric, IDP. The Pac11/Dyn2 complex is one example of a more general class of IDP systems in which two IDP monomers bind one or more LC8-like proteins to form a duplex scaffold composed of two IDP chains in parallel alignment (Barbar and Nyarko, 2014; Kidane et al., 2013; Nyarko and Barbar, 2011; Nyarko et al., 2013). Along the scaffold are multiple bivalent sites for assembly of various essential bivalent proteins. In the functional assembly, the scaffold does not fold into an arrangement of packed macromolecules; rather the overall complex retains an inherent flexibility in IDP ‘linker’ sequences between bound proteins. The developing insight is that, in these disordered linker sequences the duplex often self-organizes new domains in response to binding partner ligands. The new domains may self-associate and/or offer new binding surfaces for additional ligands, not necessarily bivalent. A ‘new domain’ in this sense is not a folded unit, but rather a dynamic ensemble in which the relative populations of local, interconverting, and widely varying conformations can be shifted when the duplex binds partner ligands. In other words, a new domain is *a sequence element of local dynamic structure* in which

conformational populations may shift when the IDP chain is constrained by binding partner ligands. In apo IDPs, these domains may or may not be experimentally detected by a tendency to form secondary structure. But when the IDP is incorporated into a duplex by binding a cross-linking ligand, the domains are detected as contiguous residues that, for example, are self-associated (one segment from each chain), and/or display increased secondary structure, and/or have NMR peaks that are sensitive to changes in field-strength. New elements of dynamic structure identified for the IDP duplex scaffold of Pac11 are LH domains in response to Pac11 binding of Dyn2 or Nip100, and an L1 domain in response to binding of Nip100.

These results reinforce several generalizations about IDP duplex scaffolds, and underscore the interplay of intrinsic disorder, bivalency and propensity to form local dynamic structure in propagating long-range effects of ligand binding to IDP scaffolds. First, binding contacts of bivalent folded partner proteins are limited to IDP residues in and near the recognition motif. Second, the flexible linkers that retain disorder in the duplex are not featureless peptide sequences connecting binding motifs; rather they contain elements of dynamic structure from which conformations specific for self-association, and/or for binding additional ligands, may be selected. Last, the biological relevance of IDP duplex scaffolds extends beyond dynein; an array of similar IDP scaffolds is expected for assemblies that function in an astonishingly broad range of biological systems (Barbar, 2008; Rapali et al., 2011a), including interactions with the DNA damage response protein ASCIZ, tumor suppressor protein Kibra, mitotic spindle assembly protein Chica, and in rabies virus replication (Dunsch et al., 2012; Rapali et al., 2011b; Raux et al., 2000; Rayala et al., 2006).

Experimental Procedures

Protein Preparation. Pac11 construct containing residues 1-87 was generated by PCR and subsequently cloned into a modified pET15da (Novagen) expression vector with an N-terminus His_6 tag followed by tobacco etch virus (Tev) protease cutting site. Pac11 1-87Δ66-73 was cloned into the same vector by the Gibson Assembly method. A

Nip100 construct containing residues 214-389 (Nip100 CC1B) was generated by PCR and subsequently cloned into a pET15b (Novagen) expression vector with a C-terminus His_6 tag. DNA sequences were verified by automated sequencing. The recombinant vectors were transformed into Rosetta (DE3) cell lines for protein expression. Dyn2 was prepared as described elsewhere (Nyarko et al., 2013) but with modifications that allow the removal of the fused SUMO tag.

Cells were grown in LB media at 37 °C to an optical density (A600) of 0.6-0.8 followed by protein induction for 3-16 h with 0.1-0.4 mM IPTG. For Pac11 1-87 WT and Δ66-73, which both form inclusion bodies under the above conditions, the cells were lysed using denaturing buffer (100 mM sodium phosphate, pH 8, 10 mM Tris, 8 M urea) followed by affinity chromatography purification on Ni-NTA resin (Qiagen) and elution in imidazole denaturing buffer (20 mM sodium phosphate, pH 8, 500 mM sodium chloride, 350 mM imidazole, 1 mM sodium azide, 6 M urea). The proteins were refolded by stepwise dilution in refolding buffer (20 mM sodium phosphate, pH 8, 500 mM sodium chloride, 1 mM sodium azide) to a urea concentration of 1 M followed by dialysis into the same refolding buffer with 0 M urea. The His_6 tag was cleaved using Tev protease followed by purification on a Superdex™ 75 (GE Healthcare) size exclusion chromatography column, resulting in purity > 95%. Dyn2 and Nip100 CC1B were purified under native conditions. The SUMO tag was cleaved from Dyn2 using SUMO protease (Cornell University). Nip100 CC1B was lysed in the affinity buffer containing 2 M urea to prevent aggregation and improve its interactions with Ni-NTA resin. Protein concentrations were determined from absorbances at 280 nm using molar extinction coefficient values (Pac11 1-87: 2980 M^{-1} cm^{-1}; Pac11 1-87Δ66-73: 1490 M^{-1} cm^{-1}; Dyn2: 12950 M^{-1} cm^{-1}; Nip100 CC1B: 8940 M^{-1} cm^{-1}) computed by expasy.org/protparam. The computed Nip100 CC1B molar extinction coefficient takes into account the C-terminal polyhistidine sequence.

Protein complex formation was assessed by analytical size exclusion chromatography (Superdex 200 10/300, GE Healthcare) at 0.2 ml/min in buffer containing 50 mM sodium phosphate, 100 mM sodium chloride, 1 mM sodium azide at

pH 7.2. In each run, 100 μl protein samples were loaded at concentrations of 100 μM for Nip100 CC1B, Pac11 1-87 and Pac11 1-87Δ66-73, and of 200 μM for Dyn2. Protein (complex) markers were selected to facilitate comparison to expected complexes of similar molecular weights: markers show elution volumes for the 114 kDa complex of ICTL7•Tctex1•LC8•LC7 (Hall et al., 2010), 35 kDa for β-lactoglobulin, and 13.7 kDa for ribonuclease A. Computed molecular weights of assigned complexes are derived using the theoretical molecular weights of Pac11 1-87 monomer (10.6 kDa), Dyn2 dimer (20.8 kDa) and Nip100 CC1B dimer (42.8 kDa). SDS PAGE was used to examine protein content in the eluted fractions.

Circular Dichroism Spectroscopy. CD spectra were recorded for Pac11 1-87 on a JASCO 720 spectropolarimeter using a 1-mm cell at a concentration of 20 μM in 10 mM sodium phosphate, pH 6, and at 25 °C, 15 °C, and 5 °C.

Isothermal Titration Calorimetry. ITC experiments of Pac11 with Dyn2 and Nip100 CC1B were performed using Microcal's VP-ITC microcalorimeter (North Hampton, MA) at 25 °C in buffer composed of 50 mM sodium phosphate, pH 7.2, 50 mM sodium chloride, 0.5 mM sodium azide. Experimental conditions were identical for Pac11 1-87 and for Pac11 1-87Δ66-73, in which residues 66-73 are deleted. In Dyn2 experiments, an initial 2 μL injection was followed by 26 injections of 10 μL Dyn2 samples (250-400 μM) into 15-28 μM Pac11 1-87 (Fig. 2A) or into Pac11 1-87Δ66-73 (Fig. 2B) in the sample cell. An initial 2 μL injection was followed by 26 injections of 10 μL Pac11 1-87 (350-500 μM) into Nip100 CC1B (36-50 μM) in the sample cell. All experiments were accompanied by a constant stirring rate of 351 rpm. Protein samples and buffer were degassed prior to data collection. Data were processed using Origin 7.0 (Microcal) and fit to a single-site binding model. Stoichiometry was computed for number of binding sites, for example, on 1 Pac11 monomer there are 2 Dyn2 sites, giving a stoichiometry of 1:2 for Pac11/Dyn2; on 1 Pac11 monomer there is 1 Nip100 site, giving a stoichiometry of 1:1 for Pac11/Nip100. Recorded data are the average of three independent experiments, with uncertainty reported as the difference between the experimental value and the average.

NMR Experiments. Doubly labeled $^{15}N/^{13}C$ and ^{15}N labeled proteins were prepared in buffer composed of 10 mM sodium phosphate at pH 6.0 with 50 mM sodium chloride, 10 mM Arg^+, 10 mM Glu^-, 1 mM sodium azide, 8% ^2H_2O, a protease inhibitor mixture (Roche Applied Science), and 2,2-dimethylsilapentane-5-sulfonic acid for ^1H chemical shifts referencing. For backbone assignments, NMR spectra of Pac11 1-87 were obtained at 15 °C and concentration of 300 μM. Dynamics data of ^{15}N labeled Pac11 1-87 were collected for the free protein, for Pac11-Dyn2 complex at molar ratio (Pac11: Dyn2) of 1:2.2, and for Pac11-Nip100 CC1B complex at molar ratio (Pac11: Nip100 CC1B) of 1:1.2. Native PAGE and ^1H-^{15}N HSQC spectra were used to assay sample integrity before and after every NMR data collection. All reported data are from HSQC spectra that remain unchanged during data collection.

Backbone resonance assignments for apo Pac11 1-87 were determined from a set of [^{15}N,^1H]-TROSY (Salzmann et al., 1998) based triple-resonance HNCACB and HN(CA)CO experiments featuring band selective-excitation-short-transient (BEST) (Favier and Brutscher, 2011) methodology to reduce measurement times. A TROSY-type HAHBHN(CACO)NH (Löhr et al., 1999) experiment was collected to determine proton assignments. These experiments were carried out with a cryo-probe on a Bruker Avance 900 MHz spectrometer, while the remaining experiments (HN(CO)CACB, HNCO) employed a room-temperature xyz-gradient probe on a Bruker Avance 500 MHz spectrometer. Backbone resonance assignments for Dyn2-bound Pac11 1-87 (Pac11:Dyn2 molar ratio of 1:2.2) were determined using a similar HN(CO)CACB experiment collected on a Bruker Avance 600 MHz spectrometer.

Longitudinal (T_1) and transverse (T_2) ^{15}N relaxation times were determined at 600 MHz ^1H frequency using standard HSQC-based pulse sequences (Farrow et al., 1994) at 15 °C. For each series, spectra with ten different relaxation periods (three in duplicate for error estimation) were recorded in an interleaved manner. Their duration varied from 10 ms to 1.3 s in T_1 measurements, from 23.2 ms to 371.2 ms in T_2 measurements of the free Pac11 1-87, and from 20 ms to 200 ms in T_2 measurements of the complex with Dyn2 or Nip100. Steady-state ^{15}N-{^1H} NOE values resulted from

application of TROSY-based pulse sequences (Lakomek et al., 2012) using a total recovery delay of 8 s. A train of 180° ^{1}H pulses spaced by 22 ms was applied to protons for 6 s in the saturation experiment.

NMR titrations of ^{15}N-labeled Pac11 with unlabeled Dyn2 and Nip100 CC1B were acquired on a Bruker Avance 700 MHz spectrometer at 25 °C. A series of BEST HSQC spectra of ^{15}N-labeled Pac11 1–87 were collected with Dyn2 at final molar ratios (Pac11: Dyn2) of 1:0.1, 1:0.3, 1:0.5, 1:0.7, 1:1.3, and 1:2.2 and with Nip100 CC1B at final molar ratios (Pac11: Nip100 CC1B) of 1:0.1, 1:0.3, 1:0.5, 1:0.7, and 1:1.2. Titration experiments with Nip100 CC1B were performed at pH 7.0. The spectra of fully bound Pac11/Dyn2 (molar ratio of Pac11:Dyn2=1:2.2) and Pac11/Nip100 (molar ratio of Pac11:Nip100=1:1.2) were also collected on a Bruker Avance 900 MHz spectrometer.

NMR Data Analysis. All spectra were processed with TopSpin (Bruker) and analyzed using Sparky (Goddard and Kneller) and NMRView (Johnson, 2004). C_{α} and C_{β} chemical shifts were used to determine the secondary structure by calculating the difference from the random coil values at specified temperature and pH (Kjaergaard et al., 2011). For Pac11-Dyn2 interactions, per-residue absolute chemical shift change was calculated using both ^{15}N and ^{1}H chemical shift changes in spectra of free Pac11 and Dyn2-bound Pac11 collected at 900 MHz. A scaling factor of 0.17 was used on ^{15}N chemical shift to eliminate the differences in the ^{15}N and ^{1}H chemical shift ranges. For titration and dynamics experiments, peak intensities were measured as peak heights. To account for differences in concentration across the titration series, a normalization factor was determined from the signal to noise ratio of residue 1 in Pac11-Dyn2 titration or residue 87 in Pac11-Nip100 titration. Changes in peak intensities I_{Bound}/I_{Free} were calculated as the ratio of signal intensity in spectra of the complex and free Pac11. T_1, T_2 and steady-state NOE values were determined as described previously (Benison et al., 2006; Morgan et al., 2011).

Chemical Shifts-Based Structures. An ensemble of structures for apo Pac11 1-87 was generated using the program CS-Rosetta (Shen et al., 2008) and chemical shift assignments for the $^{1}H_{\alpha}$, $^{1}H_{N}$, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{13}C'$, and ^{15}N atoms. The CS-Rosetta

calculations were carried out on the server at the BMRB using the chemical shifts data. The average RMSD for the 10 structural models that have the lowest energy structures is 16.2 Å. This large deviation is due to the expected differences in the disordered regions. All 10 structures have similar N-terminal α-helix spanning residues 1-24 with relatively lower average RMSD of 2.3 Å.

Sequence Analysis. Domain boundaries of Pac11 were determined using NCBI BLAST. Coiled-coil domains were predicted by the Coils program (Lupas et al., 1991) and secondary structure was predicted by PSIPRED v3.3 (Buchan et al., 2013).

Acknowledgements

This work is supported by National Institutes of Health Grant GM 084276 to EB. We acknowledge the support of the protein core facility in the OSU Environmental Health Sciences Center (NIH/NIEHS 00210), and access to the Research Infrastructure activity in the 7th Framework Programme of the EC (Project number: 261863, Bio-NMR) (Frankfurt, Germany). The authors also wish to thank Prof. Peter Knight for suggesting the presence of SAH domain in IC, Prof. Clare Woodward and the Barbar lab for valuable discussions.

Chapter 4

Regulation of Dynein IC Involves a Novel Interplay between Phosphorylation and Disordered Spliced Linkers

Jing Jie, Frank Löhr and Elisar Barbar

Abstract

Dynactin and NudE are the most prevalent regulators of cytoplasmic dynein motility and cargo binding activities. Both interact with the intrinsically disordered N-terminal domain of dynein intermediate chain IC, which also contains the phosphorylation sites that regulate these interactions. NMR and ITC studies show that the Ser84 phosphorylation site is in a disordered linker distant from the dynactin and NudE binding interface. Structural studies of a phosphomimetic Ser to Asp mutation suggest that phosphorylation stabilizes an electrostatic cluster that docks the disordered linker containing Ser84 against the N-terminal helix, resulting in a conformation that blocks access of IC to dynactin but not to NudE. Formation of this cluster is tightly regulated by the length and sequence of the disordered linkers, and the position of the phosphorylation site. This model explains the selective binding of IC of dynactin *versus* NudE and why this selection is specific for the IC-2C isoform.

Introduction

The transport activity of the 1.6-MDa dynein motor complex is essential for a variety of cellular processes including chromosome segregation during mitosis, axonogenesis, and movement of vesicles (Kardon & Vale, 2009; Vallee *et al*, 2004, 2012). The dynein domain responsible for cargo attachment and for interactions with regulatory proteins consists of multiple subunits, a key component of which is dynein intermediate chain IC, whose intrinsically disordered N-terminal 300-amino acid segment, N-IC (Makokha et al., 2002), is central to dynein assembly and function. N-IC interacts with the light chain subunits Tctex, LC8 and LC7 (Hall et al., 2010; Lo et al., 2001; Makokha et al., 2002; Mok et al., 2001; Susalka et al., 2002) as well as with the non-dynein proteins dynactin, the nuclear distribution protein NudE, and Huntingtin (Caviston et al., 2007; Karki and Holzbaur, 1995; McKenney et al., 2011; Morgan et al., 2011; Nyarko et al., 2012). N-IC embodies the unique properties of intrinsically disordered proteins (IDPs): lacking a 3D structure, they form conformational ensembles that are, on average, aperiodic and somewhat extended but not well packed, and often

present multiple ligand-binding sites in a linear scaffold that promotes protein-protein interactions (Clark et al., 2015). Many IDPs, like IC, remain partially disordered in the assembled multiplex, a feature that provides flexibility and versatility in the regulation of critical functions of the large and multifaceted assembly (Benison *et al*, 2006; Tompa & Fuxreiter, 2008; Clark *et al*, 2015). Partial disorder in an IDP assembly also provides accessible sites for post-translational modifications such as phosphorylation, and alternative splicing (Dunker et al., 2015). Other types of IDPs undergo disorder to order transitions when interacting with partners, *aka* folding-coupled binding (Wright and Dyson, 1999, 2009). Here we investigate the interplay of intrinsic disorder, phosphorylation, and tissue-specific alternative splicing of IC in dynein regulation by dynactin and NudE.

Dynactin binds to microtubules and to dynein through its largest subunit, $p150^{Glued}$ and this interaction is essential for recruitment of dynein to cargo (Karki and Holzbaur, 1999), for dynein's processivity along microtubules (Kardon et al., 2009; King and Schroer, 2000), and for correct spindle formation and cell division (Godin et al., 2010). Another essential dynein interaction is with NudE and its homolog Nudel which regulates dynein recruitment to kinetochores and membranes (Lam et al., 2010), centrosome migration (Wainman et al., 2009), and mitotic spindle orientation (Lu and Prehoda, 2013). Since both dynactin $p150^{Glued}$ and NudE bind dynein IC, at a common site (Barbar, 2012; McKenney et al., 2011; Morgan et al., 2011; Nyarko et al., 2012), the selective binding of one protein over another has to be tightly regulated especially when both are present in the same cellular compartment. One potential mechanism for selection between these partners is phosphorylation at a serine-rich region distant from the site of binding (Nyarko et al., 2012), a modification known to affect several dynein-related cellular processes (Blasier et al., 2014; Huang et al., 1999; Mitchell et al., 2012; Pullikuth et al., 2013; Runnegar et al., 1999).

In mammals, ICs are encoded by two genes *Dyncli1* and *Dyncli2* expressing two proteins IC-1 and IC-2 which share 69% sequence identity and each forms multiple isoforms (Dillman and Pfister, 1994; Kuta et al., 2010; Vaughan and Vallee, 1995a). In

rat, the expression of alternatively spliced A-C isoforms is tissue- and development specific: IC-2C is the only ubiquitously expressed isoform (Ha et al., 2008) whereas IC-1 isoforms are primarily in brain, spinal cords and embryonic tissues (Kuta et al., 2010). Vaughan *et al.* first showed that both serine 84 phosphorylation of IC-2C and the phosphomimetic mutation S84D abolish $p150^{Glued}$ binding in COS-7 cells and in 2D gel assays (Vaughan, 2001). As similar experiments were not performed on IC-1 isoforms, it is not known if the effect of phosphorylation is isoform specific. Additional studies of phosphorylation of IC-2C affecting binding to proteins besides $p150^{Glued}$ underscore the essential role of phosphorylation *in vivo*. For example, phosphorylation at residue T89 enhances binding to ZW10 resulting in targeting dynein to the kinetochore (Whyte et al., 2008), while phosphorylation at residues S88/T89 reduces dynein interaction with NudE (Gao et al., 2015) and stimulates dynein activation. Two recent reports of dynactin-dynein complex structure give interesting features of the structure of dynactin with dynein (Chowdhury et al., 2015; Urnavicius et al., 2015) but do not offer atomic resolution of the interaction of $p150^{Glued}$ with IC, nor of the Ser-rich region as these parts of the structure are highly flexible and therefore not evident in negative stain or cryo-electron microscopy.

Here we address two longstanding questions in the field of dynein regulation, using the Ser-to-Asp point mutations reported in the dynein literature to mimic *in vivo* phosphorylation of IC (Vaughan, 2001). One is whether phosphorylation underlies the selection between dynactin and NudE, and the other is whether this phosphorylation effect is modulated by alternative splicing. We offer answers to these questions and identify a subtle yet significant structural effect of phosphorylation in a disordered region distant from the interaction/functional site. The negative charge at Ser84 when it is phosphorylated or when replaced by Asp promotes electrostatically-driven long-range tertiary contacts whose effects are modulated by residues in disordered linkers distant from the phosphorylation site. Elucidation of these long-range allosteric interactions offers new insight into modulation of dynein functionalities and of dynamic IDP complexes in general.

Results

Constructs Design. Mammalian IC exists in several tissue-specific isoforms, of which the two studied here, rat 1A and 2C, are most different in sequence and function (Kuta et al., 2010). Sequence domains of isoforms IC-1A and IC-2C are shown in Fig. 1A along with equivalent domains of IC from *Drosophila melanogaster* (Dros IC) and from *Saccharomyces cerevisiae* (Pac11). Our constructs of N-terminal sequences of mammalian isoforms, IC-$1A_{112}$ and IC-$2C_{96}$ (Fig. 4.1B, C), were designed to include the single alpha helix (SAH), helix 2 (H2), the Ser-rich region, and the binding domain to both $p150^{Glued}$ and NudE as identified from alignment with Dros IC and Pac11 (Jie et al., 2015; Morgan et al., 2011; Nyarko et al., 2012). Mutations that mimic phosphorylation, chosen because they are used in *in vivo* experiments, are indicated by stars and are single amino acid replacements of Ser to Asp in each of two isoforms, namely, sites Ser100 in isoform 1A and Ser84 in isoform 2C (both shown in green, Fig. 4.2C).

Additional smaller IC-$2C_{44}$ and IC-$2C_{83}$ constructs (Fig. 4.1C) were designed to identify the domain boundaries required for binding to $p150^{Glued}$ and Nudel. The human $p150^{Glued}$ construct employed in this study, p150*, corresponds to residues 382-531 predicted to form a dimeric coiled coil, and is the smallest $p150^{Glued}$ fragment sufficient to bind IC (Vallee et al., 2012). The human Nudel 1-189 contains the predicted dimeric coiled-coil domain at its N-terminus, referred to as nNudel, and is designed based on the comparison to the domain of NudE that binds Dros IC (Nyarko et al., 2012).

Structural Studies of IC-$2C_{96}$. The ^{1}H-^{15}N HSQC spectrum of IC-$2C_{96}$ (Fig. 4.2A) shows backbone assignments for all 90 non-proline residues. Secondary chemical shifts (Fig. 4.2B, bar graph) identify high helical propensity for residues 4–38 (SAH) and residues 52-66 (H2), consistent with the CD spectrum (Fig. 4.2D), and indicated by ribbons (Fig. 4.2E). The rest of the protein, residues 39-51 linking SAH and H2 (L1) and residues 67-80 linking H2 and the Ser-rich domain (L2), is disordered. These propensities match sequence-predicted IC-$1A_{112}$ and IC-$2C_{96}$ structures, except that the first helix is 10 residues shorter than predicted (Fig. 4.2B). The SAH domain contains

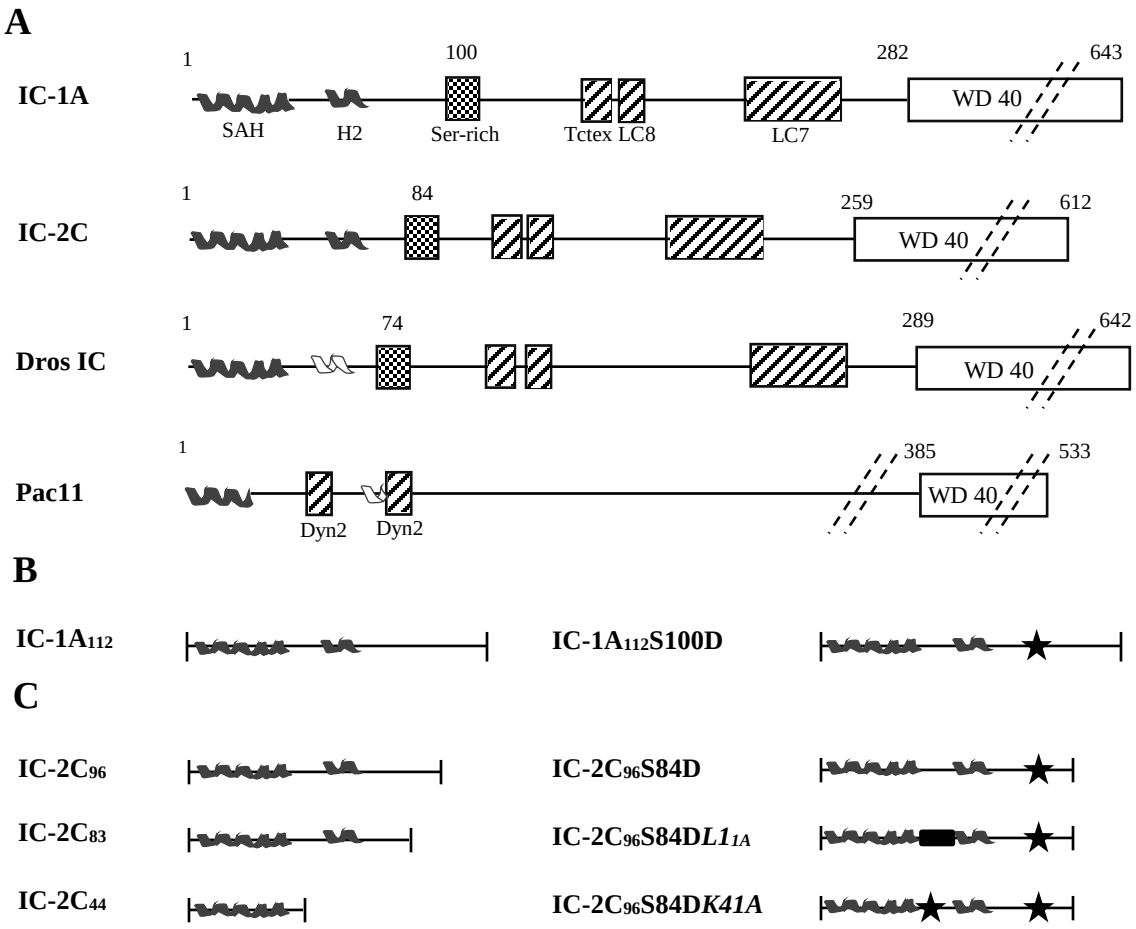
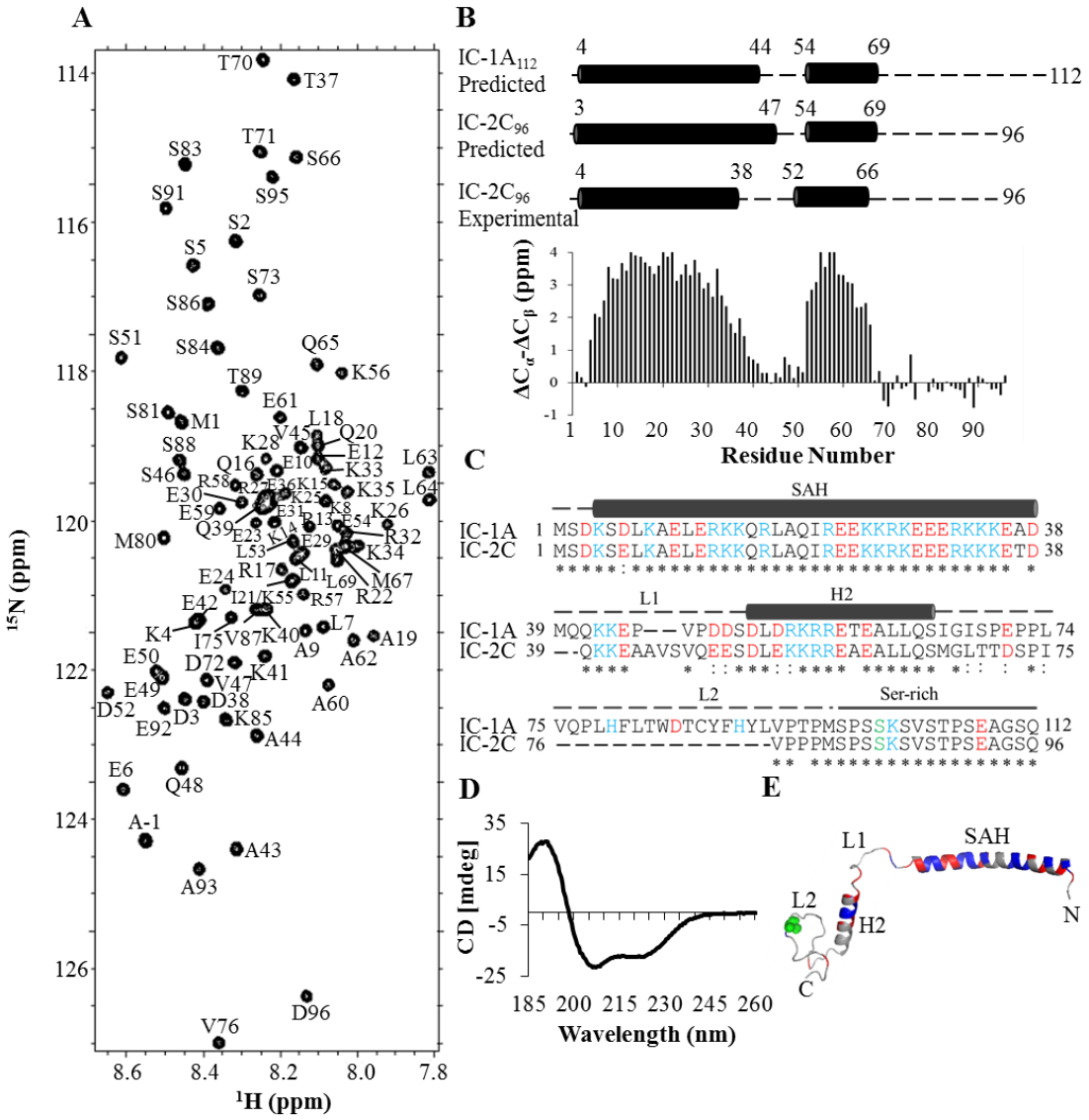


Figure 4.1. Dynein intermediate chain from two higher eukaryotes and yeast. (A) Sequence domains in IC from rat (isoforms IC-1A and IC-2C), Drosophila (Dros IC) and yeast (IC homolog Pac11) show at the N-terminus a single α helix (SAH) (Jie et al., 2015). NMR data (Jie et al., 2015; Morgan et al., 2011) also indicate in each IC the presence of a second helix (H2), which may be either nascent (white helix) or strong (black helix). The C-terminal sequence is predicted to be a WD repeat (WD40) domain (empty box). Mammalian IC and Drosophila IC have binding sites for three light chains Tctex, LC8 and LC7 (diagonal stripes), and contain a Ser-rich region (cross-hatched) preceding the light chain binding sites. In yeast, Pac11 has binding sites for two copies of Dyn2 (LC8 homolog) and no Ser-rich region. (B) and (C) show IC-1A and IC-2C constructs used in this study. Black stars indicate the site-specific mutations and the bar points to the region where linker 1 is swapped between the isoforms.

Figure 4.2. IC-2C$_{96}$ structure. (A) An [^{15}N, ^{1}H]-TROSY HSQC spectrum of IC-2C$_{96}$ is shown with backbone NH assignments. Spectra were acquired at 15 °C on a 900 MHz Bruker Avance NMR Spectrometer. (B) The algorithm-predicted secondary structures of IC-1A$_{112}$ and IC-2C$_{96}$, and the experimental secondary structure of IC-2C$_{96}$ show two stable helices 4-38 and 52-66 (black cylinders). Secondary structure prediction was obtained from PSIPRED v3.3 (Buchan et al., 2013; Jones, 1999), and experimental secondary structure was determined from the changes in C_α and C_β chemical shifts relative to random coil. (C) Sequence alignment of IC-1A$_{112}$ and IC-2C$_{96}$ showing identical (asterisk) and similar (colon) residues (D/E, K/R, S/T, and L/I). The diagram above shows black bars for stable helices (SAH and H2), dashed lines for disordered regions that contain the linker sequences L1 and L2, and solid line for the Ser-rich region. Positively and negatively charged residues are shown in blue and red, respectively, while Ser84 phosphorylation site mutated to aspartic acid is shown in green. (D) A CD spectrum of IC-2C$_{96}$ showing a combination of secondary helical and disordered structures. (E) A structural model of IC-2C$_{96}$ based on experimental secondary chemical sifts and highlighting the abundance of charged residues (red: negatively charged, blue: positively charged, green: phosphorylation site).

Figure 4.2. (*Continued*)



~71% charged residues (Fig. 4.2C) and its sequence lacks the hydrophobic residue placement to form a coiled coil.

Our NMR studies here focus on IC-2C due to aggregation of IC-1A at concentrations required for NMR. The 79% sequence identity between IC-1A and IC-2C constructs, and their similar secondary structure predictions (Fig. 4.2B, C) support the expectation that both isoforms have similar structure in the SAH and H2 helices. Their primary difference is the length of L2 and the residue composition of L1 (Fig. 4.2C).

Evidence that helix 4-38 is an independently folded SAH domain stems from the preservation of helical structure in the truncated IC-2C_{44} construct (Appendix 1, Fig. A1.1) and the p150* binding data in Fig. 4.3 (discussed below). While in IC-2C_{44} the secondary chemical shifts show considerable helical structure, they are slightly weaker than in the 1-96 construct (Appendix 1, Fig. A1.1A), suggesting some long-range interactions may contribute to SAH stability in IC-2C_{96}. Comparison of amide chemical shifts in HSQC spectra show that the largest chemical shift perturbations are for residues 40-44 at the C-terminal of IC-2C_{44}, an expected end effect. However, the same spectra also show unexpected changes at the N-terminal 1-5 residues (Appendix 1, Fig. A1.1B), suggesting interactions between residues 1-5 and a segment within residues 45-96 in IC-2C_{96}.

In summary, NMR secondary chemical shifts show that IC-2C_{96} has an independently folded SAH domain at the N-terminus, another shorter helix H2 separated from SAH by a flexible linker L1, and a fully disordered rest of the protein as modeled in Fig. 4.2D. Small chemical shift differences between IC-2C_{96} and IC-2C_{44} suggest a population in which H2 residues pack against SAH.

Interactions of Mammalian IC Isoforms with p150 and nNudel.* The interactions of IC isoforms for p150* and nNudel determined by ITC (Fig. 4.3, 4.4) show a 1:1 stoichiometry consistent with binding of one p150* or nNudel dimer to two chains of monomeric IC. With p150*, both IC-1A_{112} and IC-2C_{96} bind with apparent K_d values of 6.8 μM and 12.6 μM, respectively; associated enthalpic and entropic

contributions are given in Table 4.1. IC-2C_{44} binds p150* with a slightly lower affinity of 17 μM (Fig. 4.3 and Table 4.1) suggesting that while the SAH residues are sufficient for binding, there are other contributors to binding affinity. nNudel binds both IC-2C_{44} and IC-2C_{96} with apparent K_d values of 1 μM (Fig. 4.4A, C and Table 4.1). Taken together, the data indicate that nNudel binds with significantly higher affinity than p150*, and that the first 44 residues are sufficient for binding either protein, but only with p150* does the rest of the protein contribute to modulating the affinity. Supporting this conclusion is the reduced binding observed for IC-2C_{83} with p150* (Fig. 4.3H) but not with nNudel (data not shown).

NMR experiments identify the sites on IC that are perturbed by p150* and by nNudel binding (Fig. 4.5A, B, and 4.6A, B). Data collected at 900 MHz for ^{15}N labeled IC-2C_{96} with unlabeled p150* are shown as HSQC spectral overlay of p150*-bound IC-2C_{96} (red) and unbound IC-2C_{96} (black) (Fig. 4.5B). P150* binding results in complete peak attenuation for residues 1-46 and 51-65 and significant peak attenuation for residues 47-50 and 66-69 (Fig. 4.5H). Similar experiments collected at 700 MHz show complete peak attenuation for residues 2-40, but only small peak attenuation for the rest of these regions (bar graphs overlay of 700 and 900 MHz intensities, Fig. 4.6B). Field-strength-dependence of peak attenuation in L1 and H2 is interpreted as due to conformational exchange in the fast intermediate (sharp peaks, 700 MHz) to slow intermediate (broad peaks, 900 MHz) timescale. With nNudel, HSQC spectral overlay of bound IC-2C_{96} (cyan) and unbound IC-2C_{96} (black) (Fig. 4.5A) show similar peak disappearance as with p150* except that with nNudel smaller differences in field-dependent peak attenuation is observed for residues in L1 and H2 (Fig. 4.6A, B). Thus, while nNudel binds more tightly than p150*, fewer residues in IC-2C_{96} are significantly perturbed.

In summary, both IC isoforms bind p150* with similar moderate affinity, but bind nNudel with considerably higher affinity (order of magnitude difference in K_d). The SAH domain is the primary site for binding both proteins, but residues within L1 and H2 are significantly more affected by p150* binding than by nNudel binding. The

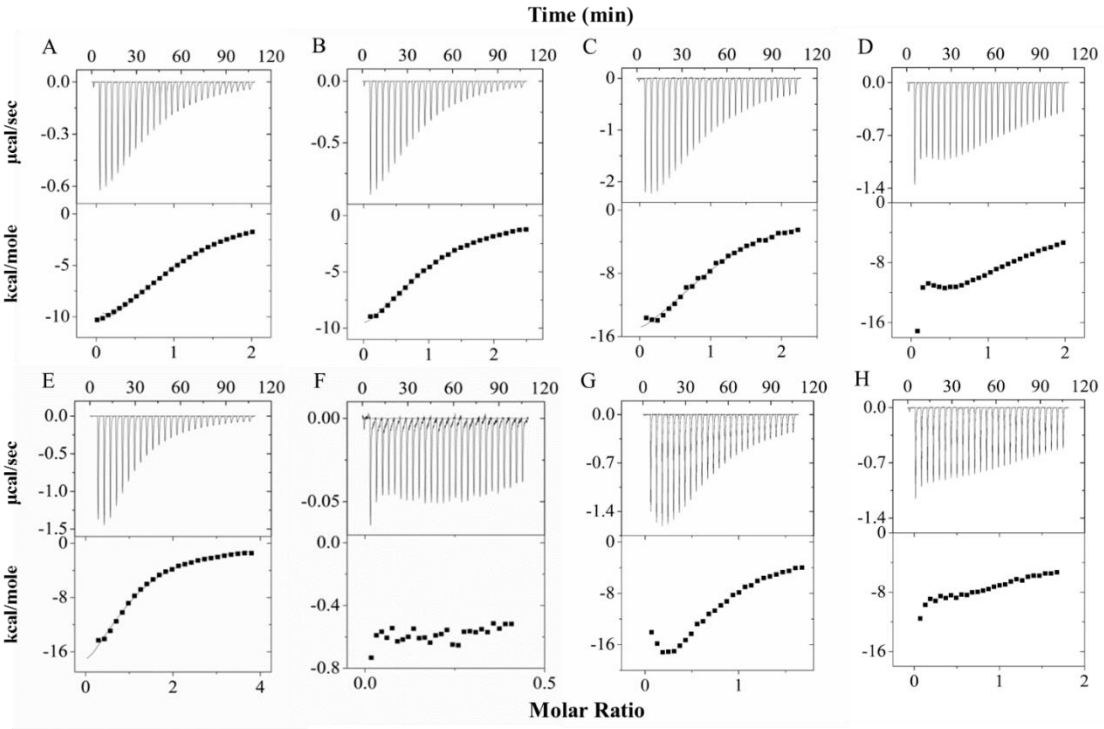


Figure 4.3. Interactions of IC isoforms with p150*. Representative thermograms (top) and binding isotherms (bottom) from ITC titrations with p150* of (A) IC-1A$_{112}$; (B) IC-1A$_{112}$S100D; (C) IC-2C$_{44}$; (D) IC-2C$_{96}$S84D*L1$_{1A}$*; (E) IC-2C$_{96}$; (F) IC-2C$_{96}$S84D; (G) IC-2C$_{96}$S84D*K41A*; and (H) IC-2C$_{83}$. Data were collected at 25 °C, pH 7.5. Thermodynamic parameters derived from fits to a single-site binding model are given in Table 4.1.

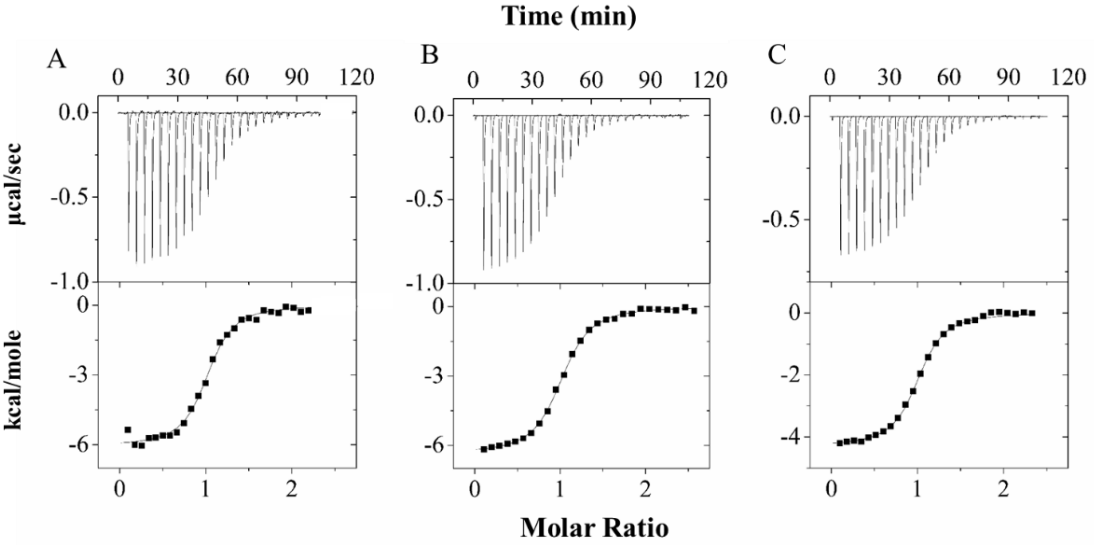


Figure 4.4. Interactions of IC isoforms with nNudel. Representative thermograms (top) and binding isotherms (bottom) from ITC titrations of (A) IC-2C_{96}; (B) IC-2C_{96}S84D and (C) IC-2C_{44} with nNudel, collected at 25 °C, pH 7.5. Thermodynamic parameters are given in Table 4.1.

Table 4.1. Thermodynamics parameters for the IC-p150* and IC-nNudel interactions recorded at 25 °C

Titrant	Protein	n	Kd	ΔG°	ΔH°	TΔS°
			uM		*kcal/mol*	
IC-1A_{112}	p150*	1.0±0.0	6.8±0.9	-7.1±0.1	-21.7±8	-14.6±8
IC-2C_{96}	p150*	1.1±0.1	12.6±3	-6.7±0.1	-24.8±3	-18.1±3
IC-1A_{112}S100D	p150*	1.0±0.0	10.8±1	-6.8±0.1	-13.1±1	-6.4±1
IC-2C_{96}S84D	p150*		ND			
IC-2C_{44}	p150*	1.0±0.0	16.7±2	-6.5±0.1	-15.8±5	-9.3±5
IC-2C_{83}	p150*		ND			
IC-2C_{96}	nNudel	1.0±0.0	0.9±0.0	-8.3±0.0	-6.2±0.1	2.1±0.1
IC-2C_{96}S84D	nNudel	1.0±0.1	1.0±0.1	-8.2±0.0	-6.4±0.0	1.8±0.0
IC-2C_{44}	nNudel	1.0±0.0	1.0±0.0	-8.2±0.0	-4.3±0.0	3.9±0.0

Figure 4.5. Comparison of WT and phosphomimetic IC-2C$_{96}$ interactions. (A) Superposed ^{1}H-^{15}N HSQC spectra of IC-2C$_{96}$ in the apo (black) and nNudel-bound (cyan) forms. (B) Superposed spectra of IC-2C$_{96}$ in the apo (black) and p150*-bound (red) forms. (C) Superposed spectra of the phosphomimetic IC-2C$_{96}$S84D in the apo (black) and p150*-bound (green) forms. (D)–(F) Changes in chemical shifts in IC-2C$_{96}$ or IC-2C$_{96}$S84D upon addition of nNudel or p150* are shown as bar graphs. (G)-(I) Changes in intensity in nNudel- or p150*-bound compared to free IC-2C$_{96}$ or IC-2C$_{96}$S84D are shown as bar graphs. All spectra were collected at 900 MHz. Samples for ligand-bound forms are at an IC: ligand molar ratio of 1:1.2. A number of peaks broadened out in spectra of the bound forms.

Figure 4.5. (*Continued*)

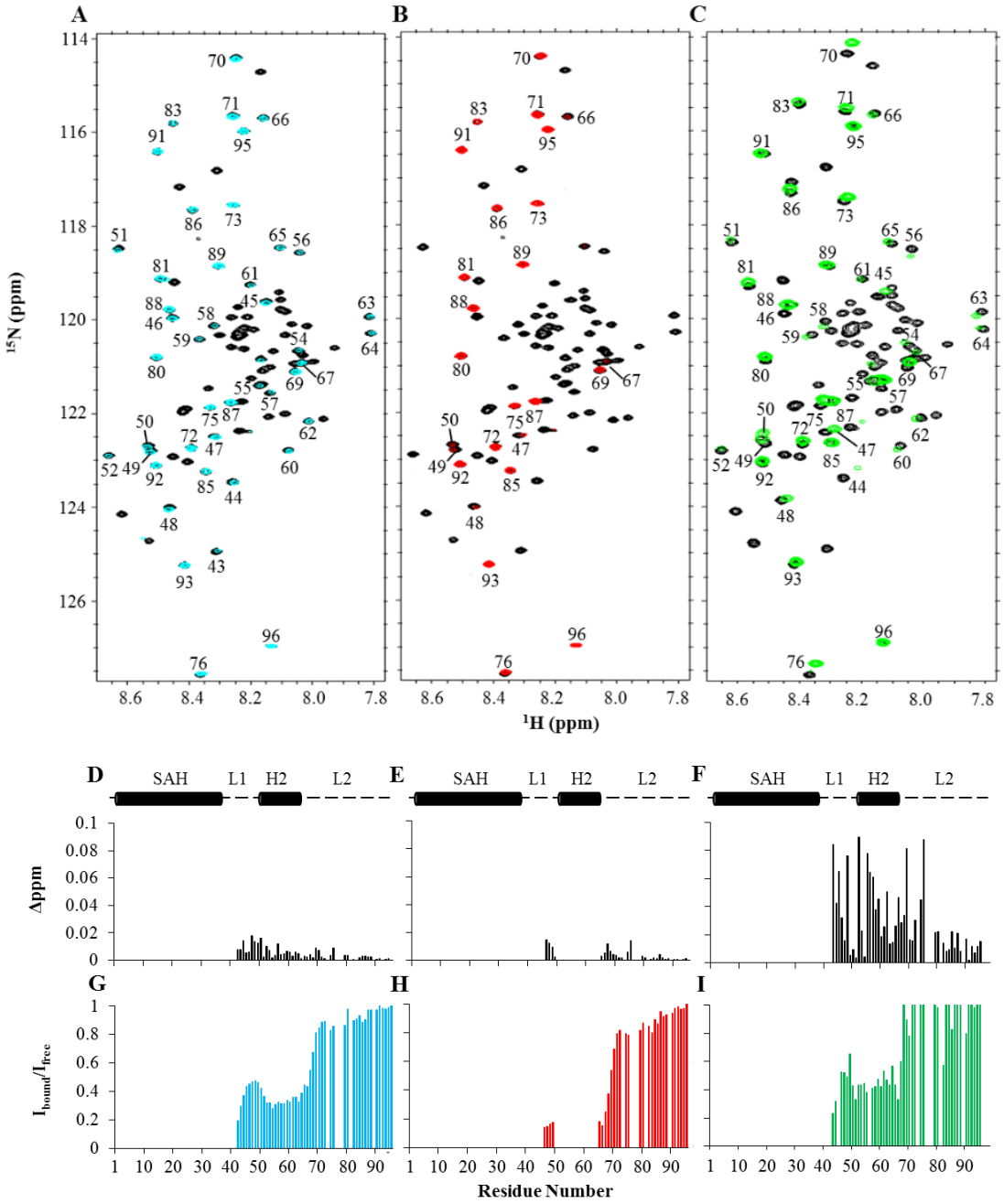
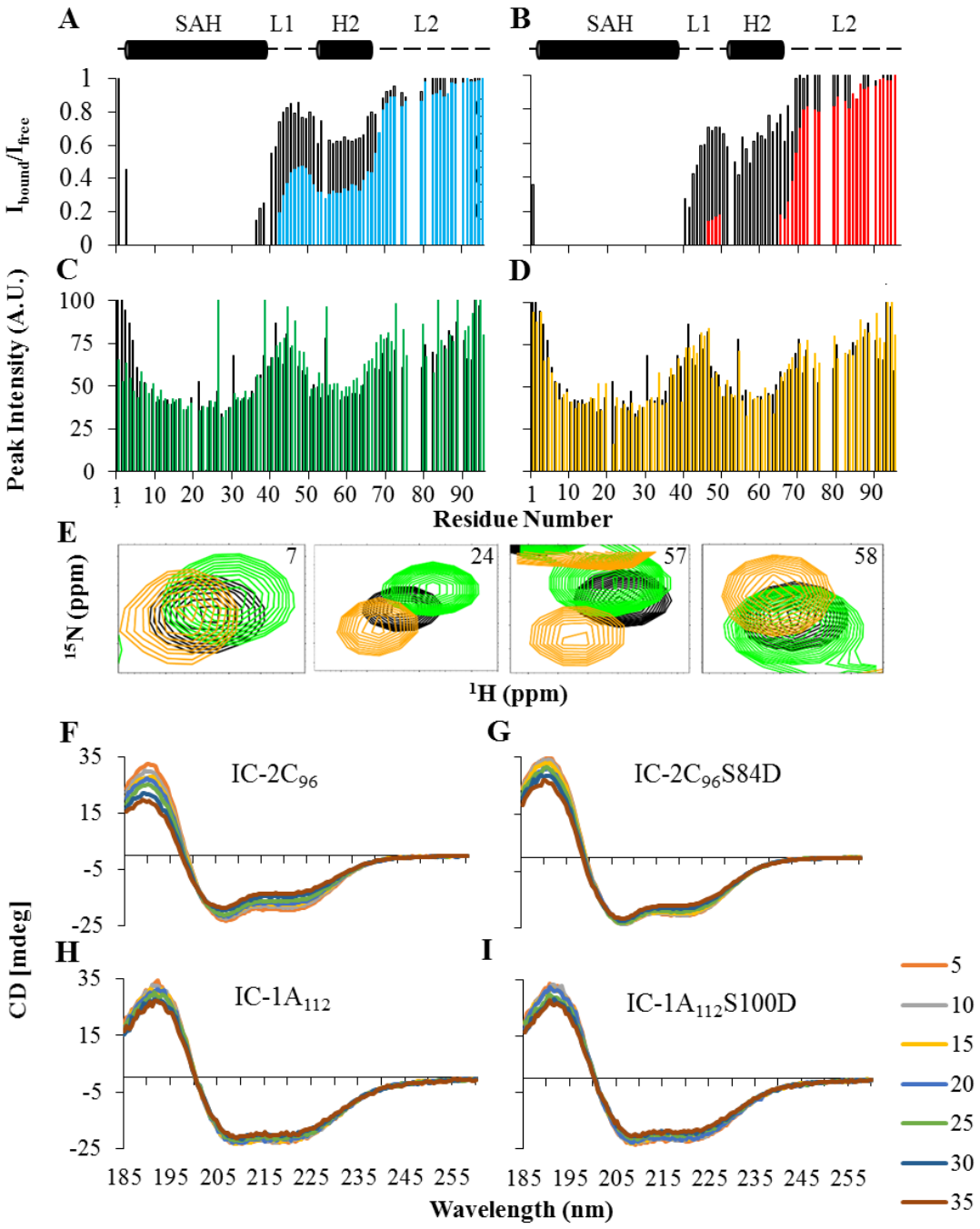


Figure 4.6. Helix packing inferred from field-dependence of NH peak intensities and from temperature-dependence as measured by circular dichroism spectra. (A) Comparison of peak intensity of nNudel-bound IC-2C_{96} at 700 MHz (black) and 900 MHz (cyan, as in Fig. 4.5G). (B) Comparison of peak intensity of p150*-bound IC-2C_{96} at 700 MHz (black) and at 900 MHz (red, as in Fig. 4.5H). Intensity of I_{bound} is expressed relative to the intensity of the same NH peak in I_{free}. (C) Overlay of peak intensities in IC-2C_{96} (black) and IC-2C_{96}S84D (green) collected at 900 MHz. (D) Overlay of peak intensities in IC-2C_{96} (black) and IC-2C_{96}S84D*K41A* (orange) collected at 900 MHz. (E) Spectral overlays of individual peaks in 900 MHz HSQC spectra that show chemical shifts perturbation. Selected peaks from SAH and H2 are shown for IC-2C_{96} (black), IC-2C_{96}S84D (green), and IC-2C_{96}S84D*K41A* (orange). (F)-(I) CD spectra of IC-2C_{96} (F), IC-2C_{96}S84D (G), IC-1A_{112} (H), and IC-1A_{112} S100D (I) at increasing temperatures in the 5 °C- 35 °C range. Significant differences are only observed for the IC-2C isoform with IC-2C_{96}S84D showing higher stability.

Figure 4.6. (*Continued*)



field-dependence of peak attenuation implies that, with p150*, residues in L1 and H2 sample multiple conformations in intermediate exchange that are different from those in the apo- and in the nNudel-bound IC.

Effect of Phosphomimetic Mutation of IC on Binding to p150 and nNudel.* Here we introduce phosphomimetic mutations used in in vivo experiments, at position 84 in IC-2C and the corresponding position in IC-1A and compare the effects on their binding affinity for p150* versus nNudel, relative to WT affinity for p150* versus nNudel. Constructs employed are IC-1A_{112}S100D and IC-2C_{96}S84D, for both of which ITC binding data show widely different affinities with p150*: IC-1A_{112}S100D binds p150* with an affinity similar to that for WT (K_d of 10.8 μM, Fig. 4.3B and Table 4.1), while IC-2C_{96}S84D binds p150* with affinity so reduced that no binding is detected by ITC (Fig. 4.3F) but detected by NMR at concentrations of 300 μM (see below). In contrast, ITC data indicate that IC-2C_{96}S84D binds nNudel (Fig. 4.4B) with the same affinity as WT.

Since the outstanding binding affinity effect of phosphomimetic mutation (Table 4.1) is on isoform IC-2C_{96} affinity for p150*, this effect was further studied by NMR. The ^{1}H-^{15}N HSQC spectra of IC-2C_{96}S84D bound to p150* (green, Fig. 4.5C) show complete peak attenuation for residues 1-43, as observed with IC-2C_{96}, but only some peak attenuation for residues 44-69 (Fig. 4.5I), contrary to that observed with IC-2C_{96}. Instead, this region shows considerable chemical shift perturbation (Fig. 4.5F), not observed in the WT IC-2C_{96} with either nNudel or p150* (Fig. 4.5D, E). Chemical shift perturbation as opposed to peak disappearance suggests faster conformational exchange in IC-2C_{96}S84D binding (sharper peaks) relative to WT (broad peaks), consistent with weaker binding.

In summary, a phosphomimetic mutation of IC suggests that the phosphorylation effect on p150* binding is isoform-specific: minimal effect with IC-1A, but significantly reduced affinity with IC-2C. Evidence of binding at NMR concentrations suggest presence of small population of p150* accessible conformation in the ensemble. Further, the resulting ensemble has no effect on binding to nNudel.

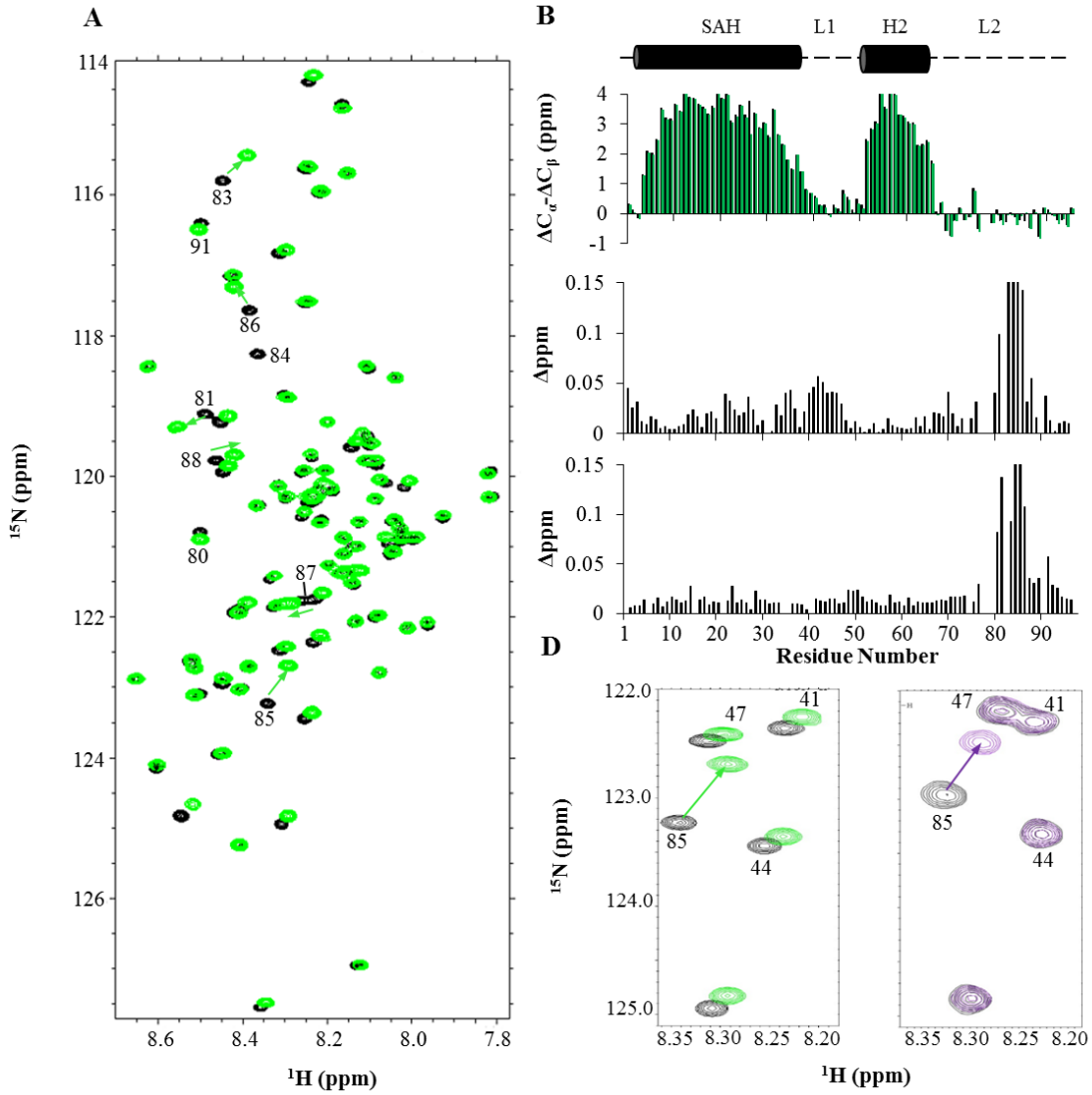
Structure and Dynamics of IC-2C$_{96}$S84D. To investigate the structural basis for reduced binding of p150* to the IC-2C phosphomimetic protein, we further compared HSQC spectra and secondary chemical shifts to those of the WT protein (Fig. 4.7A, B). Of the peaks that shift (green arrows in Fig. 4.7A), the largest effect is for residues located near the mutation site, as expected. However, peaks for residues distant from the mutation also undergo a shift, most notably residues 41 and 42 in L1 (Fig. 4.7C, top). These shifts considerably decrease in high salt buffer (200 mM NaCl) (Fig. 4.7C, bottom) suggesting some long-range charge-charge interactions are destabilized in high salt. Secondary chemical shifts indicate no change in secondary structure in S84D compared to WT (Fig. 4.7B).

An increase in tertiary contacts in IC-2C$_{96}$S84D is inferred from a small increase in its CD-detected thermal stability relative to WT (Fig. 4.6F, G). The negative ellipticity indicative of partial helical structure is significantly reduced at elevated temperature in the WT compared to the S84D mutant which shows very little change in the 5-35˚C temperature range suggesting presence of a partially stable protein core presumably formed by packing of SAH against H2. No change in stability is observed with the IC-1A$_{112}$ phosphomimetic mutant (Fig. 4.6H, I), suggesting that phosphorylation does not increase helical packing in this isoform.

Backbone dynamics of IC-2C$_{96}$ WT and IC-2C$_{96}$S84D determined from T_1, T_2 and steady-state heteronuclear NOE experiments (Fig. 4.8) show overall similar mix of ordered and disordered structures. T_2 is lowest for SAH and H2 (0.07 s and 0.09 s, respectively) compared to the overall average value of 0.20 s, consistent with nonrandom structure for these segments (Fig. 4.8B, E). Steady-state heteronuclear NOE measurements confirm that the SAH and H2 are the most ordered segments with positive average NOE values of 0.47 and 0.54, respectively (Fig. 4.8C, F), supporting the secondary chemical shifts identification of helical structures in these regions, while the linkers are fully disordered exhibiting near zero or negative NOE values. Interestingly, in IC-2C$_{96}$S84D a small increase in ordered structure is detected in L1 residues 43-47 which become slightly more positive (negative NOE average value of

Figure 4.7. Conformational effects of phosphomimetic S84D mutation on IC-2C. (A) ^{1}H-^{15}N HSQC spectral overlay of IC-2C_{96} (black) and IC-2C_{96}S84D (green) in buffered 50 mM NaCl. Peaks that shift the most are highlighted by green arrows and labeled. (B) Comparison of secondary chemical shifts of IC-2C_{96}S84D (green) and IC-2C_{96} (black) indicate unaltered secondary structure in WT *vs* mutant construct. (C) Salt effects on chemical shift changes in IC-2C_{96}S84D. Bar graphs show residue number *versus* chemical shift changes (Δppm) of the same NH peak in spectra of IC-2C_{96} *vs* IC-2C_{96}S84D, for samples in 50 mM NaCl (*top)* and 200 mM NaCl (*bottom).* (D) Enlarged regions of spectra in (A) for samples in lower (l*eft*) and higher (*right*) salt.

Figure 4.7. (*Continued*)



0.07 in WT to small positive NOE average value of 0.02, Fig. 4.8C, F).

Backbone dynamics differences are also inferred from field-dependence of peak intensities, which measure motional restriction on the intermediate timescale. Comparison of peak intensities of IC-2C_{96} in HSQC spectra collected at 900 and 700 MHz (Appendix 1, Fig. A1.2A) show SAH and H2 residues at 900 MHz are 20% and 70% less intense than those at 700 MHz, respectively. This intermediate exchange involving SAH and H2 is presumably due to domain movement facilitated by the disordered L1 linker connecting them. Residues with highest field-dependence difference are mapped onto the structure (Appendix 1, Fig. A1.2B). The IC-2C_{96}S84D peak intensities at 900 MHz show similar trend as WT with peak intensities slightly higher in H2 (Fig. 4.6C) but with major decrease in intensity in the first few residues preceding SAH. Both changes can be interpreted as due to formation of a more compact conformation in IC-2C_{96}S84D: increase in signal intensity in H2 as due to change in motional heterogeneity from multiple conformations to primarily one, and decrease in signal intensity at the N-terminus as due to a change from high flexibility in the disordered N-terminal residues to restricted motion. Both changes are considerably distant from the site of phosphorylation.

In summary, dynamics measurements identify SAH and H2 as the only ordered segments in the protein, while the linkers are fully disordered. ^{1}H-^{15}N steady-state NOEs and T_1/T_2 ratios suggest an increase in order in IC-2C_{96}S84D localized to the first part of SAH, and to a lesser extent the C-terminal end of the SAH and few residues in L1. Significantly, the phosphomimetic mutation does not change the secondary structure in the vicinity of the mutation, but causes very small changes in L1 and N-terminal SAH, and considerable changes in the domain motion of SAH and H2.

Rescue of p150 binding in S84D mutants by either a K41A or E42A mutation.* The S84D mutation in IC-2C affects chemical shifts and dynamics of L1 in a salt-dependent manner, suggesting the shielding of a charge-charge interaction. To test this idea, we introduced single site mutations that replace either the positive charge of K41or the negative charge of E42 with alanine. These mutations when introduced one at a time

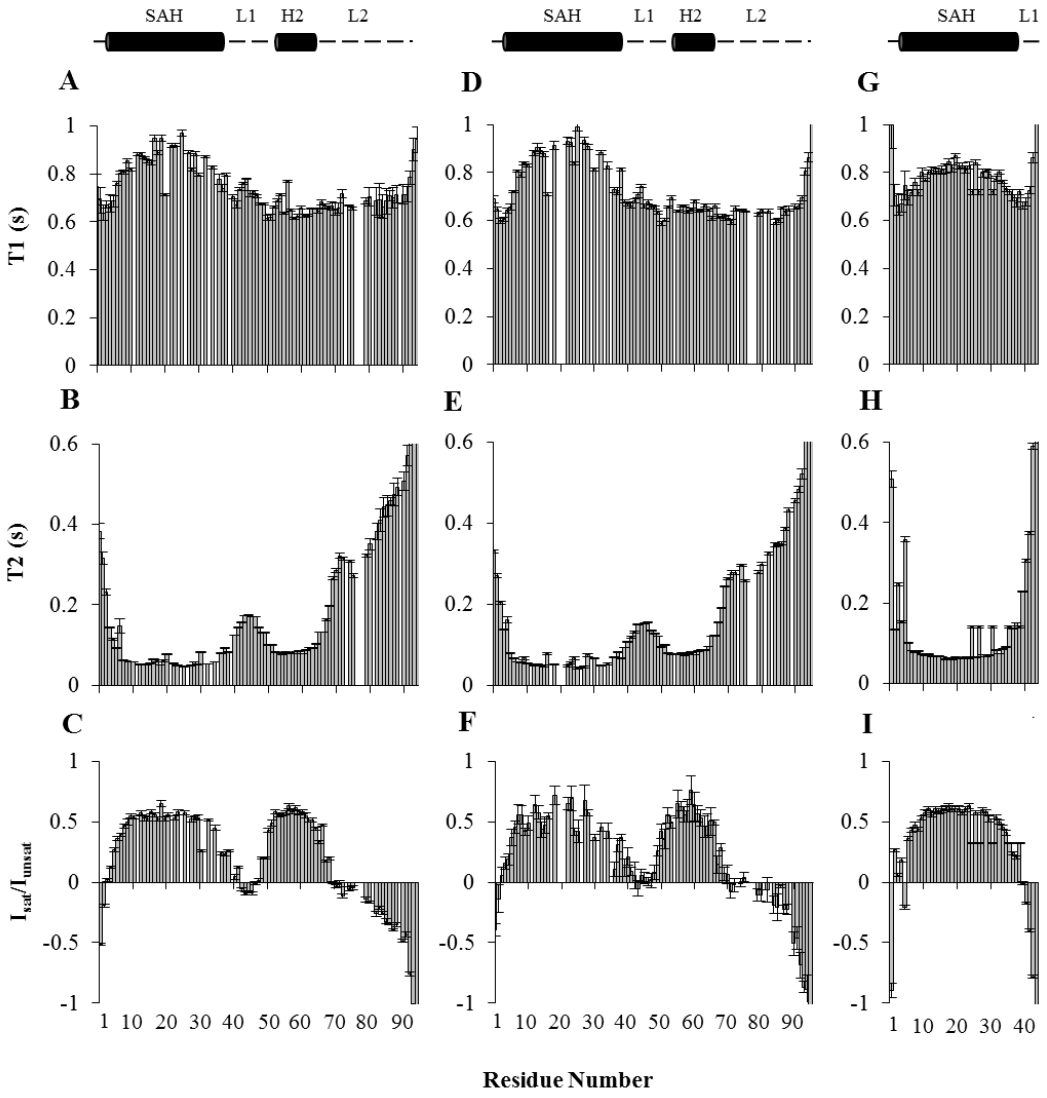


Figure 4.8. Dynamics of isoform 2C in WT, phosphomimetic and truncated constructs. For IC-2C_{96} (A-C), IC-2C_{96}S84D (D-F) and IC-2C_{44} (G-I), longitudinal relaxation (T_1), transverse relaxation (T_2) and steady-state heteronuclear NOE values are plotted as a function of residue number.

in the IC-2C_{96}S84D to generate two double mutants show similar regain of affinity with p150* (Fig. 4.3G). The single mutants K41A and E42A show similar binding to p150* as the WT IC-2C (data not shown) confirming that the enhanced binding of S84D*K41A* relative to S84D is not due to specific interaction of A41 with p150*.

Comparison of HSQC spectra of IC-2C_{96}S84D*K41A* double mutant with the WT spectra show that the SAH and H2 residues undergo the most significant chemical shift difference (Fig. 4.6E). Interestingly, a shift in S84D*K41A* (orange) in the opposite direction to S84D (green) suggests a change in the packing of SAH and H2 for S84D*K41A* that is similar to that in WT. Peak intensities at 900 MHz for S84D*K41A* are similar to WT and especially significant is the gain in intensity for peaks at the N-terminus (Fig. 4.6D) compared to S84D (Fig. 4.6C). Supporting a more extended form in S84D*K41A* is the similarity of CD-detected thermal stability to WT (data not shown) and the similarity in direction of chemical shifts to IC-2C_{44}, which lacks H2 and is therefore a model for fully extended IC.

Effect of Amino Acid Composition of IC-2C Linker 1 on IC-p150^{Glued} Interactions. To investigate the role of specific residues in L1 on p150* binding, we generated a chimera, IC-2C_{96}S84D*L1*$_{1A}$, in which L1 residues (39-51) of IC-2C_{96}S84D were replaced by L1 residues (39-50) of IC-1A_{112}, and measured its binding to p150* (Fig. 4.3D). ITC experiments show some restored binding to p150* but to a lesser extent compared to IC-2C_{96}S84D*K41A* (Fig. 4.3G). Also as with the rescue mutants above, the IC-2C_{96}S84D*L1*$_{1A}$ thermal stability is similar to WT rather than S84D (data not shown) confirming the effect of residues within L1 on the degree of SAH/H2 packing, and also suggesting that in IC-1A, the longer L2 contributes to the higher affinity, and compensates for the negative effect of the naturally occurring Asp at position 84, the same position of phosphorylation in IC-2C.

In summary, these results, together with the absence of binding in IC-2C_{83} suggest a model for IC-2C_{96}S84D that is compact and stabilized by electrostatic interactions between L1 and SAH, and between the phosphorylation site in L2 and N-terminal SAH. Chemical shift differences in IC-2C_{96}S84D compared to WT mapped on

a CS-Rosetta model (Fig. 4.9A, largest in magenta to smallest in yellow) suggest significant long-range effects on C-terminal SAH and L1 and on the N-terminal linker preceding SAH. This compact conformation was chosen out of many low energy structures as a model for a compact structure. Additional support for this model is the slightly more restricted mobility when compared to WT (magenta for higher heteronuclear NOEs and T_1/T_2 ratios), showing that restricted motion is localized to both linkers, and to the N-terminal of SAH (Fig. 4.9B, C).

Discussion

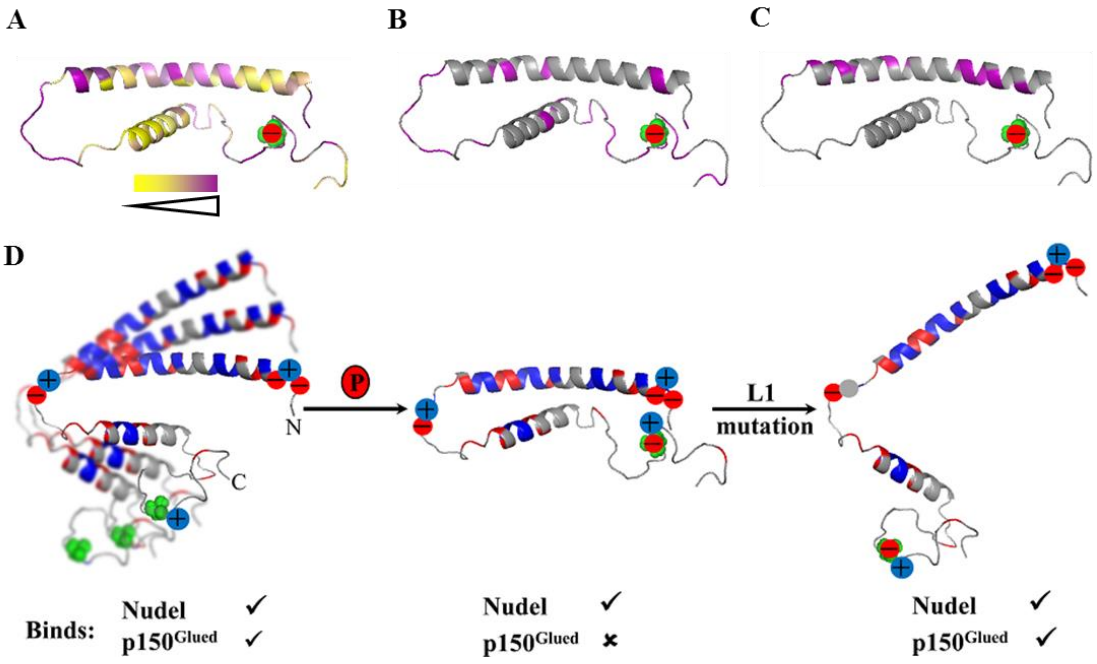
Dynactin p150^{Glued} and NudE/Nudel Bind IC at A Conserved N-terminal SAH Domain. The dimeric coiled-coil domain of p150^{Glued} (a component of dynactin) or NudE/Nudel modulates dynein activity by binding IC at its N-terminal SAH domain. The highly charged SAH (71% R, K, E) (Fig. 4.2) stabilized by intrahelical electrostatic interactions between residues i and i+4 (e.g. K8-E12) and i and i+3 (e.g. E10-R13) is also observed in yeast and *Drosophila* IC (Jie et al., 2015; Morgan et al., 2011). In all three species, a stretch of 24-35 amino acids forms a stable single helix identified by NMR secondary chemical shifts, lower T_2 relaxation, and CD, and is also the binding site for both p150^{Glued} and Nudel.

While ITC experiments on smaller constructs identify the SAH domain as sufficient for binding Nudel and p150^{Glued} (Table 4.1), the significant peak attenuation of H2 residues at 900 MHz relative to 700 MHz only with p150^{Glued} binding suggest a more complex binding process with p150^{Glued} that also involves residues within H2.

Regulation of IC Binding to p150^{Glued} By Phosphorylation Is Specific to The IC-2C Isoform. The alternatively spliced and tissue-specific IC-1A and IC-2C isoforms differ in the composition of the disordered linker L1 connecting SAH and H2 and in the length of linker L2 connecting H2 and the Ser-rich region. Cell extract studies (Prof. Erika Holzbaur, personal communication) of p150^{Glued} binding to IC isoforms show a significantly higher affinity of p150^{Glued} for IC-1A relative to IC-2C, suggesting that isoform expression is a key aspect of regulation. In IC-2C, phosphorylation at Ser84

Figure 4.9. Model for effect of phosphomimetic mutation on dynamic structure and interactions of IC-2C$_{96}$. (A) A CS-Rosetta structure of IC-2C$_{96}$ highlighting chemical shift perturbation of residues in IC-2C$_{96}$S84D (largest to smallest, purple to yellow). Highlighted in purple are residues that become more ordered in IC-2C$_{96}$S84D based on ^{1}H-^{15}N steady-state NOE values (B) and ratios of T_1/T_2 values (C). The S84D mutation is represented by a negative charge (red circle) on top of the serine residue (green spheres) in all 3 structures. (D) An ensemble model for the effect of phosphomimetic mutation on interactions of IC-2C$_{96}$ with p150* and nNudel. Charged amino acids are indicated (positive charge, red; negative charge, blue) and colored circles show the positions of D2, K3, E5, K41, E42, and K85. In apo-IC-2C$_{96}$ (left), the SAH domain and H2 loosely pack against each other in multiple conformations in intermediate exchange that are accessible to either p150* or nNudel. Upon phosphomimetic mutation or phosphorylation (red circle) of S84, the introduced negative charge interacts with residues at the N-terminus, shifting the population towards more compact conformations (center), which bind Nudel but not p150*. The binding of p150* is restored in either of two 'rescue' mutations, K41A or E42A; the former illustrated on the right.

Figure 4.9. (*Continued*)



distant from the recognition SAH domain or substitution of Ser84 by Asp results in diminished binding to $p150^{Glued}$ in cells (Vaughan, 2001), and the ITC data in Fig. 4.3 confirm a significant reduction in $p150^{Glued}$ binding to the phosphomimetic mutant. No phosphorylation or phosphomimetic mutant studies are reported on IC-1A isoform but ITC experiments with the IC-1A phosphomimetic S100D mutant show similar binding affinity as the WT (Fig. 4.3 and Table 4.1), suggesting that Ser100 phosphorylation of IC-1A does not affect its binding to $p150^{Glued}$ or Nudel. Since our ITC studies on recombinantly expressed unmodified proteins (Fig. 4.3) show similar affinity for unphosphorylated IC-1A *versus* IC-2C, we predict that cell extracts have a high population of phosphorylated IC-2C giving rise to the observed lower $p150^{Glued}$ affinity. Based on the different binding affinities between the isoforms in cell extracts and between the recombinantly expressed phosphomimetic forms of IC-2C and IC-1A, we propose that the phosphorylation effect on IC-$p150^{Glued}$ binding is specific for the shorter IC-2C isoform and has no effect on IC-1A.

Another consequence of phosphorylation in IC-2C is its effect on the selective binding of regulatory proteins that share a common IC binding site. Both $p150^{Glued}$ and NudE/Nudel are localized in cellular compartments with dynein (Vallee et al., 2012) and even though both bind at the N-terminal SAH, the S84D mutation significantly diminishes binding to $p150^{*}$ but has no effect on binding to nNudel (Fig. 4.3, 4.4), suggesting phosphorylation as a mechanism that promotes binding of NudE/Nudel when dynactin is present. Since all IC constructs of various mutations and size show similar binding affinity to Nudel as long as they contain the SAH domain, phosphorylation at residues S88/T89 should not reduce interaction with NudE as previously reported (Gao et al., 2015).

Structural Basis of the Effect of IC-2C Phosphorylation on Dynactin Binding. To elucidate how phosphorylation in a disordered linker at a distant site from binding results in reduced binding of IC-2C to $p150^{Glued}$, we compare CD and NMR data of the phosphomimetic IC-$2C_{96}$S84D with that of the WT protein. CD spectra of the two proteins are indistinguishable and indicate partial helical structure at low temperature,

and so do NMR secondary chemical shifts. While there is no detected increase in secondary structure in the disordered linker containing the S84D mutation, several pieces of evidence suggest formation of a more compact structure. With CD, IC-2C_{96}S84D shows higher thermal stability suggesting a more compact core presumably formed from packing of SAH against H2. With NMR, salt-dependent chemical shifts suggest phosphorylation-induced stabilization of long-range electrostatic interactions between L1 and C-terminal SAH residues, and between L2 and the first five N-terminal residues (Fig. 4.9A). In addition, ^{15}N heteronuclear NOE measurements show a small increase in ordered structure localized to L1, L2, and the first turn of the SAH helix (Fig. 4.9B). Similarly, T_1/T_2 ratios, which measure motional heterogeneity, show the largest differences for the N-terminal region of SAH (Fig. 4.9C) suggesting their involvement in long-range interactions. Further evidence of a stable electrostatic interaction between the phosphorylation site and the N-terminal residues is the observation that IC-2C_{83} does not bind p150*. A charged carboxyl group at the C-terminal appears to stabilize similar tertiary contacts as in S84D, forming a more compact conformation that is not accessible to p150* binding.

Length and Sequence Specificity of Disordered Linkers Modulate Phosphorylation Effects. In the IDP literature, phosphorylation events more commonly regulate binding by disrupting 3D structures (Metcalfe et al., 2005; Mitrea et al., 2014), by inducing electrostatic interactions with the partner protein, or by altering a structure in the vicinity of the binding site (Bah et al., 2015). Here, phosphorylation (using S84D as a phosphomimetic) induces an electrostatic cluster that docks the linker L2 containing the phosphorylation site against the short disordered segment preceding SAH. Taken together the ITC, NMR and published *in vivo* data (Vaughan, 2001) favor the conclusion that the reduced binding of p150^{Glued} to S84D is not because of a change in the structure or dynamics at the phosphorylation site, nor of a change in structure or dynamics at the binding site, but because of electrostatic interactions between D84/K85 in L2 and D3/K4/D6 at the N-terminus that favor closed conformations and diminish accessibility of p150^{Glued} to IC, as illustrated in Fig. 4.9D.

Our model for the modulation of phosphorylation effect by sequence specificity of the linkers involves ensembles of multiple conformations. Unphosphorylated IC (Fig. 4.9D, left) has varying degrees of SAH/H2 packing interactions associated with domain movement around linker L1. Phosphorylation of Ser84 shifts the average to conformations that pack SAH against H2, due to the combined stabilization arising from two clusters of electrostatic interactions, one between L1 residues K41 and E42 (circles) and charged residues at the C-terminal end of SAH, and another between L2 residues D84 and K85 and the N-terminal SAH residues 3-6 (Fig. 4.9D, middle). Disruption of either cluster is sufficient to shift the equilibrium away from the closed conformation (Fig. 4.9D, right and left). Evidence of stabilization of compact conformations also comes from the small increase in peak intensities at 900 MHz for H2 of IC-2C_{96}S84D relative to WT (Fig. 4.6C), suggesting less heterogeneous dynamics in IC-2C_{96}S84D. The packing interactions between SAH and H2 are in turn modulated by L1. A substitution of a charged residue with Ala abolishes the stabilizing interaction between L1 and C-terminal SAH leading to a more extended structure and dissipation of the electrostatic clustering of L2 and N-terminal SAH, thus 'rescuing' IC binding to $p150^{Glued}$ (Fig. 4.9D, right).

The sequence specificity and length of disordered linkers separating functional SAH and Ser-rich region domains control domain movements that stabilize one conformation over another. In IC-2C, phosphorylation at S81 and S83 is not sufficient to reduce binding (Vaughan, 2001) presumably due to the absence of a neighboring charged residue as in S84/K85. In IC-1A, the presence of an extended-promoting PVP triplet in L1, the absence of a charged residue next to D84 in L2, and longer L2 all disfavor electrostatic interactions between L2 and the N-terminal residues preceding SAH, thus stabilizing an open conformation accessible to $p150^{Glued}$ even in the S100D phosphomimetic IC-1A. These studies give unprecedented structural understanding of the interplay between phosphorylation and alternative splicing of dynein intermediate chain in dynein regulation, and more broadly, offer new dimensions to mechanistic and structural explanations of how phosphorylation and alternative splicing in intrinsically

disordered regions increase the functional diversity of large complexes. A model explaining these long-range allosteric interactions opens up numerous new opportunities to modulate the functionalities of protein complexes.

Experimental Procedures

Expression Construct and Protein Production. Recombinant rat IC isoforms 1A and 2C, human p150^{Glued}, and human Nudel were gifts (Erika Holzbaur, University of Pennsylvania). DNA fragments encoding residues corresponding to IC-1A$_{112}$ and IC-2C$_{96}$ were prepared by PCR and were subsequently cloned into a pMCSG9 vector using ligation independent cloning as previously reported (Nyarko et al., 2013). Site-directed mutagenesis of Ser100 and Ser84 to aspartic acid to generate IC-1A and 2C phosphomimetic mutants, K41A and E42A on IC-2C$_{96}$S84D, and K41A on IC-2C$_{96}$ were performed using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies) following the manufacturer's protocol. IC-2C$_{44}$, IC-2C$_{83}$, p150^{Glued} 382-531, and Nudel 1-189 were prepared by PCR and subsequently cloned into a modified pET15da vector (Jie et al., 2015), a pPROEX vector with an N-terminus His$_6$ tag, and a pET15b vector with a C- terminus His$_6$ tag, separately. The IC-2C*L1$_{1A}$* gene was produced by Gene-Script (Piscataway, NJ). All sequences were verified by automated sequencing. Recombinant vectors were transformed into Rosetta (DE3) cell lines for protein expression.

The transformed cell lines were grown in LB media at 37 °C to an optical density (A600) of 0.6-0.8 followed by protein induction for 3-16 h with 0.1-0.4 mM IPTG. Cells were harvested, lysed, and centrifuged to remove all cell debris. The soluble fractions were purified using Qiagen's Ni-NTA affinity chromatography protocol (Qiagen Valencia, CA) or Talon's metal affinity resin (Clontech). The His$_6$-fusion tags were cleaved by TEV protease. All proteins were further purified on a Superdex^{TM} 75 (GE healthcare) size exclusion chromatography column, resulting in purity > 95 %. Proteins were dialyzed against appropriate buffers, and their concentrations were determined from absorbances at 280 nm and computed (expasy.org/protparam) molar extinction

coefficient values. The concentrations of IC-2C constructs were determined using Bradford assay and comparison of their intensities on polyacrylamide gels with intensities of lysozyme loaded with known concentrations. Reported concentrations are those for the monomeric proteins. The computed $p150^{Glued}$ 382-531 and Nudel 1-189 molar extinction coefficients take into account the His_6 sequence from the expression vector. The proteins were subsequently stored at 4 ^{o}C and used within a week.

Isothermal Titration Calorimetry. ITC experiments using Microcal's VP-ITC microcalorimeter (North Hampton, MA), were performed at 25 °C in buffer composed of 50 mM sodium phosphate, 50 mM sodium chloride, (with 5 mM β-mercaptoethanol in IC-1A constructs that have a single cysteine), and 1 mM sodium azide, pH 7.2-7.5. A typical experiment involved an initial 2 μL injection, followed by 26 (10 μL) injections of titrants to proteins in the concentration range of 20-40 μM, accompanied by stirring at a constant rate. Added titrants were in the concentration range of 0.25 – 0.4 mM. Protein samples and buffer were degassed prior to data acquisition. Data were processed using Origin 7.0 and fit to a single-site binding model. Recorded data are the average of two or more independent experiments, with the uncertainty reported as the difference between the experimental value and the average.

NMR Measurements. Doubly labeled $^{15}N/^{13}C$ and ^{15}N labeled proteins were prepared using published protocols (Makokha et al., 2004). Protein samples were prepared in buffer composed of 10 mM sodium phosphate at pH 6.0 or pH 7.0 with 50 mM NaCl or 200 mM NaCl, 1 mM sodium azide, 8 % $^{2}H_2O$, a protease inhibitor mixture (Roche Applied Science), and 2,2-dimethylsilapentane-5-sulfonic acid for ^{1}H chemical shifts referencing. For backbone assignments of IC-$2C_{96}$, NMR spectra were obtained at 15 °C and concentration of 350 μM. Native PAGE and ^{1}H-^{15}N HSQC spectra were used to assay sample integrity before and after every NMR data collection. All reported data are for HSQC spectra that remain unchanged during the data collection time.

Backbone resonance assignments were determined from a set of [^{15}N,^{1}H]-TROSY (Pervushin et al., 1997; Salzmann et al., 1998) based triple-resonance experiments featuring band selective-excitation-short-transient (BEST) (Favier and

Brutscher, 2011; Schanda et al., 2006) methodology to reduce measurement times. The HNCACB, HA[HB, HN](CACO)NH (Löhr et al., 1999), HN(CO)CACB, HN(CA)CO and HNCO experiments were carried out with a cryo-probe on Bruker Avance 900 MHz and 950 MHz spectrometers, respectively.

Longitudinal (T_1) and transverse (T_2) ^{15}N relaxation times of the IC-2C$_{96}$ and IC-2C$_{96}$S84D were determined at 600 MHz ^{1}H frequency using standard HSQC-based pulse sequences (Farrow et al., 1994) at 15 °C. For each series, spectra with ten different relaxation periods (three in duplicate for error estimation) were recorded in an interleaved manner. Their duration varied from 10 ms to 1.3 s in T_1 measurements, and from 22.88 ms to 366.08 ms in T_2 measurements. Steady-state ^{15}N-{^{1}H} NOE values were obtained using TROSY-based pulse sequences (Ferrage et al., 2008; Lakomek et al., 2012) and a total recovery delay of 8 s. A train of 180˚ ^{1}H pulses spaced by 22 ms was applied to protons for 6 s in the saturation experiment.

NMR experiments of ^{15}N-labeled IC-2C$_{96}$ and IC-2C$_{96}$S84D with unlabeled p150^{Glued} 382-531 and Nudel 1-189 were acquired at 15 °C on Bruker Avance 900 MHz and 700 MHz spectrometers, respectively. TROSY-HSQC spectra of ^{15}N-labeled IC-2C$_{96}$ and IC-2C$_{96}$S84D were collected at final molar ratios (IC-2C$_{96}$/S84D: p150/Nudel) of 1:1.2 at pH 7.0.

NMR Data Analysis. All spectra were processed with TopSpin (Bruker) and analyzed using Sparky (Goddard and Kneller) and NMRView (Johnson, 2004). C$_\alpha$ and C$_\beta$ secondary chemical shifts were used to determine the secondary structure (Kjaergaard et al., 2011). Peak intensities were measured as peak heights with a normalization factor determined from the signal to noise ratio of residue 96 at the C-terminal. Changes in peak intensities I_{Bound}/I_{Free} were calculated as the ratio of signal intensity in spectra of the complex and free IC-2C$_{96}$. T_1, T_2 and steady-state NOE values were determined as described previously (Benison et al., 2006; Morgan et al., 2011).

Chemical Shifts-Based Structures. An ensemble of structures for apo IC-2C$_{96}$ was generated using the program CS-Rosetta (Shen et al., 2008) and chemical shift assignments for the ^{1}H$_\alpha$, ^{1}H$_N$, ^{13}C$_\alpha$, ^{13}C$_\beta$, ^{13}C', and ^{15}N atoms. The CS-Rosetta

calculations were carried out on the server at the BMRB using the chemical shifts data. All structural images were generated using PyMOL (Delano, 2002). The open and closed tertiary structures (Fig. 4.9) are inferred from multiple NMR data and CD data.

Acknowledgements

The authors wish to thank Prof. Erika Holzbaur for access to unpublished data, Prof. Clare Woodward and the Barbar lab for valuable discussions. This work is supported by National Science Foundation grant MCB 0818896 to EB. We acknowledge the support of the protein core facility in the OSU Environmental Health Sciences Center (NIH/NIEHS 00210), and access to the Research Infrastructure activity in the 7th Framework Programme of the EC (Project number: 261863, Bio-NMR) (Frankfurt, Germany).

Chapter 5

Evolution of Dynein IC Regulation by Dynactin among Mammals, Drosophila, and Yeast

Jing Jie, Frank Löhr and Elisar Barbar

Abstract

The dynein intermediate chain (IC) subunit is central to the regulation of dynein motility and to the recruitment by non-dynein proteins, such as dynactin and NudE/Nudel, which participate in a number of dynein functions and bind to an overlapping site on IC. Here, using NMR, ITC and CD-detected thermal unfolding, we investigate the mechanism of how dynein selectively interacts with either regulator. We focus on the IC-dynactin and IC-NudE/Nudel interactions in three species: rat, Drosophila and yeast. We show that only dynactin binding to IC has evolved to gain complexity: in yeast, there is no significant difference between dynactin and NudE binding on assembled IC, in Drosophila, a short nascent helix close to the common recognition site contributes to regulation of dynactin binding, while in mammals, phosphorylation in a Ser-rich region distant from the common binding site on a tissue-specific IC isoform contributes to regulation of dynactin binding. Modulation of the IC-NudE/Nudel interaction is also through the altered dynactin affinity to IC in higher eukaryotes.

Introduction

Eukaryotic cells require efficient and regulated transport of intracellular components for multi-layered temporal and spatial organization. Proper organization is essential in numerous cellular processes including cell proliferation, signaling, and metabolism. An essential component of this cytoskeleton network is cytoplasmic dynein, a ubiquitously expressed 1.6 MDa motor protein complex consisting of six subunits, the heavy chain (HC) which has the motor activity, the intermediate chain (IC), the light intermediate chain (LIC) and three light chains, Tctex, LC8 and roadblock. Acting as a cellular motor, dynein transports organelles, vesicles, proteins and mRNAs on the microtubules from cell periphery towards cell center and this transport underlies a wide range of functions, including nuclear envelope breakdown, mitotic spindle assembly and orientation, Golgi complex formation and positioning, endosomes and lysosomes distribution (Chevalier-Larsen and Holzbaur, 2006; Harada et al., 1998; Vale, 2003). A

number of non-dynein proteins are proposed to regulate its motility, and target dynein to different cargos and positions. The best-studied dynein regulators dynactin and NudE/Nudel are ubiquitously expressed and present in the same cellular compartments,

Dynactin, a large 1.2 MDa protein complex, has been shown to participate in almost all aspects of dynein functions, including dynein cargo recruitment (Karki and Holzbaur, 1999), increased dynein processivity along microtubules (Culver-Hanlon et al., 2006; King and Schroer, 2000), and the initiation of dynein-mediated transport in neurons (Moughamian and Holzbaur, 2012). The complex is formed from multiple copies of 11 different subunits (Schroer, 2004), the largest of which, $p150^{Glued}$, contains an N-terminal CAP-Gly (cytoskeleton-associated protein, glycine-rich) domain for binding microtubules, and two coiled-coil domains (CC1 and CC2) (Schroer, 2004). The C-terminal half of $p150^{Glued}$ CC1 (CC1B) is sufficient to bind the N-terminus disordered region of dynein intermediate chain (Fig. 1) (McKenney et al., 2011). Studies on Drosophila and mammalian proteins identify an N-terminal α helix on IC for $p150^{Glued}$ recognition (Morgan et al., 2011; Siglin et al., 2013), and reveal the mechanism for the negative effects of IC phosphorylation on $p150^{Glued}$ binding (Jie et al., submitted).

NudE and its closely related mammalian paralogue Nudel are required for nuclear migration (Efimov and Morris, 2000), mitotic spindle assembly (Feng and Walsh, 2004) and cortical neuronal position (Shu et al., 2004). Mammalian NudE and Nudel share ~55% sequence identity and regulate a subset of dynein functions primarily during mitosis, including centrosome duplication, spindle orientation (Chansard et al., 2011) and targeting dynein to the kinetochores (Lam et al., 2010). NudE/Nudel also recruits another regulatory protein Lis1 to dynein, and the triplex has been show to facilitate high-load dynein transport (McKenney et al., 2010). An N-terminal coiled-coil domain in NudE/Nudel is responsible for binding dynein at the same SAH domain of IC that binds $p150^{Glued}$ (Nyarko et al., 2012) (Jie et al., submitted).

The multiple dynein functions mediated by $p150^{Glued}$ and NudE/Nudel, their localization in the same cellular compartments, and the overlapping binding sites on

dynein IC raise the question of how and when dynein selectively interacts with either regulator. Here we provide insights into these questions by characterizing the interactions of IC with dynactin and NudE/Nudel for rat, Drosophila and yeast with the goal of identifying common themes that govern these interactions. We show a progression in complexity with the requirements in the species: in yeast, there is no significant difference in binding in assembled IC, while in Drosophila, a short transient helix proximal to the binding interface regulates binding, while in mammalian, regulation requires phosphorylation in a Ser-rich region distant from the recognition site. All of these regulatory events only affect the binding of $p150^{Glued}$.

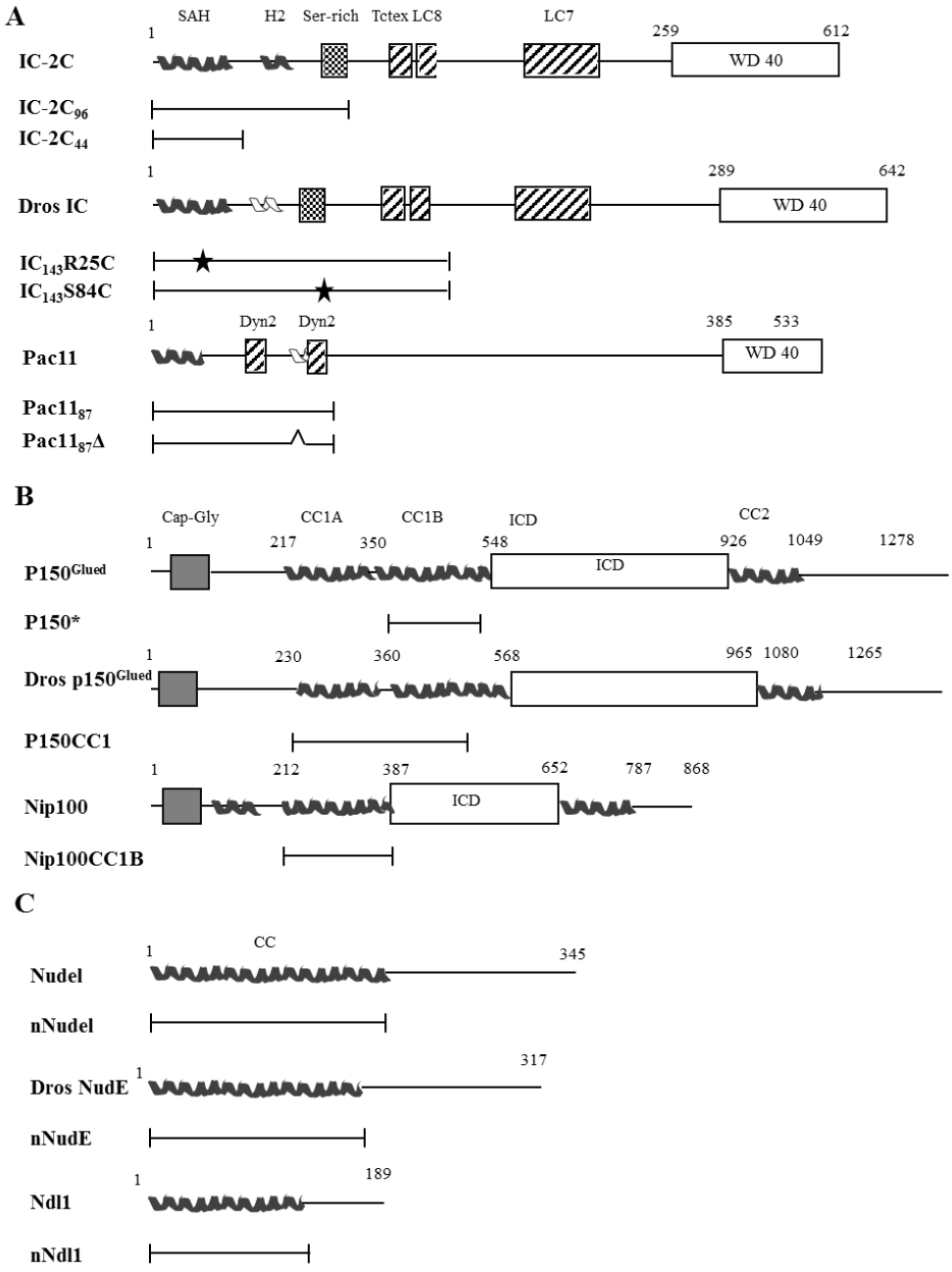
Results

Constructs Design. The amino-terminal portion of dynein intermediate chain IC contains a single α helix (SAH) domain, which is a common binding domain for both $p150^{Glued}$ and NudE, and a second helix (H2) with different NMR-detected helical propensities (Fig. 5.1, black for folded helix, white for transient helix) (Jie et al., 2015; McKenney et al., 2011; Morgan et al., 2011; Nyarko et al., 2012) . It also contains light chains Tctex and LC8 recognition sites. IC from three eukaryotes (rat, Drosophila, and yeast) studies here are shown in Fig. 5.1A. Constructs of rat IC isoform 2C 1-96, 1-44, Drosophila IC 1-143 and Pac11 1-87 were designed to include either only the SAH domain or the SAH, H2 and light chain binding sites. The two Drosophila IC_{143} cysteine mutants were designed to be cross-linked within (R25C) or outside (S84C) the common binding domain of $p150^{Glued}$ and NudE/Nudel. Crosslinking at R25 will create a dimer that will test whether the SAH domain favor formation of a homodimeric coiled coil in the complex. Crosslinking at S84 will mimic the effect of dimeric light chain binding on the interaction with IC-p150 CC1 and IC-nNudE. Corresponding $p150^{Glued}$ and NudE/Nudel constructs from the three eukaryotes were designed to contain the coiled-coil regions identified to be responsible for binding IC (McKenney et al., 2011; Nyarko et al., 2012).

Secondary Structures and Stability of $p150^{Glued}$- and NudE/Nudel- IC

Figure 5.1. Dynein intermediate chain, dynactin p150^{Glued} and NudE/Nudel from two higher eukaryotes and yeast. (A) Sequence-based domains in IC from rat (isoform IC-2C), Drosophila (Dros IC) and yeast (IC homolog Pac11) show at the N-terminus a single α helix (SAH) (Jie et al., 2015). NMR data (Jie et al., 2015; Morgan et al., 2011) also indicate in each IC the presence of a second helix (H2), which may be either nascent (white helix) or strong (black helix). The C-terminal sequence is predicted to be a WD repeat (WD40) domain (empty box). Mammalian IC and Drosophila IC have binding sites for three light chains Tctex, LC8 and LC7 (diagonal stripes), and contain a Ser-rich region (cross-hatched) preceding the light chain binding sites. In yeast, Pac11 has binding sites for two copies of Dyn2 (LC8 homolog) and no Ser-rich region. Below each full-length IC diagram are the constructs used in this study, which are rat IC-2C$_{96}$ and IC-2C$_{44}$, Dros IC$_{143}$ dimer cross-linked by a single point mutation at either residue R25 or S84 to cysteine, Pac11$_{87}$ and Pac11$_{87}$Δ66-73, of which residues 66-73 are deleted. (B) Sequence domains in p150^{Glued} from human, Drosophila (Dros p150^{Glued}) and yeast (p150^{Glued} homolog Nip100) show a Cap-Gly domain (grey box) at the N-terminus, a coiled-coil domain which is further divided into CC1A and CC1B, an intercoiled domain (ICD, empty box) (Urnavicius et al., 2015), and a second coiled-coil domain CC2. Below each full-length p150^{Glued} diagram are the constructs used in this study, which are human p150^{Glued} 382-531 (p150*), Dros p150^{Glued} 221-509 (p150CC1), and Nip100 214-389 (Nip100 CC1B). (C) Primary sequence domains in NudE/Nudel from human (Nudel), Drosophila (Dros NudE) and yeast (NudE homolog Ndl1) show a coiled-coil domain domain (CC) at the N-terminus, and an unstructured C-terminal domain. Below each full-length NudE/Nudel diagram are the constructs used in this study, which are human Nudel 1-189 (nNudel), Dros NudE 1-174 (nNudE), and Ndl1 1-128.

Figure 5.1. (*Continued*)



complexes. Mammalian IC-2C$_{44}$ and nNudel both show typical CD-detected helical structure indicated by negative ellipticity at 208 and 222 nm (Fig. 5.2). In addition, nNudel shows a [θ_{222}]/[θ_{208}] of greater than 1 (Fig. 5.2B) indicative of supercoiling as expected for a coiled-coil dimer. Contrary to prediction, p150* in low salt buffer shows little helical structure as indicated by ellipticity close to zero at 222 nm, and a shift of the 208 nm signal towards 200 nm (Fig. 5.2A), suggesting that p150* is primarily disordered and not a coiled-coil as predicted. Interestingly a large difference in structure is observed for the IC$_{44}$-p150* complex suggesting folding of p150* upon binding to IC as indicated by a CD spectrum for the complex with a [θ_{222}]/[θ_{208}] higher than 1.

IC$_{143}$ Cysteine Mutants Interactions Monitored by ITC. IC$_{143}$R25C dimer binds p150CC1 with a stoichiometry of 1:1, and apparent K_d values of 27.5 μM (Fig. 5.3A), considerably weaker than that of the WT monomeric IC$_{143}$ (3.5 μM) (Morgan et al., 2011). IC$_{143}$R25C binds nNudE with a stoichiometry of 0.5:1 and apparent K_d values of 3.9 μM (Fig. 5.3B). IC$_{143}$S84C binds p150CC1 (Fig. 5.3C) in a two-step manner and with a 50-fold enhanced affinity of 0.07μM for the first step. IC$_{143}$S84C binds nNudE (Fig. 5.3D) with a stoichiometry of 1:1, and with apparent K_d values of and 5.1 μM. Associated enthalpic and entropic contributions are reported in Table 5.1.

Pac11$_{87}$-nNdl1 Interactions Monitored by ITC and NMR. Pac11$_{87}$ and Pac11$_{87}$Δ66-73 bind to nNdl1 (Fig. 5.4A, B) with the same stoichiometry of 1:1, and similar apparent K_d values of 2.3 and 2.5 μM, respectively. In the presence of Dyn2 (LC8 ortholog in yeast), the binding affinity of Pac11$_{87}$ and nNdl1 increased to 0.34 μM (Fig. 5.4C and Table 5.1). Dyn2 presence also changes Pac11$_{87}$-Nip100 CC1B binding to a two-step interaction (Fig. 5.4D). NMR spectra of ^{15}N labeled Pac11 1-87, taken before and after addition of unlabeled nNdl1 (Fig. 5.5A), show segments of contiguous Pac11 residues for which NH peaks disappear. In the nNdl1titration (Fig. 5.5B), peaks that markedly decrease in intensity with added nNdl1 are for residues 1-27 in the N-terminus SAH domain. Consistent with the ITC data, Ndl1 binding on Pac11 only requires the SAH domain, whereas Nip100 binding causes attenuation of peaks in SAH domain, the linker between SAH and first Dyn2 binding site, and H2 (Fig. 5.5B, red).

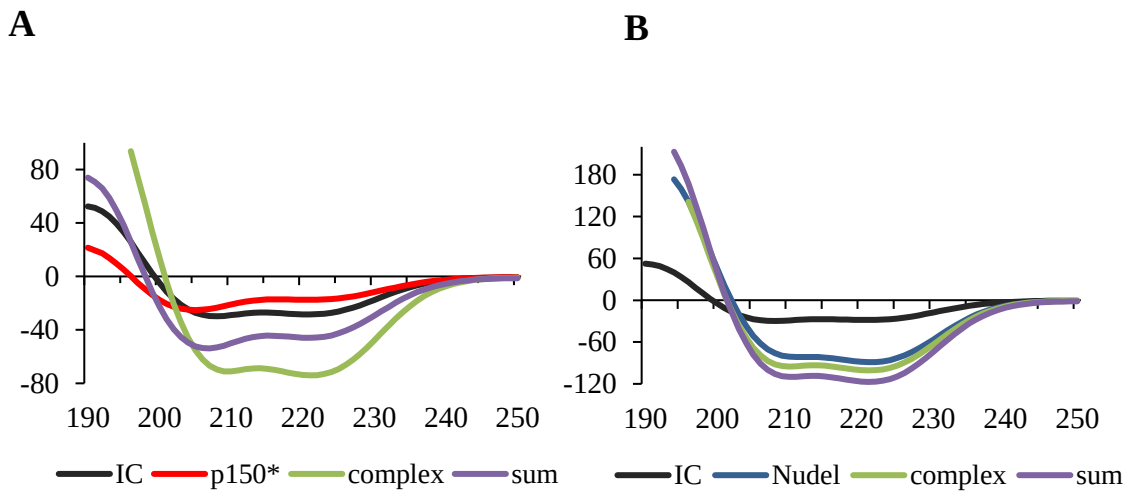


Figure 5.2. CD-detected secondary structure and stability of p150^{Glued} and NudE/Nudel from human. (A) CD spectra overlay of IC_{44} (black), p150* (red), complex of IC_{44}-p150* (green), and the sum of IC_{44} and p150* (purple) are shown. (B) CD spectra overlay of IC_{44} (black), Nudel (blue), complex of IC_{44}-Nudel (green), and the sum of IC_{44} and Nudel (purple) are shown. All spectra were collected at 5 °C.

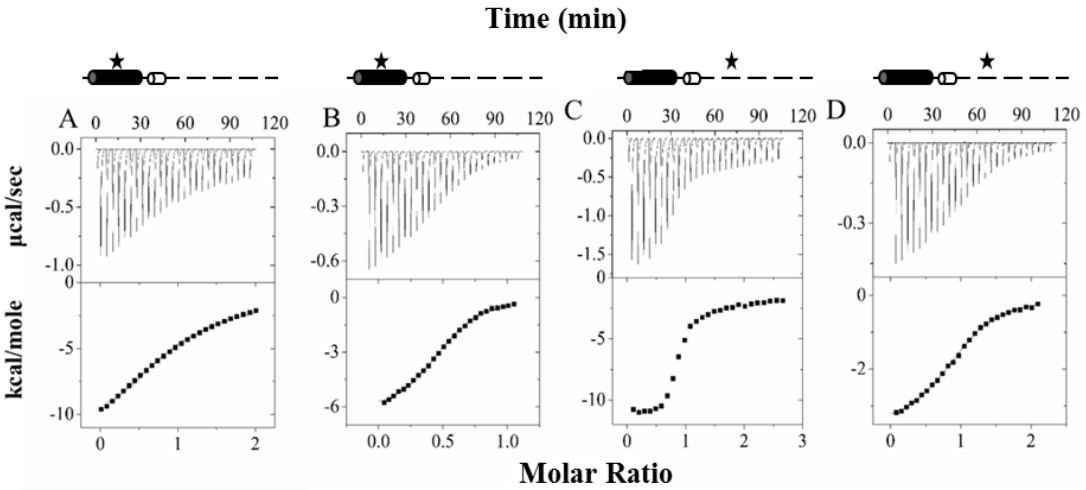


Figure 5.3. Thermodynamic measurements of interactions of IC_{143} cysteine mutants with p150CC1 and nNudE. Representative ITC experiments are shown as thermograms (top) and binding isotherms (bottom) of IC_{143}R25C titrated into p150CC1 (A), nNudE (B), and IC_{143}S84C titrated into p150CC1 (C) and nNudE (D) are shown. Data were collected at 25 °C in 50 mM sodium phosphate, 50 mM sodium chloride, 0.5 mM sodium azide, pH 7.2, and fit to a single-site or a two-site binding model using Origin 7.0 (Microcal) to compute the parameters listed in Table 5.1 except for (C). Shown above each thermogram is the domain diagram of the IC_{143} construct with the star indicating the position of the crosslinking cysteine.

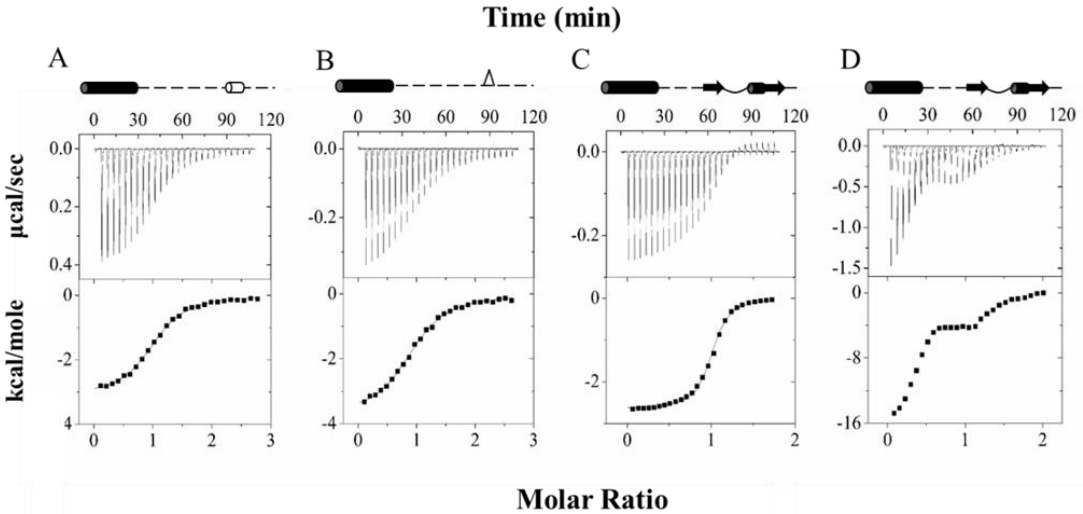


Figure 5.4. Thermodynamic measurements of interactions of Pac11$_{87}$ and Pac11$_{87}$Δ66-73 with nNdl1 and Nip100CC1B. Representative ITC experiments are shown as thermograms (top) and binding isotherms (bottom) of Pac11$_{87}$ (A), Pac11$_{87}$Δ66-73 (B), Pac11$_{87}$ prebound to Dyn2 (C) titrated into nNdl1, and Pac11$_{87}$ prebound to Dyn2 (D) titrated into Nip100CC1B. Data were collected at 25 °C in 50 mM sodium phosphate, 50 mM sodium chloride, 0.5 mM sodium azide, pH 7.2, and fit to a single-site binding model using Origin 7.0 (Microcal) to compute the parameters listed in Table 5.1 except for (D). Shown above each thermogram is the domain diagram of the Pac11$_{87}$ construct (SAH: black cylinder; H2: white cylinder; β strands at Dyn2 binding sites and ordered linker indicated by solid line are representing the Pac11$_{87}$ – Dyn2 complex).

Table 5.1. Thermodynamics parameters for the IC-p150 and IC-NudE interactions recorded at 25 °C

Titrant	Protein	n	K_d	ΔG°	ΔH°	TΔS°
			uM		*kcal/mol*	
IC1–143 R25C[d]	p150CC1	1.0±0.0	28.4±1	-6.30±0.02	-15.9±0.3	-9.6±0.1
IC1–143 R25C[d]	nNudE	0.50±0.04	3.7±1	-7.4±0.2	-6.2±0.5	1.2±0.4
IC1–143 S84C[d]	p150CC1	0.39/0.36	0.07/5.1	-9.9/-7.2	-9.4/-1.8	0.5/5.4
IC1–143 S84C[d]	nNudE	1.0±0.0	5.9±1	-7.2±0.1	-3.70±0.04	3.5±0.1
Pac11 1-87	Nip100	0.97±0.01	5.6±0.1	-7.10±0.02	-8.0±0.5	-0.8±0.5
Pac11 1-87	Ndl1 1-128	1.00±0.02	2.3±0.1	-7.70±0.02	-3.10±0.01	4.60±0.03
Pac11 1-87 Δ66-73	Ndl1 1-128	0.95±0.01	2.5±0.2	-7.60±0.03	-3.70±0.03	3.9±0.1
Pac11 1-87 + Dyn2	Ndl1 1-128	1.00±0.01	0.34±0.02	-8.80±0.04	-2.6±0.1	6.3±0.1

To examine the effect of Ndl1 binding, NMR dynamics experiments were carried out for the Pac11-Ndl1 complex. Relaxation parameters for the remaining residues 28-87 (Fig. 5.5C, blue) are essentially the same as the free Pac11 (Fig. 5.5C, black) and the slightly lower T_2 values for all the remaining residues (Fig. 5.5C, middle) are due to increased protein size thus faster R_2 relaxation. In contrast, Nip100 binding causes increased flexibility in L2 and the C-terminus as evidenced by negative NOE values and increased T_2 values for these regions (Jie et al., 2015).

Interaction Mechanisms of IC-2C with p150 and nNudel*. To examine the residue specificity of IC required for p150* and nNudel, ^{1}H-^{15}N TROSY HSQC and HSQC CLEANEX-PM spectra were collected on the two complexes. In both HSQC spectra of p150*-bound IC_{44} and nNudel-bound IC_{44}, almost all signals were significantly attenuated. When comparing the two HSQC spectra, peaks 3, 5, 32, 38, 39 (Fig. 5.6, red) are still present with low intensity in the nNudel-complex but not in the p150*-complex, indicating the participation of these residues in p150* binding but not nNudel. In the CLEANEX-PM spectra of the two complexes, among the remaining signals in the HSQC spectrum of each complex, residues 40 and 41 are missing in the p150*-complex and residues 5 and 40 are missing in the nNudel-complex, indicating that these residues which are not participating in the binding directly still become more protected in the individual complex.

Discussion

Folding Coupled to Binding as Hallmark of p150^{Glued} Binding. For the interactions of IC and p150^{Glued}, we use a domain of p150^{Glued} (p150*) that is predicted to form a dimeric coiled-coil and corresponds to the smallest domain required for binding (McKenney et al., 2011). A part of this domain has been crystallized for the chicken species and shows indeed a coiled-coil (PDB code: 4RFX). Similarly for IC, we use a domain (IC_{44}) that forms a single alpha helix (SAH) identified by NMR to contain the residues specifically involved in binding and shown by ITC to retain similar binding affinity as the longer construct (Jie et al., 2015) (Jie et al., submitted).

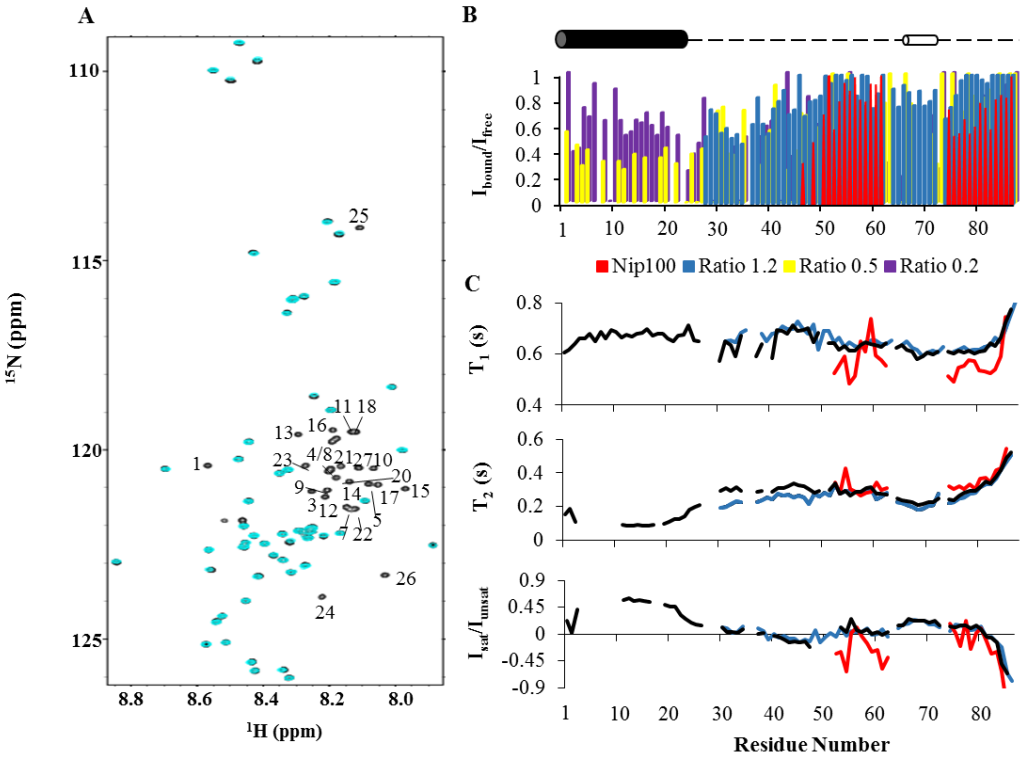


Figure 5.5. NMR titrations and dynamics of Pac11$_{87}$ with nNdl1 and comparison to those of Nip100CC1B. (A) Superposition of 900 MHz ^{1}H-^{15}N HSQC spectra of Pac11$_{87}$ in the apo form (black) and with added nNdl1 at a Pac11:nNdl1 ratio of 1:1.2 (cyan). Peaks that broaden beyond detection in the Ndl1 1-128 bound Pac11 are labeled. (B) Changes in relative intensity of Pac11$_{87}$ NH peaks upon addition of nNdl1 at 700 MHz are shown as a bar graph of residue number *versus* relative peak intensity (I_{bound}/I_{free}). Peak intensities in the presence of nNdl1 are normalized to the reference peak in spectra of free Pac11. Same peak intensity data for Nip100CC1B-bound Pac11 are shown in red with molar ratio Pac11: Nip100CC1B of 1:0.2. (C) Longitudinal ^{15}N relaxation (T_1), transverse relaxation (T_2) and steady-state heteronuclear NOE values (I_{sat}/I_{unsat}) are plotted in for apo Pac11$_{87}$ (black), Nip100CC1B-bound Pac11$_{87}$ with Pac11: Nip100 molar ratio of 1:1.2 (red), and nNdl1-bound Pac11$_{87}$ with Pac11: nNdl1 molar ratio of 1:1.2 (cyan). Dynamics parameters of residues whose peaks disappear in bound forms are left blank in the graphs. The diagram (top) depict apo Pac11$_{87}$ being primarily disordered (dotted lines) with secondary structure the same as Fig. 5.1A.

Figure 5.6. Different interacting mechanisms of the IC-p150 and IC-Nudel complexes. (A) A CS-Rosetta of IC_{44} showing residues that are involved in p150 binding but not Nudel (red), residues that are not involved in either p150 or Nudel binding but become more ordered in the CLEANEX-PM experiments only when p150 is bound (pink), and residues that are not involved in either p150 or Nudel binding but become more ordered in the CLEANEX-PM experiments only when Nudel is bound (blue circles).

The choice of a construct is simpler with NudE/Nudel; sequence prediction shows two independently folding domains, an N-terminal domain that is predicted to form a coiled coil and binds IC and a disordered C-terminal domain that does not bind IC in Drosophila species (Nyarko et al., 2012). A crystal structure for the N-terminal domain is available for murine Nudel (Derewenda et al., 2007). The full N-terminal predicted coiled-coil domain for both human and yeast was included in the studies here.

Comparison of the IC interactions with $p150^{Glued}$ and NudE/Nudel shows that the binding affinity is ~10-fold weaker for $p150^{Glued}$ in the human species (12 μM with $p150^{Glued}$ to 1 μM with Nudel) (Jie et al., 2015) (Jie et al., submitted), but only ~10-fold weaker in yeast (5.6 μM to 2.3 μM). Most striking about this comparison is that even when the affinities are comparable as in yeast, the changes in entropy and enthalpy are significantly different. In human, TΔS is -18 kcal/mol in one and 2 in the other (Table 4.1), while in yeast, TΔS is -0.8 kcal/mol in one and 4.6 in the other (Table 5.1) (Jie et al., 2015). The large entropic penalty associated with binding $p150^{Glued}$ cannot be attributed only to increase in ordered structure of IC because in human, when IC_{44} which already forms an ordered SAH domain (from NMR studies) still binds with a high entropic penalty and high favorable enthalpy (TΔS of -9 kcal/mol in one and 4 in the other). We attribute the large entropic penalty observed primarily with binding $p150^{Glued}$ to an increase in ordered structure of $p150^{Glued}$ in the complex and suggests a folding/coiled-coil stabilization in $p150^{Glued}$ coupled to its binding.

This suggestion is verified by the stark difference observed in CD-detected structure between $p150^{Glued}$ and Nudel. While nNudel shows a stable dimeric coiled-coil, p150* shows some helical structure only at low temperature and no evidence of coiled-coil formation at the low concentrations studied by CD. When IC_{44} is bound, a significant formation of a helical coiled-coil structure is observed and a sufficiently stable complex that unfolds cooperatively above 35˚C. The increase in stability between Nudel and IC/Nudel complex reflects the high binding affinity to IC, while the cooperative unfolding of the IC/$p150^{Glued}$ complex reflects the folding/dimerization of $p150^{Glued}$ upon binding.

Long Range Effects on IC Structure is Hallmark of p150^{Glued} Binding. In the previous paragraph, we described evidence that the structure of p150^{Glued} is changed upon its binding to IC suggesting that the plasticity in p150^{Glued} is required for its regulation of IC binding. Here we show that the effect of binding on the IC structure is in turn only pronounced upon interactions with p150^{Glued}. NMR studies identify that the SAH domain is the recognition site for both p150^{Glued} and NudE/Nudel (Jie et al., 2015; Morgan et al., 2011; Nyarko et al., 2012), but both crosslinking and ITC experiments show that segments completely distant from SAH contribute to binding only with p150^{Glued}. In Drosophila IC, crosslinking at residue 84 in the linker distant from binding, significantly enhances p150^{Glued} binding but has no effect on binding to NudE. Removal of helix 2 (H2) in Drosophila IC and Pac11 significantly reduces binding with p150^{Glued}, but has no effect on binding to NudE/Ndl1. NMR titration experiments show peak attenuation in H2 and other distant residues only upon interactions with p150^{Glued} for both Drosophila and yeast proteins. NMR dynamics experiments performed on yeast proteins show remarkable and unexpected increase in flexibility in the linkers distant from binding only upon binding to p150^{Glued}/Nip100 (Jie et al., 2015), while no change is observed with NudE/Ndl1. Most importantly, experiments that mimic phosphorylation at Ser84 in mammalian IC show significant reduction in binding of p150^{Glued}, but no change in the binding of Nudel (Jie et al., submitted). These differences suggest a model where binding of p150^{Glued} requires the rest of the protein to regulate access to the recognition SAH domain.

P150^{Glued} and NudE Bind to Opposite Surfaces of the IC SAH Domain. For one protein to require access to the binding surface while the other does not, suggests that while they share the same recognition domain, they bind to somewhat opposite surfaces on the same SAH region. Experiments to determine the structures of IC/p150^{Glued} and IC/NudE by X-ray crystallography have not been successful. NMR titration experiments for detecting ^{1}H-^{15}N chemical shifts perturbations have also been unsuccessful because of peak attenuations across the whole SAH domain with both p150^{Glued} and NudE in

any species studied. We summarize here experiments that suggest opposite surfaces for binding on the IC SAH domain.

In human proteins, Cleanex experiments on bound $p150^{Glued}$ and bound Nudel compared to the free IC identify boundaries of the surface that is engaged in $p150^{Glued}$ to be larger than that with Nudel (3-39 for $p150^{Glued}$ and 6-37 for Nudel) even though $p150^{Glued}$ binds with considerably weaker affinity.

Using Drosophila proteins, crosslinking in the SAH domain not only shows weaker binding indicating that the monomeric SAH domain prefers to bind as a monomer and does not form a homodimeric coiled coil as viewed in the literature (Siglin et al., 2013), but more importantly, the stoichiometry of binding changes with NudE indicating that one NudE dimer binds two dimers of IC, instead of one as in $p150^{Glued}$. This suggests that crosslinking covers a site on IC that is not accessible to NudE but still accessible to $p150^{Glued}$.

Using yeast proteins, NMR titration experiments on bound Nip100 and bound Ndl1 compared to the free Pac11 also identify boundaries of the interaction surface on IC for $p150^{Glued}$ to be larger than that with Ndl1. In Nip100-bound Pac11, SAH residues 1-24 disappear as well as linker residues 25-45. In Ndl1-bound Pac11, only residues 1-27 disappear.

Disorder in Helix 2 and Binding Regulation. Our NMR-based structural features of IC from mammals, Drosophila, and yeast show interesting similarities, but striking differences in the structure and interactions of a short helix in IC, H2. IC from all 3 organisms contains a SAH domain that is the binding site for both $p150^{Glued}$ and NudE/Nudel. They also contain a shorter second helix, H2, the features of which are notably different. Among the three organisms H2 has varying structure and distance from SAH (Fig. 5.1). In both mammals and Drosophila, H2 is separated from SAH by a short 9-13-amino acid linker, while in yeast, H2 is much shorter (8 aa, compared to 15 aa in mammalian IC) but distant from SAH by a 43-amino acid linker. Only mammalian IC has a well-formed H2, while in both Drosophila and yeast, H2 is

disordered and has only a small propensity to form a helix (Jie et al., 2015; Morgan et al., 2011).

The importance of H2 for binding p150^{Glued} also varies among species. In mammalian IC, binding of a construct containing only SAH is only slightly weaker when compared to a construct containing SAH and H2. In contrast, in either Drosophila or yeast IC, IC constructs that lack H2 have significantly diminished binding to p150^{Glued} (Jie et al., 2015; Morgan et al., 2011). All the binding experiments were performed on proteins of the same species (i.e. mammalian IC with mammalian p150^{Glued} and NudE/Nudel, etc.).

These data raise interesting questions as to the function of H2 in Drosophila and yeast IC, even when they are not at the recognition site as identified by NMR, and when both only form short transient helices. The fully formed H2 in mammalian IC does not contribute to binding affinity suggesting that there may be a correlation between intrinsic disorder in H2 and binding affinity such that when H2 is ordered it does not contribute to the binding affinity. While this correlation has not been observed in intrinsically disordered complexes, it is tempting to speculate that the disorder and flexibility of H2 may be necessary to allow access to p150^{Glued}, while in mammalian IC phosphorylation offers another layer of control of this access (see below).

Evolution of Regulation by Dynactin among Mammals, Drosophila *and Yeast.* In yeast Pac11, the effect of H2 on binding is overshadowed by the presence of light chains. H2 is sandwiched between the two recognition motifs for Dyn2 (LC8 ortholog in yeast), and therefore when Dyn2 is present in the assembled dynein, as we expect it to be considering its very high affinity, whether H2 is present or not, Pac11 binds to Nip100 with higher affinity. Therefore, it appears that in the assembled complex, there is no apparent method to differentiate between p150^{Glued} binding and NudE.

In Drosophila IC, the light chain binding regions are considerably C-terminal of H2 (Fig. 5.1) and therefore we speculate that the presence of H2 does contribute to regulation of p150^{Glued} binding in the assembled complex. An additional layer of control is the presence of light chains binding as modeled by S84C crosslinking, which shows

differential effect on binding to p150^{Glued} where its affinity is enhanced but with no change to its binding to NudE.

In mammals, a higher level of control has evolved by insertion of a serine-rich region distant from the SAH domain that can undergo phosphorylation significantly reducing p150^{Glued} binding in one isoform IC-2C but without affecting Nudel binding (Jie et al., submitted). In Drosophila, while there is a serine-rich sequence in a similar region, there is no evidence of its phosphorylation *in vivo*. In yeast, there is no serine-rich region, and therefore it appears that phosphorylation is only a mammalian development that replaces the involvement of H2. In mammalian IC, H2 is well folded and its presence does not contribute appreciably to binding.

Selection of binding partners is a highly pertinent question, considering the multiplicity of regulators of dynein that bind IC. Significantly, the variation of p150^{Glued} binding to IC among organisms indicates different regulatory mechanisms have evolved, and these studies using proteins from yeast, Drosophila and mammalian origin, are the first step in identifying the molecular basis of the strategic role for protein disorder in regulation of dynein function.

Experimental Procedures

Constructs Design. The *D. melanogaster* IC 1–143 was cloned into pET SUMO (Invitrogen) vector. Site-directed mutagenesis of Arg-25 and Ser-84 to generate IC_{143} single cysteine mutants were performed using the Stratagene QuikChange ™ kit (Agilent Technologies) following the manufacturer's protocol (ref). Ndl1 1-128 was generated by PCR and was subsequently cloned into a pET15b vector with a C-terminus 6Xhistidine tag. All sequences were verified by automated sequencing. Recombinant Drosophila p150^{Glued} CC1, nNudE, rat IC isoform 2C 1-96, 1-44, human p150 382-531 (P150*), human Nudel 1-189 (nNudel), Pac11 1-87, Pac11 1-87Δ66-73, Dyn2, and Nip100 CC1B were described previously (Jie et al., 2015; Morgan et al., 2011; Nyarko et al., 2012). Recombinant vectors were transformed into BL21 (DE3) or Rosetta (DE3) cell lines for protein expression.

Protein Expression and Purification. The transformed cell lines were grown in LB media at 37 °C to an optical density (A600) of 0.6-0.8 followed by protein induction for 4-6 h with 0.1-0.4 mM IPTG. Cells were harvested, lysed, and centrifuged to remove all cell debris. The soluble fractions were purified using Qiagen's Ni-NTA affinity chromatography protocol (Qiagen Valencia, CA) or Talon's metal affinity resin (Clontech).

The region encoding the 6Xhistidine and SUMO sequences was removed by incubating the proteins with SUMO protease (Cornell University). The IC cysteine mutants were added diamide (SIGMA) to enhance dimerization while concentrating. Other proteins were purified as previously described (Jie et al., 2015; Morgan et al., 2011; Nyarko et al., 2012). All proteins were further purified on a Superdex^{TM} 75 (GE healthcare) size exclusion chromatography column, resulting in purity > 95 %.

Proteins were dialyzed against appropriate buffers, and their concentrations determined from absorbances at 280 nm and computed (expasy.org/protparam) molar extinction coefficient values (IC_{143}R25C, IC_{143}S84C: 2980 M^{-1} cm^{-1}; Ndl1 1-128: 1490 M^{-1} cm^{-1}). Reported concentrations are those for the monomeric proteins. The computed protein molar extinction coefficient takes into account the polyhistidine sequence from the expression vector. The proteins were subsequently stored at 4 ºC and used within a week.

Isothermal Titration Calorimetry. ITC experiments using Microcal's VP-ITC microcalorimeter (North Hampton, MA), were performed at 25 °C in buffer composed of 50 mM sodium phosphate, 50 mM sodium chloride, and 1 mM sodium azide, pH 7.2-7.5. A typical experiment involved an initial 2 μL injection, followed by 26 - 10 μL injections of titrants to proteins in the concentration range of 18-60 μM, accompanied by stirring at a constant rate of 351 rpm. Added titrants were in the concentration range of 0.3 – 0.4 mM. Protein samples and buffer were degassed prior to data acquisition. Data were processed using Origin 7.0 and fit to a single-site binding model or a two single-site binding model. Recorded data are the average of two or more independent

experiments, with the uncertainty reported as the difference between the experimental value and the average.

CD Experiments. CD spectra were recorded on a JASCO 720 spectropolarimeter using a 1-mm cell. Proteins were prepared in 10 mM sodium phosphate or 1X PBS, pH 6. Data on the native proteins were collected at 25 °C and a protein concentration of 10 μM. Temperature unfolding data by monitoring the ellipticity at 222 nm was acquired in the temperature range of 5–70 °C. A temperature-regulated water bath connected to the sample chamber allowed accurate monitoring of the sample temperature.

NMR Experiments. ^{15}N labeled $Pac11_{87}$ and IC-$2C_{44}$ were prepared using published protocols (Jie et al., 2015). Protein samples were prepared in buffer composed of 10 mM sodium phosphate at pH 7.0 with 50 mM NaCl, 1 mM sodium azide, 8 % $^{2}H_2O$, a protease inhibitor mixture (Roche Applied Science), and 2,2-dimethylsilapentene-5-sulfonic acid for ^{1}H chemical shifts referencing. NMR titrations of ^{15}N-labeled Pac11 1-87 with unlabeled Ndl1 1-128 were acquired on a Bruker Avance 700 MHz spectrometer at 25 °C. A series of BEST HSQC spectra of ^{15}N-labeled Pac11 1-87 were collected with Ndl1 1-128 at final molar ratios (Pac11: Ndl1) of 1:0.2, 1:0.5, and 1:1.2. HSQC spectra and dynamics experiments of ^{15}N labeled Pac11 1-87 were collected for Ndl1 1-128 bound at molar ratio (Pac11: Ndl1) of 1:1.2 at 15 °C on 900 MHz and 600 MHz Bruker Avance spectrometers, respectively (Jie et al., 2015). Spectra were collected for apo-, p150*-bound, nNudel-bound IC_{44} at 15 °C using the CLEANEX-PM-FHSQC pulse sequence with a 20-ms mixing time on a 700 MHz Bruker Avance spectrometer.

NMR Data Analysis. All spectra were processed with TopSpin (Bruker) and analyzed using Sparky (Goddard and Kneller). For titration experiments, peak intensities were measured as peak heights. To account for differences in concentration across the titration series, a normalization factor was determined from the signal to noise ratio of residue 86 which is not affected by Ndl1 1-128 binding. Changes in peak intensities I_{Bound}/I_{Free} were calculated as the ratio of signal intensity in spectra of the complex and free Pac11 1-87.

Acknowledgements

The authors wish to thank Dr. Helena Kovacs, Andrew Popchock and Yujuan Song for the help in data collection. This work is supported by National Science Foundation grant MCB 0818896 to EB. We acknowledge the support of the protein core facility in the OSU Environmental Health Sciences Center (NIH/NIEHS 00210), and access to the Research Infrastructure activity in the 7th Framework Programme of the EC (Project number: 261863, Bio-NMR) (Frankfurt, Germany).

Chapter 6

Summary and Future Work

The dynein intermediate chain (IC) subunit is central to both dynein assembly and dynein regulation. The N-terminal domain of IC that contains the binding sites for dynein light chain subunits and two dynein regulators ($p150^{Glued}$ and NudE/Nudel) is intrinsically disordered, an transitions from a monomeric protein to a duplex scaffold when binds to its dimeric partners. $P150^{Glued}$ and NudE/Nudel both participate in numerous dynein functions and are able to regulate dynein motor activity and target dynein to cargos and subcellular locations via their interactions with the disordered N-IC. The studies in this dissertation were aimed at investigating the interactions of N-IC with $p150^{Glued}$ and NudE/Nudel in different species to ultimately understand the mechanism and evolution of dynein regulation. This chapter is the summary of the results and conclusions in chapters 3, 4 and 5. In addition, some questions remaining in the field and proposed future work to potentially answer those questions are presented.

Summary

A Single α Helix (SAH) and a Second Helix Are Conserved in the Disordered Region of Dynein N-IC. We have experimentally determined the secondary structures using NMR spectroscopy of yeast, Drosophila and mammalian dynein N-IC and shown that they all contain a SAH domain (Jie et al., 2015; Morgan et al., 2011) that is an α helix enriched in charged residues and that does not meet the requirements for the formation of a coiled coil in contrast to speculation (Vaughan and Vallee, 1995b). They also all contain a second shorter α helix (H2) comparing to the SAH that is separated by a disordered linker with different lengths (~14 residues in Drosophila and mammalian IC and 42 residues in yeast IC). There are striking differences in the helicity of H2 among species, namely H2 in mammalian IC is as strong as SAH whereas H2 in yeast and Drosophila IC are weak, nascent helices.

The SAH Domain in Dynein N-IC is Necessary for Binding both $p150^{Glued}$ and NudE/Nudel but Only $p150^{Glued}$ Requires Long-range Conformational Coordination for Optimal Binding. We have shown using NMR spectroscopy and ITC that the binding site for NudE/Nudel on dynein N-IC is restricted to the SAH domain in all three species.

In contrast, H2 is required for optimal p150^{Glued} binding as evidenced by the reduced affinity determined by ITC in all three species. In yeast, Nip100 (p150^{Glued}) binds to N-Pac11 (IC) through direct interactions with SAH and forms a duplex in which H2 in two Pac11 chains self-associate and become more ordered (Jie et al., 2015). In Drosophila, SAH is the major site of p150^{Glued} interaction and H2 contributes to binding possibly through conformational changes as evidenced by different peak attenuation patterns in NMR titration experiments (Morgan et al., 2011). In mammals, removal of H2 in N-IC slightly decreases the binding affinity to p150^{Glued}, which is quite different from the dramatic weaker binding of yeast and Drosophila N-IC without H2 to p150^{Glued}. H2 contributes to p150^{Glued} binding as evidenced by peak attenuation in NMR titration experiments at different field strengths.

Light Chain Binding on N-IC Regulates p150^{Glued} and NudE/Nudel Interactions. In yeast, the dynein light chain Dyn2 (LC8) binding on N-Pac11 forms an aligned bivalent scaffold and restricts the flexibility of H2 residues. This conformation enhances both Nip100 and Ndl1 (NudE) binding by increasing the effective concentration of Pac11 dimer. The Dyn2 recognition motif is 20 residues from SAH, the binding site for Nip100 and Ndl1. In Drosophila, crosslinking at residue 84 by point mutation to cysteine on N-IC forms an IC dimer which mimics the light chain-binding conformation. The crosslinking site is 47 residues from SAH and it enhances only p150^{Glued} binding but not NudE.

Phosphorylation on Tissue-specific IC Isoform Regulates p150^{Glued} and NudE/Nudel Interactions with Mammalian IC. The phosphomimetic IC-2C$_{96}$S84D has significantly decreased binding affinity to p150^{Glued} but not NudE/Nudel, and the corresponding phosphomimetic mutant on another IC isoform IC-1A$_{112}$S100D does not affect either p150^{Glued} or NudE/Nudel binding. We have shown that S84D mutation does not alter the secondary structure of IC but shifts the population of free IC towards a more compact structure in which the SAH packs against H2 and electrostatic interactions occur between the C-terminus of SAH and the linker (L1) separating SAH and H2. This conformation is further stabilized by the electrostatic interactions between

the introduced negative charge at residue 84 and the few residues N-terminal to the SAH domain. This compact conformation makes the SAN inaccessible to $p150^{Glued}$ but does not impact NudE/Nudel interacting with IC. The amino acid composition and length of L1 on IC-2C are optimal to achieve this conformation when phosphorylation takes place at residue 84, whereas IC-1A isoform does not meet the requirements thus explaining the isoform-specificity of regulated $p150^{Glued}$ and NudE/Nudel binding to IC by phosphorylation. The observation above is restricted to mammalian dynein IC because there is no serine-rich region that contains potential phosphorylation sites in Pac11 and no tissue-specific IC isoforms that are different in L1 sequence in Drosophila.

Taken together, the results of these studies provide detailed structural insights into the disordered N-terminus of dynein intermediate chain and regulation of dynein activities mediated by the interactions of N-IC with dynein regulators $p150^{Glued}$ and NudE/Nudel in different eukaryotic species, which provides the first evolutionary view of dynein regulation in the field. We have shown for the first time, from yeast to mammals, dynein regulation by dynactin has evolved to gain complexity whereas dynein interactions with NudE/Nudel only require the SAH domain in all three species and not affected by helix 2 or phosphorylation on IC. We have also shown for the first time that phosphorylation distant from the binding site does not change the structure of the binding site but promotes electrostatically-driven long-range tertiary contacts forming a conformation that blocks access of $p150^{Glued}$. This phosphorylation effect involves coupling between phosphorylation and sequence specificity and length of disordered linkers. These studies give novel structural understanding of the interplay of phosphorylation and alternative splicing in intrinsically disordered regions that increases the functional diversity of large complexes. The unprecedented observation that the interaction between the neuron-specific IC isoform 1A in mammals and dynactin $p150^{Glued}$ is not modulated by phosphorylation provides new opportunities for researchers in the dynein field to investigate the physiological relevance.

Future Work

Structure of the IC-Nudel Complex and IC-p150^{Glued} Complex. To gain insights into the structure of IC when in complex with Nudel and p150^{Glued} and how the structures of the two complexes may differ, x-ray crystallography could be used to obtain the 3D structures of the two complexes. Crystal structures of the coiled-coil domains of mammalian Nudel and p150^{Glued} that are responsible for binding N-IC are available (PDB codes 4RFX and 2V71). The SAH domain of N-IC is a good candidate since it is a single helix without disorder. I have made the complex of mammalian IC-2C$_{44}$ in complex with nNudel and got micro-crystals (Appendix 2, Figure A2.1). Further optimization on pH, concentration of PEG and if necessary a new crystal screen should be performed. The crystallization of IC-p150^{Glued} complex is more challenging as the binding is so weak that the complex dissociates during the size exclusion chromatography. We plan to look at the IC and p150^{Glued} homologues in the thermophilic fungus *Chaetomium thermophilum*, which lives at temperatures up to 60 °C. The overall protein stability and binding affinities should be desirable at 25 °C. Sequence alignments show that *Chaetomium* N-IC also contains an N-terminus SAH domain and *Chaetomium* p150^{Glued} contains a CC1B region. Future work would include the cloning, expression and purification of *Chaetomium* IC SAH domain and p150^{Glued} CC1B, determination of their binding affinity, and if it is high enough, growing crystals of the complex so as to analyze the structure.

Incorporation of Phosphoserine into Mammalian IC-2C Position 84. We have demonstrated the phosphorylation effects on IC-p150^{Glued} interaction using the phosphomimetic mutant IC-2C$_{96}$S84D. A better construct would be IC-2C$_{96}$ with phosphoserine incorporated at position 84. With the maturation of methods for genetically incorporating unnatural amino acids, the protocol and the materials (Rogerson et al., 2015) are available. Future work after successfully incorporating phosphoserine include ITC with p150^{Glued} to determine binding constants, CD and NMR experiments to measure secondary structure and thermal stability, and NMR paramagnetic relaxation enhancement experiments to measure long-range structures.

The Structural Basis for the Interaction of N-IC with Huntingtin. Huntingtin (Htt) is a ubiquitously expressed multidomain protein and plays roles in numerous cellular processes including gene expression, endocytosis (Harjes and Wanker, 2003), vesicle trafficking, and mitotic spindle orientation (Godin et al., 2010). Htt is shown to facilitate dynein mediated cargo transport along microtubules and contribute to the recruitment of the dynein-dynactin complex at the cell cortex (Caviston et al., 2011). A direct dynein-Htt interaction has been demonstrated between dynein N-IC and human huntingtin 536-698 via yeast two-hybrid assays (Caviston et al., 2007). No detailed information of this interaction is available. How Htt regulates dynein's function and how it is coordinated with other dynein regulatory proteins remain elusive.

Preliminary protein purification data (Appendix 3, Figure A3.1) show that mammalian Htt 536-698 is soluble but difficult to purify. However, the production of Htt protein could be attempted again using optimized protein sequences. ITC and NMR experiments could be performed to investigate the thermodynamics, dynamics of the interaction and the recognition sites on IC. The potential binding sites of Htt on N-IC are residues in the segment separating H2 and first light chain (Tctex) binding sites. This segment contains the serine-rich region that has the phosphorylation site. Potential research questions are: will light chain binding, $p150^{Glued}$ and NudE/Nudel binding, and IC phosphorylation affect this interaction? What is the structure of IC in complex with all binding partners above?

Htt in Drosophila (DmHtt) has been shown to interact with dynein and vesicles and the function of Htt in axonal transport is conserved (Zala et al., 2013). I have expressed and purified the DmHtt construct and obtained preliminary ITC data with DmIC (Appendix 3, Figure A3.2). The ITC experiment could be done again with higher protein concentrations to obtain thermodynamic parameters. NMR titration experiments could be performed to map Htt recognition sites and effect on IC. With the structural and dynamics information of the IC-Htt complex, a more complete image of dynein N-IC regulation in different species will be available.

Bibliography

Adames, N.R., and Cooper, J.A. (2000). Microtubule interactions with the cell cortex causing nuclear movements in Saccharomyces cerevisiae. J. Cell Biol. *149*, 863–874.

Asai, D.J., and Koonce, M.P. (2001). The dynein heavy chain: structure, mechanics and evolution. Trends Cell Biol. *11*, 196–202.

Baboolal, T.G., Sakamoto, T., Forgacs, E., White, H.D., Jackson, S.M., Takagi, Y., Farrow, R.E., Molloy, J.E., Knight, P.J., Sellers, J.R., et al. (2009). The SAH domain extends the functional length of the myosin lever. Proc. Natl. Acad. Sci. U. S. A. *106*, 22193–22198.

Babu, M.M., van der Lee, R., de Groot, N.S., and Gsponer, J. (2011). Intrinsically disordered proteins: regulation and disease. Curr. Opin. Struct. Biol. *21*, 432–440.

Babu, M.M., Kriwacki, R.W., and Pappu, R.V. (2012). Structural biology. Versatility from protein disorder. Science *337*, 1460–1461.

Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L.E., and Forman-Kay, J.D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature *519*, 106–109.

Bananis, E., Nath, S., Gordon, K., Satir, P., Stockert, R.J., Murray, J.W., and Wolkoff, A.W. (2004). Microtubule-dependent movement of late endocytic vesicles in vitro: requirements for Dynein and Kinesin. Mol. Biol. Cell *15*, 3688–3697.

Barbar, E. (2008). Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. Biochemistry (Mosc.) *47*, 503–508.

Barbar, E. (2012). Native disorder mediates binding of dynein to NudE and dynactin. Biochem. Soc. Trans. *40*, 1009–1013.

Barbar, E., and Nyarko, A. (2014). Polybivalency and disordered proteins in ordering macromolecular assemblies. Semin. Cell Dev. Biol.

Belle, V., Rouger, S., Costanzo, S., Liquière, E., Strancar, J., Guigliarelli, B., Fournel, A., and Longhi, S. (2008). Mapping alpha-helical induced folding within the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by site-directed spin-labeling EPR spectroscopy. Proteins *73*, 973–988.

Benison, G., Nyarko, A., and Barbar, E. (2006). Heteronuclear NMR Identifies a Nascent Helix in Intrinsically Disordered Dynein Intermediate Chain: Implications for Folding and Dimerization. J. Mol. Biol. *362*, 1082–1093.

Benison, G., Karplus, P.A., and Barbar, E. (2007). Structure and dynamics of LC8 complexes with KXTQT-motif peptides: swallow and dynein intermediate chain compete for a common site. J. Mol. Biol. *371*, 457–468.

Beveridge, R., Chappuis, Q., Macphee, C., and Barran, P. (2013). Mass spectrometry methods for intrinsically disordered proteins. The Analyst *138*, 32–42.

Blasier, K.R., Humsi, M.K., Ha, J., Ross, M.W., Smiley, W.R., Inamdar, N.A., Mitchell, D.J., Lo, K.W.-H., and Pfister, K.K. (2014). Live cell imaging reveals differential modifications to cytoplasmic dynein properties by phospho- and dephosphomimic mutations of the intermediate chain 2C S84. J. Neurosci. Res. *92*, 1143–1154.

Bowman, A.B., Patel-King, R.S., Benashski, S.E., McCaffery, J.M., Goldstein, L.S., and King, S.M. (1999). Drosophila roadblock and Chlamydomonas LC7: a conserved family of dynein-associated proteins involved in axonal transport, flagellar motility, and mitosis. J. Cell Biol. *146*, 165–180.

Brown, C.J., Takayama, S., Campen, A.M., Vise, P., Marshall, T.W., Oldfield, C.J., Williams, C.J., and Dunker, A.K. (2002). Evolutionary rate heterogeneity in proteins with long disordered regions. J. Mol. Evol. *55*, 104–110.

Buchan, D.W.A., Minneci, F., Nugent, T.C.O., Bryson, K., and Jones, D.T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. *41*, W349-357.

Buljan, M., Chalancon, G., Eustermann, S., Wagner, G.P., Fuxreiter, M., Bateman, A., and Babu, M.M. (2012). Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. Mol. Cell *46*, 871–883.

Buljan, M., Chalancon, G., Dunker, A.K., Bateman, A., Balaji, S., Fuxreiter, M., and Babu, M.M. (2013). Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. Curr. Opin. Struct. Biol. *23*, 443–450.

Burgess, S.A., Walker, M.L., Sakakibara, H., Knight, P.J., and Oiwa, K. (2003). Dynein structure and power stroke. Nature *421*, 715–718.

Campbell, K.S., Cooper, S., Dessing, M., Yates, S., and Buder, A. (1998). Interaction of p59fyn kinase with the dynein light chain, Tctex-1, and colocalization during cytokinesis. J. Immunol. Baltim. Md 1950 *161*, 1728–1737.

Carter, A.P., Cho, C., Jin, L., and Vale, R.D. (2011). Crystal structure of the dynein motor domain. Science *331*, 1159–1165.

Caviston, J.P., Ross, J.L., Antony, S.M., Tokito, M., and Holzbaur, E.L.F. (2007). Huntingtin facilitates dynein/dynactin-mediated vesicle transport. Proc. Natl. Acad. Sci. *104*, 10045–10050.

Caviston, J.P., Zajac, A.L., Tokito, M., and Holzbaur, E.L.F. (2011). Huntingtin coordinates the dynein-mediated dynamic positioning of endosomes and lysosomes. Mol. Biol. Cell *22*, 478–492.

Chansard, M., Hong, J.-H., Park, Y.-U., Park, S.K., and Nguyen, M.D. (2011). Ndel1, Nudel (Noodle): flexible in the cell? Cytoskelet. Hoboken NJ *68*, 540–554.

Chen, X.-J., Levedakou, E.N., Millen, K.J., Wollmann, R.L., Soliven, B., and Popko, B. (2007). Proprioceptive sensory neuropathy in mice with a mutation in the cytoplasmic dynein heavy chain 1 gene. J. Neurosci. *27*, 14515–14524.

Chevalier-Larsen, E., and Holzbaur, E.L.F. (2006). Axonal transport and neurodegenerative disease. Biochim. Biophys. Acta *1762*, 1094–1108.

Chowdhury, S., Ketcham, S.A., Schroer, T.A., and Lander, G.C. (2015). Structural organization of the dynein-dynactin complex bound to microtubules. Nat. Struct. Mol. Biol. *22*, 345–347.

Clark, S., Nyarko, A., Löhr, F., Karplus, P.A., and Barbar, E. (2016). The Anchored Flexibility Model in LC8 Motif Recognition: Insights from the Chica Complex. Biochemistry (Mosc.) *55*, 199–209.

Clark, S.A., Jespersen, N., Woodward, C., and Barbar, E. (2015). Multivalent IDP assemblies: Unique properties of LC8-associated, IDP duplex scaffolds. FEBS Lett. *589*, 2543–2551.

Coletta, A., Pinney, J.W., Solís, D.Y.W., Marsh, J., Pettifer, S.R., and Attwood, T.K. (2010). Low-complexity regions within protein sequences have position-dependent roles. BMC Syst. Biol. *4*, 43.

Cortese, M.S., Uversky, V.N., and Dunker, A.K. (2008). Intrinsic disorder in scaffold proteins: getting more from less. Prog. Biophys. Mol. Biol. *98*, 85–106.

Criswell, P.S., and Asai, D.J. (1998). Evidence for four cytoplasmic dynein heavy chain isoforms in rat testis. Mol. Biol. Cell *9*, 237–247.

Culver-Hanlon, T.L., Lex, S.A., Stephens, A.D., Quintyne, N.J., and King, S.J. (2006). A microtubule-binding domain in dynactin increases dynein processivity by skating along microtubules. Nat. Cell Biol. *8*, 264–270.

Delano, W.L. (2002). The PyMOL Molecular Graphics System, DeLano Scientific LLC, San Carlos, CA.

Derewenda, U., Tarricone, C., Choi, W.C., Cooper, D.R., Lukasik, S., Perrina, F., Tripathy, A., Kim, M.H., Cafiso, D.S., Musacchio, A., et al. (2007). The structure of the coiled-coil domain of Ndel1 and the basis of its interaction with Lis1, the causal protein of Miller-Dieker lissencephaly. Struct. Lond. Engl. 1993 *15*, 1467–1481.

Dhani, S.U., Mohammad-Panah, R., Ahmed, N., Ackerley, C., Ramjeesingh, M., and Bear, C.E. (2003). Evidence for a functional interaction between the ClC-2 chloride channel and the retrograde motor dynein complex. J. Biol. Chem. *278*, 16262–16270.

Dillman, J.F., and Pfister, K.K. (1994). Differential phosphorylation in vivo of cytoplasmic dynein associated with anterogradely moving organelles. J. Cell Biol. *127*, 1671–1681.

Dosztányi, Z., Csizmók, V., Tompa, P., and Simon, I. (2005). The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. J. Mol. Biol. *347*, 827–839.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., et al. (2001). Intrinsically disordered protein. J. Mol. Graph. Model. *19*, 26–59.

Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., and Obradović, Z. (2002). Intrinsic disorder and protein function. Biochemistry (Mosc.) *41*, 6573–6582.

Dunker, A.K., Bondos, S.E., Huang, F., and Oldfield, C.J. (2015). Intrinsically disordered proteins and multicellular organisms. Semin. Cell Dev. Biol. *37*, 44–55.

Dunsch, A.K., Hammond, D., Lloyd, J., Schermelleh, L., Gruneberg, U., and Barr, F.A. (2012). Dynein light chain 1 and a spindle-associated adaptor promote dynein asymmetry and spindle orientation. J. Cell Biol. *198*, 1039–1054.

Dyson, H.J., and Wright, P.E. (2002). Coupling of folding and binding for unstructured proteins. Curr. Opin. Struct. Biol. *12*, 54–60.

Efimov, V.P., and Morris, N.R. (2000). The LIS1-related NUDF protein of Aspergillus nidulans interacts with the coiled-coil domain of the NUDE/RO11 protein. J. Cell Biol. *150*, 681–688.

Farkasovsky, M., and Küntzel, H. (2001). Cortical Num1p interacts with the dynein intermediate chain Pac11p and cytoplasmic microtubules in budding yeast. J. Cell Biol. *152*, 251–262.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry (Mosc.) *33*, 5984–6003.

Favier, A., and Brutscher, B. (2011). Recovering lost magnetization: polarization enhancement in biomolecular NMR. J. Biomol. NMR *49*, 9–15.

Feng, Y., and Walsh, C.A. (2004). Mitotic spindle regulation by Nde1 controls cerebral cortical size. Neuron *44*, 279–293.

Ferrage, F., Piserchio, A., Cowburn, D., and Ghose, R. (2008). On the measurement of 15N-{1H} nuclear Overhauser effects. J. Magn. Reson. San Diego Calif 1997 *192*, 302–313.

Fu, M., and Holzbaur, E.L.F. (2014). Integrated regulation of motor-driven organelle transport by scaffolding proteins. Trends Cell Biol. *24*, 564–574.

Galea, C.A., Wang, Y., Sivakolundu, S.G., and Kriwacki, R.W. (2008). Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. Biochemistry (Mosc.) *47*, 7598–7609.

Gao, F.J., Hebbar, S., Gao, X.A., Alexander, M., Pandey, J.P., Walla, M.D., Cotham, W.E., King, S.J., and Smith, D.S. (2015). GSK-3β Phosphorylation of Cytoplasmic Dynein Reduces Ndel1 Binding to Intermediate Chains and Alters Dynein Motility. Traffic Cph. Den.

Gauthier-Fisher, A., Lin, D.C., Greeve, M., Kaplan, D.R., Rottapel, R., and Miller, F.D. (2009). Lfc and Tctex-1 regulate the genesis of neurons from cortical precursor cells. Nat. Neurosci. *12*, 735–744.

Gill, S., Schroer, T., Szilak, I., Steuer, E., Sheetz, M., and Cleveland, D. (1991). Dynactin, a Conserved, Ubiquitously Expressed Component of an Activator. J. Cell Biol. *115*, 1639–1650.

Goddard, T.D., and Kneller, D.G. SPARKY3, University of California, San Francisco.

Godin, J.D., Colombo, K., Molina-Calavita, M., Keryer, G., Zala, D., Charrin, B.C., Dietrich, P., Volvert, M.-L., Guillemot, F., Dragatsis, I., et al. (2010). Huntingtin is required for mitotic spindle orientation and mammalian neurogenesis. Neuron *67*, 392–406.

Gsponer, J., and Babu, M.M. (2009). The rules of disorder or why disorder rules. Prog. Biophys. Mol. Biol. *99*, 94–103.

Gsponer, J., and Babu, M.M. (2012). Cellular strategies for regulating functional and nonfunctional protein aggregation. Cell Rep. *2*, 1425–1437.

Gsponer, J., Futschik, M.E., Teichmann, S.A., and Babu, M.M. (2008). Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. Science *322*, 1365–1368.

Ha, J., Lo, K.W.-H., Myers, K.R., Carr, T.M., Humsi, M.K., Rasoul, B.A., Segal, R.A., and Pfister, K.K. (2008). A neuron-specific cytoplasmic dynein isoform preferentially transports TrkB signaling endosomes. J. Cell Biol. *181*, 1027–1039.

Haas, E. (2012). Ensemble FRET methods in studies of intrinsically disordered proteins. Methods Mol. Biol. Clifton NJ *895*, 467–498.

Hafezparast, M. (2003). Mutations in Dynein Link Motor Neuron Degeneration to Defects in Retrograde Transport. Science *300*, 808–812.

Hall, J., Karplus, P.A., and Barbar, E. (2009). Multivalency in the Assembly of Intrinsically Disordered Dynein Intermediate Chain. J. Biol. Chem. *284*, 33115–33121.

Hall, J., Song, Y., Karplus, P.A., and Barbar, E. (2010). The crystal structure of dynein intermediate chain-light chain roadblock complex gives new insights into dynein assembly. J. Biol. Chem. *285*, 22566–22575.

Harada, A., Takei, Y., Kanai, Y., Tanaka, Y., Nonaka, S., and Hirokawa, N. (1998). Golgi vesiculation and lysosome dispersion in cells lacking cytoplasmic dynein. J. Cell Biol. *141*, 51–59.

Harjes, P., and Wanker, E.E. (2003). The hunt for huntingtin function: interaction partners tell many different stories. Trends Biochem. Sci. *28*, 425–433.

Hart, M.J., de los Santos, R., Albert, I.N., Rubinfeld, B., and Polakis, P. (1998). Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. Curr. Biol. CB *8*, 573–581.

Hauer, J.A., Taylor, S.S., and Johnson, D.A. (1999). Binding-dependent disorder-order transition in PKI alpha: a fluorescence anisotropy study. Biochemistry (Mosc.) *38*, 6774–6780.

Haynes, C., Oldfield, C.J., Ji, F., Klitgord, N., Cusick, M.E., Radivojac, P., Uversky, V.N., Vidal, M., and Iakoucheva, L.M. (2006). Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput. Biol. *2*, e100.

He, Y., Chen, Y., Mooney, S.M., Rajagopalan, K., Bhargava, A., Sacho, E., Weninger, K., Bryan, P.N., Kulkarni, P., and Orban, J. (2015). Phosphorylation-Induced Conformational Ensemble Switching in an Intrinsically Disordered Cancer/Testis Antigen. J. Biol. Chem.

Holzbaur, E.L.F., and Vallee, R.B. (1994). Dyneins: Molecular Structure and Cellular Function. Annu. Rev. Cell Biol. *10*, 339–372.

Hook, P., and Vallee, R.B. (2006). The dynein family at a glance. J. Cell Sci. *119*, 4369–4371.

Huang, C.Y., Chang, C.P., Huang, C.L., and Ferrell, J.E. (1999). M phase phosphorylation of cytoplasmic dynein intermediate chain and p150(Glued). J. Biol. Chem. *274*, 14262–14269.

Hwang, T.L., van Zijl, P.C., and Mori, S. (1998). Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. J. Biomol. NMR *11*, 221–226.

Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradović, Z., and Dunker, A.K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. J. Mol. Biol. *323*, 573–584.

Iakoucheva, L.M., Radivojac, P., Brown, C.J., O'Connor, T.R., Sikes, J.G., Obradovic, Z., and Dunker, A.K. (2004). The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res. *32*, 1037–1049.

Ichikawa, M., Saito, K., Yanagisawa, H.-A., Yagi, T., Kamiya, R., Yamaguchi, S., Yajima, J., Kushida, Y., Nakano, K., Numata, O., et al. (2015). Axonemal dynein light chain-1 locates at the microtubule-binding domain of the γ heavy chain. Mol. Biol. Cell *26*, 4236–4247.

Ilangovan, U., Ding, W., Zhong, Y., Wilson, C.L., Groppe, J.C., Trbovich, J.T., Zúñiga, J., Demeler, B., Tang, Q., Gao, G., et al. (2005). Structure and dynamics of the homodimeric dynein light chain km23. J. Mol. Biol. *352*, 338–354.

Jahn, T.R., and Radford, S.E. (2005). The Yin and Yang of protein folding. FEBS J. *272*, 5962–5970.

Jensen, M.R., Ruigrok, R.W.H., and Blackledge, M. (2013). Describing intrinsically disordered proteins at atomic resolution by NMR. Curr. Opin. Struct. Biol. *23*, 426–435.

Jie, J., Löhr, F., and Barbar, E. (2015). Interactions of yeast dynein with dynein light chain and dynactin: General implications for intrinsically disordered duplex scaffolds in multi-protein assemblies. J. Biol. Chem.

Johnson, B.A. (2004). Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol. Biol. *278*, 313–352.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. *292*, 195–202.

Kardon, J.R., and Vale, R.D. (2009). Regulators of the cytoplasmic dynein motor. Nat. Rev. Mol. Cell Biol. *10*, 854–865.

Kardon, J.R., Reck-Peterson, S.L., and Vale, R.D. (2009). Regulation of the processivity and intracellular localization of Saccharomyces cerevisiae dynein by dynactin. Proc. Natl. Acad. Sci. U. S. A. *106*, 5669–5674.

Karki, S., and Holzbaur, E.L. (1995). Affinity chromatography demonstrates a direct binding between cytoplasmic dynein and the dynactin complex. J. Biol. Chem. *270*, 28806–28811.

Karki, S., and Holzbaur, E.L. (1999). Cytoplasmic dynein and dynactin in cell division and intracellular transport. Curr. Opin. Cell Biol. *11*, 45–53.

Kathiriya, J.J., Pathak, R.R., Clayman, E., Xue, B., Uversky, V.N., and Davé, V. (2014). Presence and utility of intrinsically disordered regions in kinases. Mol. Biosyst. *10*, 2876–2888.

Kidane, A.I., Song, Y., Nyarko, A., Hall, J., Hare, M., Löhr, F., and Barbar, E. (2013). Structural features of LC8-induced self-association of swallow. Biochemistry (Mosc.) *52*, 6011–6020.

King, S.J., and Schroer, T.A. (2000). Dynactin increases the processivity of the cytoplasmic dynein motor. Nat. Cell Biol. *2*, 20–24.

King, S.J., Brown, C.L., Maier, K.C., Quintyne, N.J., and Schroer, T.A. (2003). Analysis of the dynein-dynactin interaction in vitro and in vivo. Mol. Biol. Cell *14*, 5089–5097.

King, S.M., Barbarese, E., Dillman, J.F., Patel-King, R.S., Carson, J.H., and Pfister, K.K. (1996a). Brain cytoplasmic and flagellar outer arm dyneins share a highly conserved Mr 8,000 light chain. J. Biol. Chem. *271*, 19358–19366.

King, S.M., Dillman, J.F., Benashski, S.E., Lye, R.J., Patel-King, R.S., and Pfister, K.K. (1996b). The Mouse t-Complex-encoded Protein Tctex-1 Is a Light Chain of Brain Cytoplasmic Dynein. J. Biol. Chem. *271*, 32281–32287.

Kini, A.R., and Collins, C.A. (2001). Modulation of cytoplasmic dynein ATPase activity by the accessory subunits. Cell Motil. Cytoskeleton *48*, 52–60.

Kjaergaard, M., Brander, S., and Poulsen, F.M. (2011). Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. J. Biomol. NMR *49*, 139–149.

Knight, P.J., Thirumurugan, K., Xu, Y., Wang, F., Kalverda, A.P., Stafford, W.F., 3rd, Sellers, J.R., and Peckham, M. (2005). The predicted coiled-coil domain of myosin 10 forms a novel elongated domain that lengthens the head. J. Biol. Chem. *280*, 34702–34708.

Kon, T., Oyama, T., Shimo-Kon, R., Imamula, K., Shima, T., Sutoh, K., and Kurisu, G. (2012). The 2.8 Å crystal structure of the dynein motor domain. Nature *484*, 345–350.

Koonin, E.V., and Aravind, L. (2000). Dynein light chains of the Roadblock/LC7 group belong to an ancient protein superfamily implicated in NTPase regulation. Curr. Biol. CB *10*, R774-776.

Kuta, A., Deng, W., Morsi El-Kadi, A., Banks, G.T., Hafezparast, M., Pfister, K.K., and Fisher, E.M.C. (2010). Mouse Cytoplasmic Dynein Intermediate Chains: Identification of New Isoforms, Alternative Splicing and Tissue Distribution of Transcripts. PLoS ONE *5*, e11682.

Lakomek, N.-A., Ying, J., and Bax, A. (2012). Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. J. Biomol. NMR *53*, 209–221.

Lam, C., Vergnolle, M.A.S., Thorpe, L., Woodman, P.G., and Allan, V.J. (2010). Functional interplay between LIS1, NDE1 and NDEL1 in dynein-dependent organelle positioning. J. Cell Sci. *123*, 202–212.

Lareau, L.F., Green, R.E., Bhatnagar, R.S., and Brenner, S.E. (2004). The evolving roles of alternative splicing. Curr. Opin. Struct. Biol. *14*, 273–282.

Latysheva, N.S., Flock, T., Weatheritt, R.J., Chavali, S., and Babu, M.M. (2015). How do disordered regions achieve comparable functions to structured domains? Protein Sci. Publ. Protein Soc. *24*, 909–922.

Li, null, Romero, null, Rani, null, Dunker, null, and Obradovic, null (1999). Predicting Protein Disorder for N-, C-, and Internal Regions. Genome Inform. Workshop Genome Inform. *10*, 30–40.

Liang, J., Jaffrey, S.R., Guo, W., Snyder, S.H., and Clardy, J. (1999). Structure of the PIN/LC8 dimer with a bound peptide. Nat. Struct. Biol. *6*, 735–740.

Liu, J.-F., Wang, Z.-X., Wang, X.-Q., Tang, Q., An, X.-M., Gui, L.-L., and Liang, D.-C. (2006). Crystal structure of human dynein light chain Dnlc2A: structural insights into the interaction with IC74. Biochem. Biophys. Res. Commun. *349*, 1125–1129.

Lo, K.W., Naisbitt, S., Fan, J.S., Sheng, M., and Zhang, M. (2001). The 8-kDa dynein light chain binds to its targets via a conserved (K/R)XTQT motif. J. Biol. Chem. *276*, 14059–14066.

Löhr, F., Schmidt, J.M., and Rüterjans, H. (1999). Simultaneous Measurement of 3JHN,Hα and 3JHα,Hβ Coupling Constants in 13C,15N-Labeled Proteins. J. Am. Chem. Soc. *121*, 11821–11826.

Lu, M.S., and Prehoda, K.E. (2013). A NudE/14-3-3 pathway coordinates dynein and the kinesin Khc73 to position the mitotic spindle. Dev. Cell *26*, 369–380.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science *252*, 1162–1164.

Maehama, T., and Dixon, J.E. (1998). The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. J. Biol. Chem. *273*, 13375–13378.

Makokha, M., Hare, M., Li, M., Hays, T., and Barbar, E. (2002). Interactions of cytoplasmic dynein light chains Tctex-1 and LC8 with the intermediate chain IC74. Biochemistry (Mosc.) *41*, 4302–4311.

Makokha, M., Huang, Y.J., Montelione, G., Edison, A.S., and Barbar, E. (2004). The solution structure of the pH-induced monomer of dynein light-chain LC8 from Drosophila. Protein Sci. Publ. Protein Soc. *13*, 727–734.

Malaney, P., Pathak, R.R., Xue, B., Uversky, V.N., and Davé, V. (2013). Intrinsic disorder in PTEN and its interactome confers structural plasticity and functional versatility. Sci. Rep. *3*, 2035.

Manning, G., Plowman, G.D., Hunter, T., and Sudarsanam, S. (2002). Evolution of protein kinase signaling from yeast to man. Trends Biochem. Sci. *27*, 514–520.

Mark, W.-Y., Liao, J.C.C., Lu, Y., Ayed, A., Laister, R., Szymczyna, B., Chakrabartty, A., and Arrowsmith, C.H. (2005). Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? J. Mol. Biol. *345*, 275–287.

Markus, S.M., Punch, J.J., and Lee, W.-L. (2009). Motor- and Tail-Dependent Targeting of Dynein to Microtubule Plus Ends and the Cell Cortex. Curr. Biol. *19*, 196–205.

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

McKenney, R.J., Weil, S.J., Scherer, J., and Vallee, R.B. (2011). Mutually Exclusive Cytoplasmic Dynein Regulation by NudE-Lis1 and Dynactin. J. Biol. Chem. *286*, 39615–39622.

Metcalfe, E.E., Traaseth, N.J., and Veglia, G. (2005). Serine 16 phosphorylation induces an order-to-disorder transition in monomeric phospholamban. Biochemistry (Mosc.) *44*, 4386–4396.

Miller, K.G. (2016). Keeping Neuronal Cargoes on the Right Track: New Insights into Regulators of Axonal Transport. Neurosci. Rev. J. Bringing Neurobiol. Neurol. Psychiatry.

Mische, S., He, Y., Ma, L., Li, M., Serr, M., and Hays, T.S. (2008). Dynein light intermediate chain: an essential subunit that contributes to spindle checkpoint inactivation. Mol. Biol. Cell *19*, 4918–4929.

Mitchell, D.J., Blasier, K.R., Jeffery, E.D., Ross, M.W., Pullikuth, A.K., Suo, D., Park, J., Smiley, W.R., Lo, K.W.-H., Shabanowitz, J., et al. (2012). Trk activation of the ERK1/2 kinase pathway stimulates intermediate chain phosphorylation and recruits cytoplasmic dynein to signaling endosomes for retrograde axonal transport. J. Neurosci. Off. J. Soc. Neurosci. *32*, 15495–15510.

Mitrea, D.M., Grace, C.R., Buljan, M., Yun, M.-K., Pytel, N.J., Satumba, J., Nourse, A., Park, C.-G., Madan Babu, M., White, S.W., et al. (2014). Structural polymorphism in the N-terminal oligomerization domain of NPM1. Proc. Natl. Acad. Sci. U. S. A. *111*, 4466–4471.

Mittag, T., and Forman-Kay, J.D. (2007). Atomic-level characterization of disordered protein ensembles. Curr. Opin. Struct. Biol. *17*, 3–14.

Mok, Y.K., Lo, K.W., and Zhang, M. (2001). Structure of Tctex-1 and its interaction with cytoplasmic dynein intermediate chain. J. Biol. Chem. *276*, 14067–14074.

Morgan, J.L., Song, Y., and Barbar, E. (2011). Structural Dynamics and Multiregion Interactions in Dynein-Dynactin Recognition. J. Biol. Chem. *286*, 39349–39359.

Moughamian, A.J., and Holzbaur, E.L.F. (2012). Dynactin is required for transport initiation from the distal axon. Neuron *74*, 331–343.

Nagano, F., Orita, S., Sasaki, T., Naito, A., Sakaguchi, G., Maeda, M., Watanabe, T., Kominami, E., Uchiyama, Y., and Takai, Y. (1998). Interaction of Doc2 with tctex-1, a light chain of cytoplasmic dynein. Implication in dynein-dependent vesicle transport. J. Biol. Chem. *273*, 30065–30068.

Neudecker, P., Lundström, P., and Kay, L.E. (2009). Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding. Biophys. J. *96*, 2045–2054.

Nishi, H., Fong, J.H., Chang, C., Teichmann, S.A., and Panchenko, A.R. (2013). Regulation of protein-protein binding by coupling between phosphorylation and intrinsic disorder: analysis of human protein complexes. Mol. Biosyst. *9*, 1620–1626.

Nurminsky, D.I., Nurminskaya, M.V., Benevolenskaya, E.V., Shevelyov, Y.Y., Hartl, D.L., and Gvozdev, V.A. (1998). Cytoplasmic dynein intermediate-chain isoforms with different targeting properties created by tissue-specific alternative splicing. Mol. Cell. Biol. *18*, 6816–6825.

Nyarko, A., and Barbar, E. (2011). Light chain-dependent self-association of dynein intermediate chain. J. Biol. Chem. *286*, 1556–1566.

Nyarko, A., Song, Y., and Barbar, E. (2012). Intrinsic Disorder in Dynein Intermediate Chain Modulates Its Interactions with NudE and Dynactin. J. Biol. Chem. *287*, 24884–24893.

Nyarko, A., Song, Y., Nováček, J., Žídek, L., and Barbar, E. (2013). Multiple recognition motifs in nucleoporin Nup159 provide a stable and rigid Nup159-Dyn2 assembly. J. Biol. Chem. *288*, 2614–2622.

Oates, M.E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M.J., Xue, B., Dosztányi, Z., Uversky, V.N., Obradovic, Z., Kurgan, L., et al. (2013). D^2P^2: database of disordered protein predictions. Nucleic Acids Res. *41*, D508-516.

Ochiai, K., Watanabe, M., Ueki, H., Huang, P., Fujii, Y., Nasu, Y., Noguchi, H., Hirata, T., Sakaguchi, M., Huh, N.-H., et al. (2011). Tumor suppressor REIC/Dkk-3 interacts with the dynein light chain, Tctex-1. Biochem. Biophys. Res. Commun. *412*, 391–395.

Oldfield, C.J., and Dunker, A.K. (2014). Intrinsically disordered proteins and intrinsically disordered protein regions. Annu. Rev. Biochem. *83*, 553–584.

Oldfield, C.J., Cheng, Y., Cortese, M.S., Romero, P., Uversky, V.N., and Dunker, A.K. (2005). Coupled folding and binding with alpha-helix-forming molecular recognition elements. Biochemistry (Mosc.) *44*, 12454–12470.

Ori-McKenney, K.M., Xu, J., Gross, S.P., and Vallee, R.B. (2010). A cytoplasmic dynein tail mutation impairs motor processivity. Nat. Cell Biol. *12*, 1228–1234.

Pajkos, M., Mészáros, B., Simon, I., and Dosztányi, Z. (2012). Is there a biological cost of protein disorder? Analysis of cancer-associated mutations. Mol. Biosyst. *8*, 296–307.

Pazour, G.J., Wilkerson, C.G., and Witman, G.B. (1998). A dynein light chain is essential for the retrograde particle movement of intraflagellar transport (IFT). J. Cell Biol. *141*, 979–992.

Peckham, M., and Knight, P.J. (2009). When a predicted coiled coil is really a single α-helix, in myosins and other proteins. Soft Matter *5*, 2493–2503.

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. U. S. A. *94*, 12366–12371.

Pfister, K.K., Shah, P.R., Hummerich, H., Russ, A., Cotton, J., Annuar, A.A., King, S.M., and Fisher, E.M.C. (2006). Genetic analysis of the cytoplasmic dynein subunit families. PLoS Genet. *2*, e1.

Pierce, W.K., Grace, C.R., Lee, J., Nourse, A., Marzahn, M.R., Watson, E.R., High, A.A., Peng, J., Schulman, B.A., and Mittag, T. (2016). Multiple Weak Linear Motifs Enhance Recruitment and Processivity in SPOP-Mediated Substrate Ubiquitination. J. Mol. Biol. *428*, 1256–1271.

Pullikuth, A.K., Ozdemir, A., Cardenas, D., Bailey, E., Sherman, N.E., Pfister, K.K., and Catling, A.D. (2013). Epidermal growth factor stimulates extracellular-signal regulated kinase phosphorylation of a novel site on cytoplasmic Dynein intermediate chain 2. Int. J. Mol. Sci. *14*, 3595–3620.

Purohit, A., Tynan, S.H., Vallee, R., and Doxsey, S.J. (1999). Direct interaction of pericentrin with cytoplasmic dynein light intermediate chain contributes to mitotic spindle organization. J. Cell Biol. *147*, 481–492.

Qiu, R., Zhang, J., and Xiang, X. (2013). Identification of a novel site in the tail of dynein heavy chain important for dynein function in vivo. J. Biol. Chem. *288*, 2271–2280.

Raaijmakers, J.A., and Medema, R.H. (2014). Function and regulation of dynein in mitotic chromosome segregation. Chromosoma *123*, 407–422.

Rao, L., Romes, E.M., Nicholas, M.P., Brenner, S., Tripathy, A., Gennerich, A., and Slep, K.C. (2013). The yeast dynein Dyn2-Pac11 complex is a dynein dimerization/processivity factor: structural and single-molecule characterization. Mol. Biol. Cell *24*, 2362–2377.

Rapali, P., Szenes, Á., Radnai, L., Bakos, A., Pál, G., and Nyitray, L. (2011a). DYNLL/LC8: a light chain subunit of the dynein motor complex and beyond. FEBS J. *278*, 2980–2996.

Rapali, P., García-Mayoral, M.F., Martínez-Moreno, M., Tárnok, K., Schlett, K., Albar, J.P., Bruix, M., Nyitray, L., and Rodriguez-Crespo, I. (2011b). LC8 dynein light chain (DYNLL1) binds to the C-terminal domain of ATM-interacting protein (ATMIN/ASCIZ) and regulates its subcellular localization. Biochem. Biophys. Res. Commun. *414*, 493–498.

Raux, H., Flamand, A., and Blondel, D. (2000). Interaction of the rabies virus P protein with the LC8 dynein light chain. J. Virol. *74*, 10212–10216.

Rayala, S.K., den Hollander, P., Manavathi, B., Talukder, A.H., Song, C., Peng, S., Barnekow, A., Kremerskothen, J., and Kumar, R. (2006). Essential role of KIBRA in co-activator function of dynein light chain 1 in mammalian cells. J. Biol. Chem. *281*, 19092–19099.

Receveur-Brechot, V., and Durand, D. (2012). How random are intrinsically disordered proteins? A small angle scattering perspective. Curr. Protein Pept. Sci. *13*, 55–75.

Reck-Peterson, S.L., Yildiz, A., Carter, A.P., Gennerich, A., Zhang, N., and Vale, R.D. (2006). Single-molecule analysis of dynein processivity and stepping behavior. Cell *126*, 335–348.

Roberts, A.J., Numata, N., Walker, M.L., Kato, Y.S., Malkova, B., Kon, T., Ohkura, R., Arisaka, F., Knight, P.J., Sutoh, K., et al. (2009). AAA+ Ring and linker swing mechanism in the dynein motor. Cell *136*, 485–495.

Rogerson, D.T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaite, A., Hancock, S.M., Huguenin-Dezot, N., Muqit, M.M.K., Fry, A.M., Bayliss, R., et al. (2015). Efficient

genetic encoding of phosphoserine and its nonhydrolyzable analog. Nat. Chem. Biol. *11*, 496–503.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. Proteins *42*, 38–48.

Romero, P.R., Zaidi, S., Fang, Y.Y., Uversky, V.N., Radivojac, P., Oldfield, C.J., Cortese, M.S., Sickmeier, M., LeGall, T., Obradovic, Z., et al. (2006). Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. Proc. Natl. Acad. Sci. U. S. A. *103*, 8390–8395.

Romes, E.M., Tripathy, A., and Slep, K.C. (2012). Structure of a yeast Dyn2-Nup159 complex and molecular basis for dynein light chain-nuclear pore interaction. J. Biol. Chem. *287*, 15862–15873.

Runnegar, M.T., Wei, X., and Hamm-Alvarez, S.F. (1999). Increased protein phosphorylation of cytoplasmic dynein results in impaired motor function. Biochem. J. *342 (Pt 1)*, 1–6.

Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. (1998). TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc. Natl. Acad. Sci. U. S. A. *95*, 13585–13590.

Sasaki, S., Shionoya, A., Ishida, M., Gambello, M.J., Yingling, J., Wynshaw-Boris, A., and Hirotsune, S. (2000). A LIS1/NUDEL/cytoplasmic dynein heavy chain complex in the developing and adult nervous system. Neuron *28*, 681–696.

Schanda, P., Van Melckebeke, H., and Brutscher, B. (2006). Speeding up three-dimensional protein NMR experiments to a few minutes. J. Am. Chem. Soc. *128*, 9042–9043.

Schiavo, G., Greensmith, L., Hafezparast, M., and Fisher, E.M.C. (2013). Cytoplasmic dynein heavy chain: the servant of many masters. Trends Neurosci. *36*, 641–651.

Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. Cell *103*, 211–225.

Schmitz, M.L., dos Santos Silva, M.A., Altmann, H., Czisch, M., Holak, T.A., and Baeuerle, P.A. (1994). Structural and functional analysis of the NF-kappa B p65 C terminus. An acidic and modular transactivation domain with the potential to adopt an alpha-helical conformation. J. Biol. Chem. *269*, 25613–25620.

Schroeder, C.M., Ostrem, J.M.L., Hertz, N.T., and Vale, R.D. (2014). A Ras-like domain in the light intermediate chain bridges the dynein motor to a cargo-binding region. eLife *3*, e03351.

Schroer, T.A. (2004). Dynactin. Annu. Rev. Cell Dev. Biol. *20*, 759–779.

Sharma, R., Raduly, Z., Miskei, M., and Fuxreiter, M. (2015). Fuzzy complexes: Specific binding without complete folding. FEBS Lett. *589*, 2533–2542.

Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J.M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K.K., Lemak, A., et al. (2008). Consistent blind protein structure generation from NMR chemical shift data. Proc. Natl. Acad. Sci. U. S. A. *105*, 4685–4690.

Shoemaker, B.A., Portman, J.J., and Wolynes, P.G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc. Natl. Acad. Sci. U. S. A. *97*, 8868–8873.

Showalter, S.A. (2007). Intrinsically Disordered Proteins: Methods for Structure and Dynamics Studies. In eMagRes, (John Wiley & Sons, Ltd), p.

Shu, T., Ayala, R., Nguyen, M.-D., Xie, Z., Gleeson, J.G., and Tsai, L.-H. (2004). Ndel1 operates in a common pathway with LIS1 and cytoplasmic dynein to regulate cortical neuronal positioning. Neuron *44*, 263–277.

Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V.N., et al. (2007). DisProt: the Database of Disordered Proteins. Nucleic Acids Res. *35*, D786-793.

Siglin, A.E., Sun, S., Moore, J.K., Tan, S., Poenie, M., Lear, J.D., Polenova, T., Cooper, J.A., and Williams, J.C. (2013). Dynein and dynactin leverage their bivalent character to form a high-affinity interaction. PloS One *8*, e59453.

Slevin, L.K., Romes, E.M., Dandulakis, M.G., and Slep, K.C. (2014). The mechanism of dynein light chain LC8-mediated oligomerization of the Ana2 centriole duplication factor. J. Biol. Chem. *289*, 20727–20739.

Song, J., Tyler, R.C., Lee, M.S., Tyler, E.M., and Markley, J.L. (2005). Solution structure of isoform 1 of Roadblock/LC7, a light chain in the dynein complex. J. Mol. Biol. *354*, 1043–1051.

Straube, A., Enard, W., Berner, A., Wedlich-Söldner, R., Kahmann, R., and Steinberg, G. (2001). A split motor domain in a cytoplasmic dynein. EMBO J. *20*, 5091–5100.

Stuchell-Brereton, M.D., Siglin, A., Li, J., Moore, J.K., Ahmed, S., Williams, J.C., and Cooper, J.A. (2011). Functional interaction between dynein light chain and intermediate chain is required for mitotic spindle positioning. Mol. Biol. Cell *22*, 2690–2701.

Susalka, S.J., Nikulina, K., Salata, M.W., Vaughan, P.S., King, S.M., Vaughan, K.T., and Pfister, K.K. (2002). The roadblock light chain binds a novel region of the cytoplasmic Dynein intermediate chain. J. Biol. Chem. *277*, 32939–32946.

Tai, A.W., Chuang, J.Z., and Sung, C.H. (2001). Cytoplasmic dynein regulation by subunit heterogeneity and its role in apical transport. J. Cell Biol. *153*, 1499–1509.

Talbott, M., Hare, M., Nyarko, A., Hays, T.S., and Barbar, E. (2006). Folding is coupled to dimerization of Tctex-1 dynein light chain. Biochemistry (Mosc.) *45*, 6793–6800.

Tang, Q., Staub, C.M., Gao, G., Jin, Q., Wang, Z., Ding, W., Aurigemma, R.E., and Mulder, K.M. (2002). A novel transforming growth factor-beta receptor-interacting protein that is also a light chain of the motor protein dynein. Mol. Biol. Cell *13*, 4484–4496.

Theillet, F.-X., Binolfi, A., Bekei, B., Martorana, A., Rose, H.M., Stuiver, M., Verzini, S., Lorenz, D., van Rossum, M., Goldfarb, D., et al. (2016). Structural disorder of monomeric α-synuclein persists in mammalian cells. Nature *530*, 45–50.

Tompa, P. (2002). Intrinsically unstructured proteins. Trends Biochem. Sci. *27*, 527–533.

Tompa, P., and Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. Trends Biochem. Sci. *33*, 2–8.

Tynan, S.H., Purohit, A., Doxsey, S.J., and Vallee, R.B. (2000). Light intermediate chain 1 defines a functional subfraction of cytoplasmic dynein which binds to pericentrin. J. Biol. Chem. *275*, 32763–32768.

Urnavicius, L., Zhang, K., Diamant, A.G., Motz, C., Schlager, M.A., Yu, M., Patel, N.A., Robinson, C.V., and Carter, A.P. (2015). The structure of the dynactin complex and its interaction with dynein. Science *347*, 1441–1446.

Uversky, V.N. (2002). Natively unfolded proteins: a point where biology waits for physics. Protein Sci. Publ. Protein Soc. *11*, 739–756.

Uversky, V.N. (2013a). Unusual biophysics of intrinsically disordered proteins. Biochim. Biophys. Acta *1834*, 932–951.

Uversky, V.N. (2013b). A decade and a half of protein intrinsic disorder: Biology still waits for physics. Protein Sci. Publ. Protein Soc. *22*, 693–724.

Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are “natively unfolded” proteins unstructured under physiologic conditions? Proteins *41*, 415–427.

Uversky, V.N., Oldfield, C.J., and Dunker, A.K. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu. Rev. Biophys. *37*, 215–246.

Vale, R.D. (2003). The molecular motor toolbox for intracellular transport. Cell *112*, 467–480.

Vallee, R.B., Williams, J.C., Varma, D., and Barnhart, L.E. (2004). Dynein: An ancient motor protein involved in multiple modes of transport. J. Neurobiol. *58*, 189–200.

Vallee, R.B., McKenney, R.J., and Ori-McKenney, K.M. (2012). Multiple modes of cytoplasmic dynein regulation. Nat. Cell Biol. *14*, 224–230.

van der Lee, R., Lang, B., Kruse, K., Gsponer, J., Sánchez de Groot, N., Huynen, M.A., Matouschek, A., Fuxreiter, M., and Babu, M.M. (2014). Intrinsically Disordered Segments Affect Protein Half-Life in the Cell and during Evolution. Cell Rep. *8*, 1832–1844.

Varadi, M., Vranken, W., Guharoy, M., and Tompa, P. (2015). Computational approaches for inferring the functions of intrinsically disordered proteins. Front. Mol. Biosci. *2*, 45.

Vaughan, P.S. (2001). Cytoplasmic Dynein Intermediate Chain Phosphorylation Regulates Binding to Dynactin. J. Biol. Chem. *276*, 26171–26179.

Vaughan, K.T., and Vallee, R.B. (1995a). Cytoplasmic dynein binds dynactin through a direct interaction between the intermediate chains and p150Glued. J. Cell Biol. *131*, 1507–1516.

Vaughan, K.T., and Vallee, R.B. (1995b). Cytoplasmic dynein binds dynactin through a direct interaction between the intermediate chains and p150Glued. J. Cell Biol. *131*, 1507–1516.

Vetter, S.W., and Leclerc, E. (2001). Phosphorylation of serine residues affects the conformation of the calmodulin binding domain of human protein 4.1. Eur. J. Biochem. FEBS *268*, 4292–4299.

Wagner, W., Fodor, E., Ginsburg, A., and Hammer, J.A., 3rd (2006). The binding of DYNLL2 to myosin Va requires alternatively spliced exon B and stabilizes a portion of the myosin's coiled-coil domain. Biochemistry (Mosc.) *45*, 11564–11577.

Wainman, A., Creque, J., Williams, B., Williams, E.V., Bonaccorsi, S., Gatti, M., and Goldberg, M.L. (2009). Roles of the Drosophila NudE protein in kinetochore function and centrosome migration. J. Cell Sci. *122*, 1747–1758.

Walsh, I., Martin, A.J.M., Di Domenico, T., and Tosatto, S.C.E. (2012). ESpritz: accurate and fast prediction of protein disorder. Bioinforma. Oxf. Engl. *28*, 503–509.

Wang, S., and Zheng, Y. (2011). Identification of a novel dynein binding domain in nudel essential for spindle pole organization in Xenopus egg extract. J. Biol. Chem. *286*, 587–593.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Wang, L., Hare, M., Hays, T.S., and Barbar, E. (2004). Dynein light chain LC8 promotes assembly of the coiled-coil domain of swallow protein. Biochemistry (Mosc.) *43*, 4611–4620.

Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., and Jones, D.T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J. Mol. Biol. *337*, 635–645.

Whyte, J., Bader, J.R., Tauhata, S.B.F., Raycroft, M., Hornick, J., Pfister, K.K., Lane, W.S., Chan, G.K., Hinchcliffe, E.H., Vaughan, P.S., et al. (2008). Phosphorylation regulates targeting of cytoplasmic dynein to kinetochores during mitosis. J. Cell Biol. *183*, 819–834.

Williams, J.C., Xie, H., and Hendrickson, W.A. (2005). Crystal structure of dynein light chain TcTex-1. J. Biol. Chem. *280*, 21981–21986.

Williams, J.C., Roulhac, P.L., Roy, A.G., Vallee, R.B., Fitzgerald, M.C., and Hendrickson, W.A. (2007). Structural and thermodynamic characterization of a cytoplasmic dynein light chain-intermediate chain complex. Proc. Natl. Acad. Sci. U. S. A. *104*, 10028–10033.

Williams, R.M., Obradovi, Z., Mathura, V., Braun, W., Garner, E.C., Young, J., Takayama, S., Brown, C.J., and Dunker, A.K. (2001). The protein non-folding problem: amino acid determinants of intrinsic order and disorder. Pac. Symp. Biocomput. Pac. Symp. Biocomput. 89–100.

Williamson, M.P. (2013). Using chemical shift perturbation to characterise ligand binding. Prog. Nucl. Magn. Reson. Spectrosc. *73*, 1–16.

Wodarz, A., and Nusse, R. (1998). Mechanisms of Wnt signaling in development. Annu. Rev. Cell Dev. Biol. *14*, 59–88.

Wright, P.E., and Dyson, H.J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J. Mol. Biol. *293*, 321–331.

Wright, P.E., and Dyson, H.J. (2009). Linking folding and binding. Curr. Opin. Struct. Biol. *19*, 31–38.

Xie, H., Vucetic, S., Iakoucheva, L.M., Oldfield, C.J., Dunker, A.K., Obradovic, Z., and Uversky, V.N. (2007). Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. J. Proteome Res. *6*, 1917–1932.

Xu, W., Doshi, A., Lei, M., Eck, M.J., and Harrison, S.C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. Mol. Cell *3*, 629–638.

Yang, Z., Tulu, U.S., Wadsworth, P., and Rieder, C.L. (2007). Kinetochore dynein is required for chromosome motion and congression independent of the spindle checkpoint. Curr. Biol. *17*, 973–980.

Ye, G.J., Vaughan, K.T., Vallee, R.B., and Roizman, B. (2000). The herpes simplex virus 1 U(L)34 protein interacts with a cytoplasmic dynein intermediate chain and targets nuclear membrane. J. Virol. *74*, 1355–1363.

Yoder, J.H., and Han, M. (2001). Cytoplasmic dynein light intermediate chain is required for discrete aspects of mitosis in Caenorhabditis elegans. Mol. Biol. Cell *12*, 2921–2933.

Zala, D., Hinckelmann, M.-V., and Saudou, F. (2013). Huntingtin's Function in Axonal Transport Is Conserved in Drosophila melanogaster. PloS One *8*, e60162.

Zhang, J., Li, S., Fischer, R., and Xiang, X. (2003). Accumulation of cytoplasmic dynein and dynactin at microtubule plus ends in Aspergillus nidulans is kinesin dependent. Mol. Biol. Cell *14*, 1479–1488.

Zhang, J., Li, S., Musa, S., Zhou, H., and Xiang, X. (2009). Dynein light intermediate chain in Aspergillus nidulans is essential for the interaction between heavy and intermediate chains. J. Biol. Chem. *284*, 34760–34768.

Appendices

Appendix 1

Regulation of Dynein IC Involves a Novel Interplay between Phosphorylation and Length of Disordered Spliced Linkers - supplemental material

Jing Jie, Frank Löhr, and Elisar Barbar

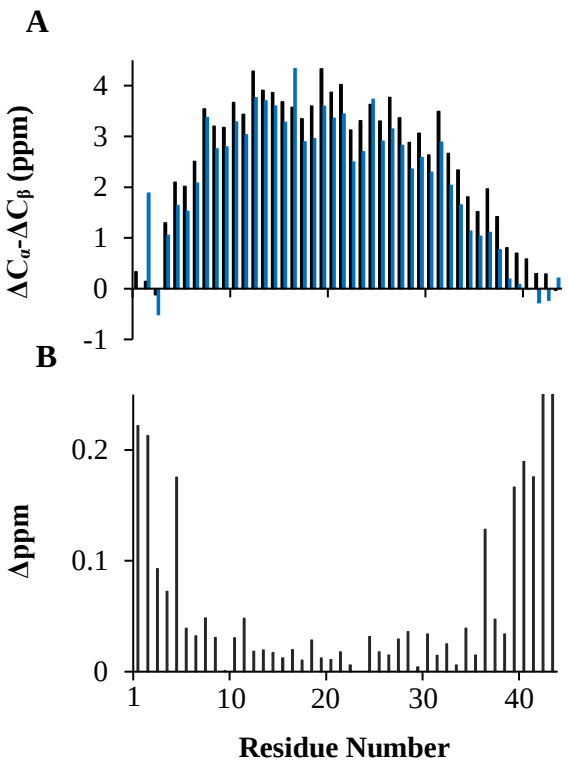


Figure A1.1. IC-2C$_{44}$ structure. (A) Secondary chemical shifts of IC-2C$_{44}$ (blue) compared to those of IC-2C$_{96}$ (black of Fig. 4.1B). (B) Chemical shift perturbation of IC-2C$_{44}$ compared to IC-2C$_{96}$.

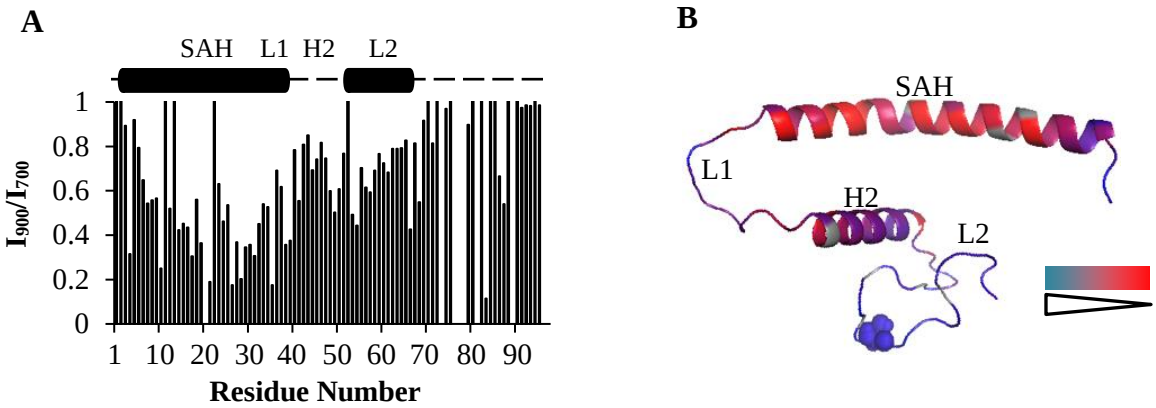


Figure A1.2. SAH and H2 pack in multiple conformations in intermediate exchange in apo IC-2C$_{96}$. (A) Ratio of peak intensities in apo-IC-2C$_{96}$ collected at 700 MHz *vs* those at 900 MHz. Residues in SAH and H2 have lower ratios indicating they are in intermediate exchange. (B) These differences are mapped on a CS-Rosetta structure of IC-2C$_{96}$ (smallest to largest ratio, red to blue). S84 residue is shown as spheres.

Appendix 2

Crystallization of the mammalian dynein intermediate chain (IC) in complex with Nudel

Jing Jie

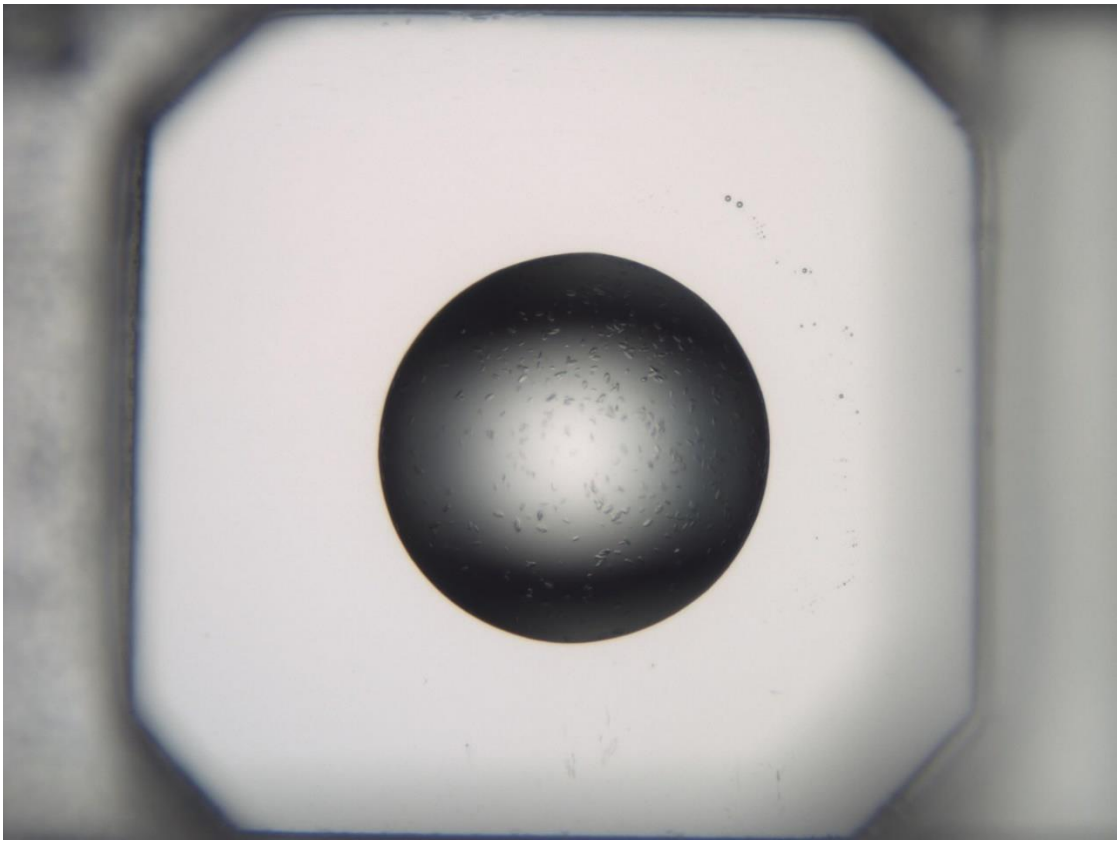


Figure A2.1. Crystal formation of the IC-2C$_{44}$-nNudel protein complex (8 mg/ml). Crystallization buffer is 0.1 M HEPES, 30% PEG 300, pH 7.5 at 20°C.

Appendix 3

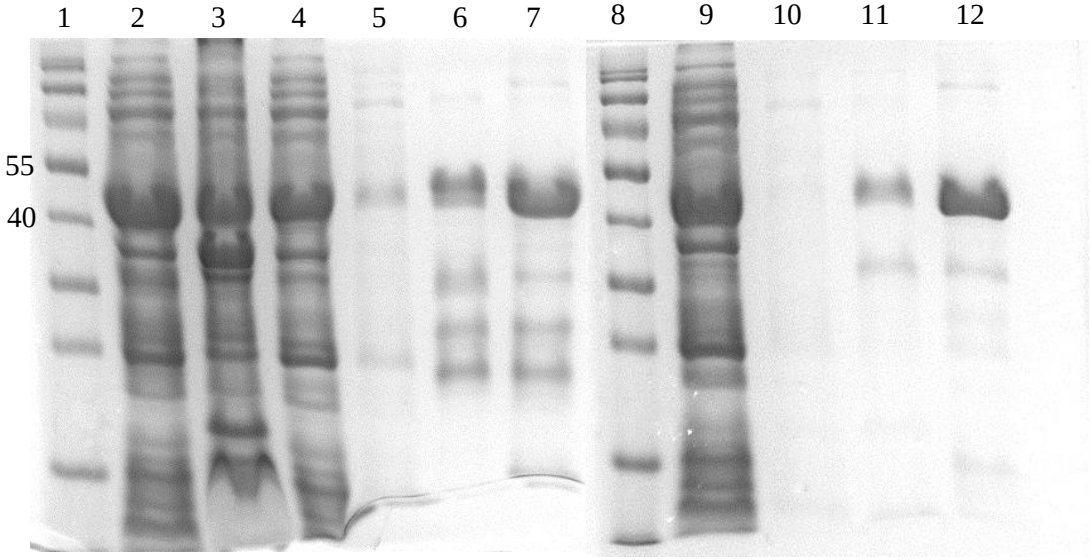
Purification of Human and Drosophila Huntingtin and its interaction with the Dynein Intermediate Chain

Jing Jie

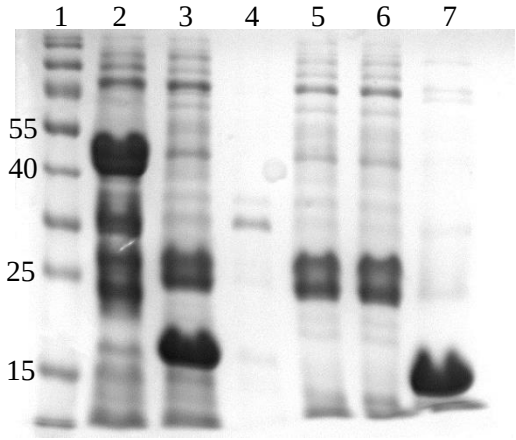
Figure A3.1. Purification of human Htt 536-698. (A) SDS-PAGE analysis of the Ni-NTA purification of His_6-sumo-Htt. Gel labels are protein marker (1), supernatant (2), pellet (3), flow through-a (4), 10 mM imidazole wash-a (5), 50 mM imidazole wash-a (6), 350 mM imidazole wash-a (7), protein marker (8), flow through-b (9), 10 mM imidazole wash-b (10), 50 mM imidazole wash-b (11), 350 mM imidazole wash-b (12). The His6-sumo-Htt 536-698 (31 KDa) was expressed at 26 °C with 0.4 mM IPTG. The protein is soluble (lane 2). Due to incomplete binding to Ni-NTA resin as evidenced by protein in lane 4, rebinding to Ni-NTA resin was performed. Protein presence in lane 9 in combination with the contamination in lane 7 suggests the weak binding affinity of the His_6 tag on this protein to Ni-NTA resin. Proteins in lanes 6 and 12 were combined, concentrated and subjected to sumo cleavage. (B) SDS-PAGE analysis of sumo cleavage and subsequent Ni-NTA purification of Htt. Gel labels are protein marker (1), before cutting (2), supernatant after cutting (3), pellet after cutting (4), flow through (5), 10 mM imidazole wash (6), 350 mM imidazole wash (7). Precipitation were observed after sumo cleavage but did not contain Htt protein (lane 3). Due to impurity in the sample prior to sumo cleavage (lane 2), the fractions containing Htt (17 KDa) have two bands (lanes 5 and 6) around 25 KDa and which one is Htt is unclear. (C) SDS-PAGE analysis of size exclusion chromatography (SEC) fractions of Htt.

Figure A3.1. (*Continued*)

A



B



A

1 2 3 4 5 6 7

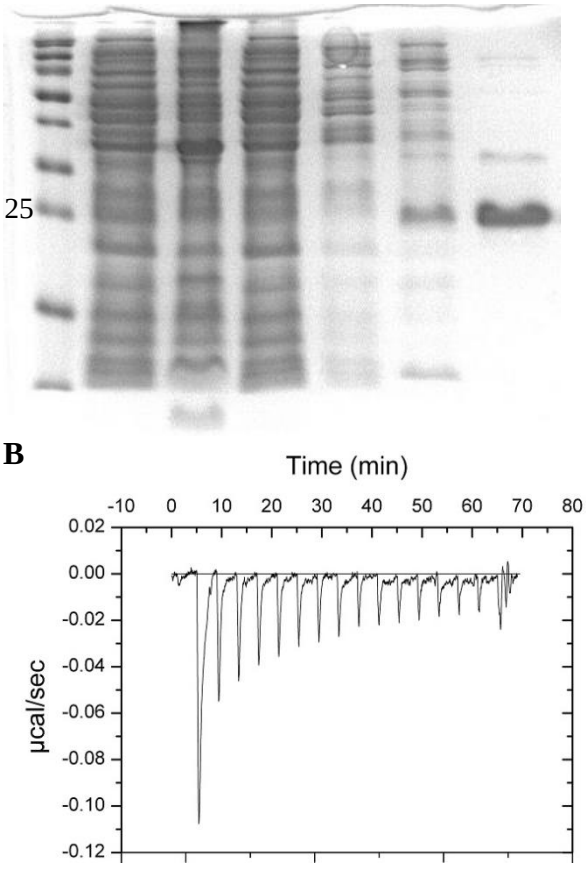


Figure A3.2. Purification of Drosophila Htt 380-481 and its interaction with dynein N-IC. (A) SDS-PAGE analysis of the Ni-NTA purification of His_6-Tev-Htt. Gel labels are protein marker (1), supernatant (2), pellet (3), flow through-a (4), 10 mM imidazole wash-a (5), 50 mM imidazole wash-a (6), 350 mM imidazole wash-a (7). The His6-Tev-Htt 380-481 (14 KDa) was expressed at 26 °C with 0.4 mM IPTG. Protein in lane 7 was concentrated used in ITC experiments with Drosophila IC 1-143. (B) Representative ITC thermogram for IC-Htt interaction at 25 °C. ITC buffer contains 50 mM sodium phosphate, 50 mM NaCl, 0.5 mM NaN_3, 5 mM BME, pH 7.5. The heat release associated with each injection is small.