Biophysical*Journal*
Article

Biophysical Society

# Interplay of Disorder and Sequence Specificity in the Formation of Stable Dynein-Dynactin Complexes

Nikolaus M. Loening,[1] Sanjana Saravanan,[2] Nathan E. Jespersen,[2] Kayla Jara,[2] and Elisar Barbar[2,*]
[1]Department of Chemistry, Lewis & Clark College, Portland, Oregon and [2]Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon

ABSTRACT Cytoplasmic dynein is a eukaryotic motor protein complex that, along with its regulatory protein dynactin, is essential to the transport of organelles within cells. The interaction of dynein with dynactin is regulated by binding between the intermediate chain (IC) subunit of dynein and the p150$^{Glued}$ subunit of dynactin. Even though in the rat versions of these proteins this interaction primarily involves the single $\alpha$-helix region at the N-terminus of the IC, in *Drosophila* and yeast ICs the removal of a nascent helix (H2) downstream of the single $\alpha$-helix considerably diminishes IC-p150$^{Glued}$ complex stability. We find that for ICs from various species, there is a correlation between disorder in H2 and its contribution to binding affinity, and that sequence variations in H2 that do not change the level of disorder show similar binding behavior. Analysis of the structure and interactions of the IC from *Chaetomium thermophilum* demonstrates that the H2 region of *C. thermophilum* IC has a low helical propensity and establishes that H2 binds directly to the coiled-coil 1B (CC1B) domain of p150$^{Glued}$, thus explaining why H2 is necessary for tight binding. Isothermal titration calorimetry, circular dichroism, and NMR studies of smaller CC1B constructs localize the region of CC1B most essential for a tight interaction with IC. These results suggest that it is the level of disorder in H2 of IC along with its charge, rather than sequence specificity, that underlie its importance in initiating tight IC-p150$^{Glued}$ complex formation. We speculate that the nascent H2 helix may provide conformational flexibility to initiate binding, whereas those species that have a fully folded H2 have co-opted an alternative mechanism for promoting p150$^{Glued}$ binding.

SIGNIFICANCE The intrinsically disordered N-terminal domain of the dynein intermediate chain is thought to be central to the function of this motor protein because it specifically binds cargo and regulates the assembly and function of the whole complex through its binding to the p150$^{Glued}$ subunit of dynactin. We have found that the structural propensity of the helix 2 (H2) region of the intermediate chain correlates with a decrease in the strength of the binding interaction with p150$^{Glued}$ and that the H2 region interacts directly with p150$^{Glued}$ rather than indirectly via the single $\alpha$-helix region. One possible model that explains our results is that H2 makes multiple nonspecific contacts with p150$^{Glued}$ before the single $\alpha$-helix locks into a more specific binding interaction.

## INTRODUCTION

Cytoplasmic dynein is a 1.6 MDa microtubule-based motor protein complex that is essential for centrosome separation during mitosis, mitotic spindle assembly, axonogenesis, and the retrograde transport of organelles and other cellular cargo (1–3). The dynein complex is composed of pairs of subunits that segregate the roles of generating motion and cargo attachment; the heavy chain subunits contain the microtubule binding domains and generate motion via ATPase activity, whereas the intermediate chain (IC), light IC, and light chain subunits play roles in assembly and stabilization of the dynein complex, cargo attachment, and modulating dynein activity (4–6). Dynein carries out its functions by binding to various cofactors and regulatory proteins (2), including dynactin, a 1.2 MDa protein complex composed of more than 20 subunits, the largest of which is p150$^{Glued}$ (7). p150$^{Glued}$ is present in two copies in the dynactin complex and forms coiled-coil domains at multiple locations, one of which, coiled-coil 1B (CC1B), directly interacts with the dynein IC (8–10). Because dynein is involved in numerous critical cellular processes, inhibition of dynein severely disrupts normal functioning of cells and, in particular, disruption of the interaction between

---

*Submitted February 28, 2020, and accepted for publication July 27, 2020.*

*Correspondence: barbare@oregonstate.edu

Nikolaus M. Loening and Sanjana Saravanan contributed equally to this work.

Editor: Scott Showalter.

https://doi.org/10.1016/j.bpj.2020.07.023

dynein and dynactin causes a motor neuron degeneration pathology resembling amyotrophic lateral sclerosis (11). Axonal transport, autophagy, and the clearance of aggregated proteins are also affected by defective dynein regulation, which suggests that dynein misregulation plays a role in the development of several neurodegenerative diseases, including Huntington's, Parkinson's, and Alzheimer's diseases (12).

The locus of dynein's interaction with p150$^{Glued}$ is the N-terminal region of the IC, a region that contains $\alpha$-helical segments flanked by intrinsically disordered linkers (13). Although cryo-electron microscopy structures of the entire dynactin-dynein-BicD2 complex exist (14,15), the p150$^{Glued}$-IC portion in those structures is invisible because of the disorder inherent in the system. Therefore, the p150$^{Glued}$-IC interaction is one of the final missing pieces for a complete description of the complex. Intrinsically disordered proteins like the N-terminal region of the IC (16–18) and the C-terminal region of the dynein light intermediate chain (19,20) maintain partial disorder even when interacting with partners, providing flexibility and functional versatility.

The interactions between IC and p150$^{Glued}$ from rat, *Drosophila*, and yeast primarily depend on a highly conserved, N-terminal single $\alpha$-helix (SAH) of IC (8,13,21) and a coiled-coil region (CC1B) of p150$^{Glued}$ (9,10). Results from nuclear magnetic resonance (NMR) spectroscopy have identified a second helix (H2) (8,21) that is nascent in *Drosophila* and yeast ICs but fully formed in rat IC (Figs. 1 *A* and 2 *B*; (9,13)). Isothermal titration calorimetry (ITC) studies showed that whereas removing H2 did not significantly affect the binding of rat p150$^{Glued}$ to IC (13), its removal weakened this interaction in yeast (21) and *Drosophila* (8) proteins (Table S1). These

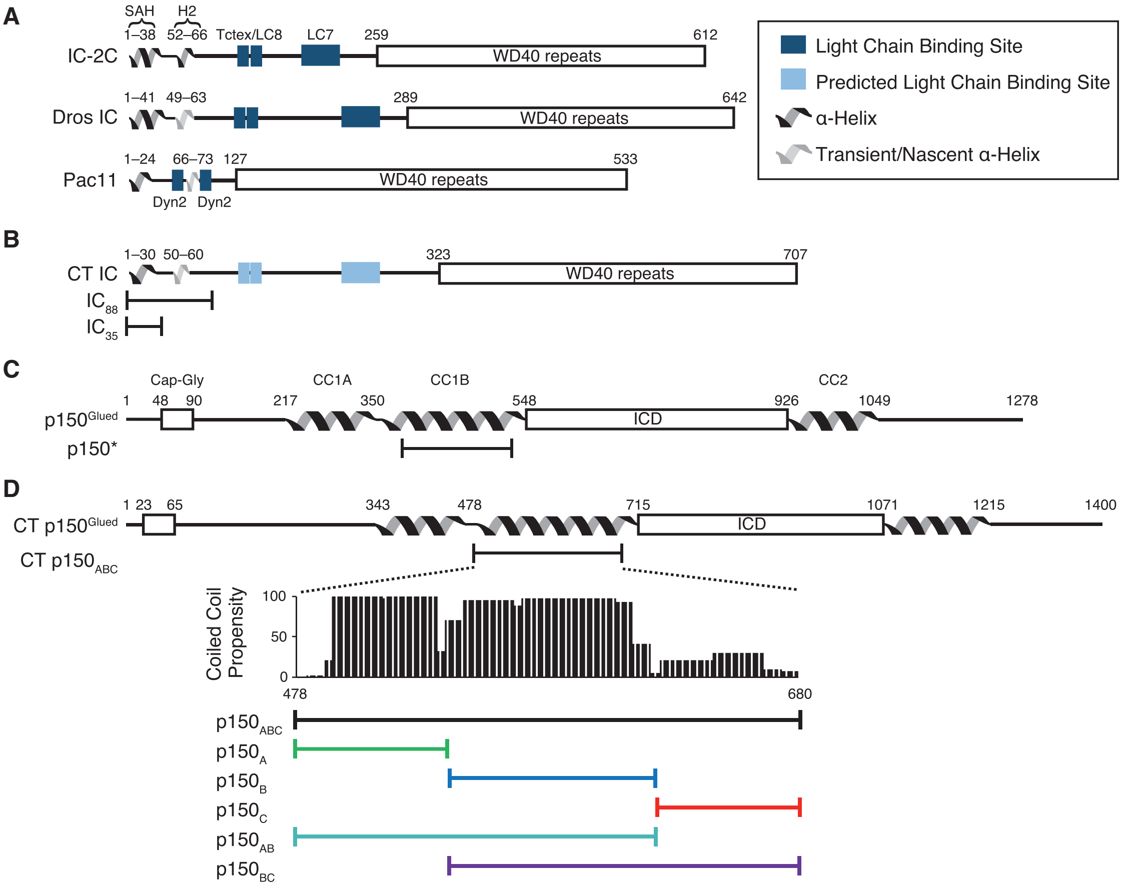
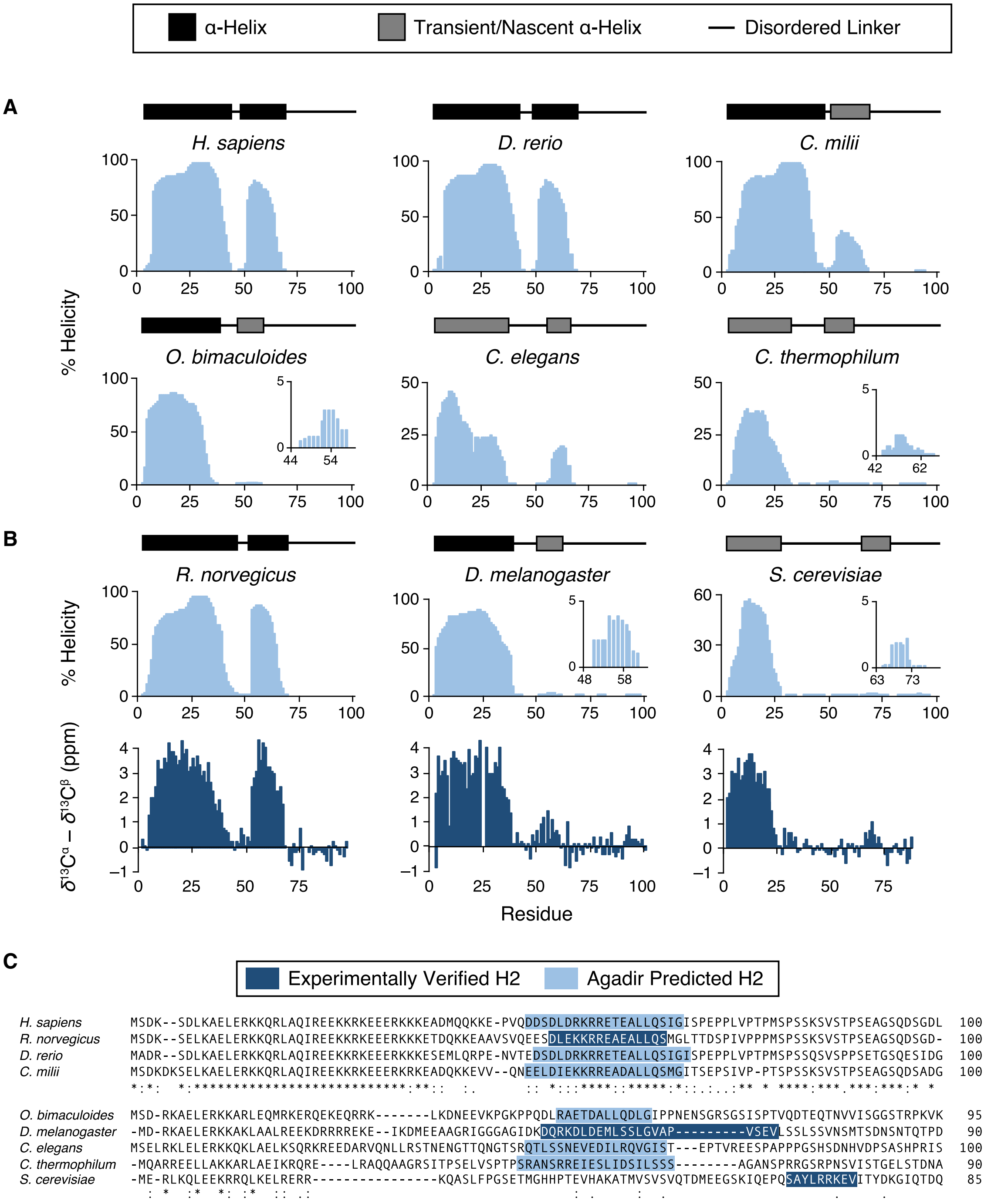


FIGURE 1 Dynein intermediate chain and dynactin p150$^{Glued}$ domain architecture. (*A*) ICs from rat (IC-2C), *Drosophila* (Dros IC), and yeast (IC homolog Pac 11) have an N-terminal single $\alpha$-helix (SAH) as well as a second helix (H2 (8,13,21)). H2 is either fully formed (*black helix*) or transient/nascent (*gray helix*). Rat and *Drosophila* ICs have binding sites for three light chains (Tctex, LC8, and LC7), whereas the yeast IC has binding sites for two copies of an LC8 homolog (DYN2). The C-terminal region for all ICs studied is predicted to contain seven WD40 repeat domains. (*B*) The IC from *C. thermophilum* (CT IC) is predicted to have an $\alpha$-helical SAH region and a nascent helix in the H2 region. Based on sequence motifs, the Tctex, LC8, and LC7 binding sites are predicted to be in regions similar to Dros IC. $IC_{88}$ and $IC_{35}$ are the constructs used in this work. (*C*) Mammalian p150$^{Glued}$ has a Cap-Gly domain near the N-terminus and two coiled-coil domains, CC1 and CC2, that are separated by an intercoil domain (ICD). CC1 is further divided into two regions called CC1A and CC1B. The construct p150* was used in previous studies. (*D*) CT p150$^{Glued}$ is predicted to have similar domains as mammalian p150$^{Glued}$. The COILS Server prediction tool was used to predict the coiled-coil propensity of each residue in the CC1B region ($p150_{ABC}$), which informed the selection of the smaller p150$^{Glued}$ constructs. To see this figure in color, go online.



```
H. sapiens        MSDK--SDLKAELERKKQRLAQIREEKKRKEEERKKKEADMQQKKE-PVQDDSDLDRKRRETEALLQSIGISPEPPLVPTPMSPSSKSVSTPSEAGSQDSGDL  100
R. norvegicus     MSDK--SELKAELERKKQRLAQIREEKKRKEEERKKKETDQKKEAAVSVQEESDLEKKRREAEALLQSMGLTTDSPIVPPPMSPSSKSVSTPSEAGSQDSGD-  100
D. rerio          MADR--SDLKAELERKKQRLAQIREEKKRKEEERKKKESEMLQRPE-NVTEDSDLDRKRRETEALLQSIGISPEPPLVPTPMSPSSQSVSPPSETGSQESIDG  100
C. milii          MSDKDKSELKAELERKKQRLAQIREEKKRKEEERKRKEADQKKEVV--QNEELDIEKKRREADALLQSMGITSEPSIVP-PTSPSSKSVSTPSEAGSQDSADG  100
                  *:*:  *:***************************:**::  :.      :: *:::****::*****:*::.:..:** * ****:***.***:***:* *

O. bimaculoides   MSD-RKAELERKKARLEQMRKERQEKEQRRK-------LKDNEEVKPGKPPQDLRAETDALLQDLGIPPNENSGRSGSISPTVQDTEQTNVVISGGSTRPKVK   95
D. melanogaster   -MD-RKAELERKKAKLAALREEKDRRRREKE--IKDMEEAAGRIGGGAGIDKDQRKDLDEMLSSLGVAP---------VSEVLSSLSSVNSMTSDNSNTQTPD   90
C. elegans        MSELRKLELERKKQKLAELKSQRKREEDARVQNLLRSTNENGTTQNGTSRQTLSSNEVEDILRQVGIST---EPTVREESPAPPPGSHSDNHVDPSASHPRIS  100
C. thermophilum   -MQARREELLAKKARLAEIKRQRE----LRAQQAAGRSITPSELVSPTPSRANSRREIESLIDSILSSS--------AGANSPRRGSRPNSVISTGELSTDNA   90
S. cerevisiae     -ME-RLKQLEEKRRQLKELRERR---------------KQASLFPGSETMGHHPTEVHAKATMVSVSVQTDMEEGSKIQEPQSAYLRRKEVITYDKGIQTDQ   85
                    : *  :*  *: :*  :: .:                                 : .     :  .                    .     .
```

FIGURE 2 Disorder in H2 is species dependent. (*A*) Residue-level percent helicity predictions generated using the Agadir algorithm for the first 100 amino acids of IC from *H. sapiens* (human), *D. rerio* (zebrafish), *C. milii* (Australian ghostshark), *O. bimaculoides* (Californian two-spot octopus), *C. elegans* (nematode), and CT (thermophilic fungus). The insets for *O. bimaculoides* and CT show detail for the predicted H2 region. (*B*) The top graphs show the percent helicity prediction for the ICs from *R. norvegicus* (rat), *D. melanogaster* (fruit fly), and *S. cerevisiae* (yeast). The insets for *Drosophila* and yeast show detail for the predicted H2 region. The bottom graphs show NMR secondary chemical shift data for IC constructs from rat, *Drosophila*, and yeast from prior studies (8,13,21). (*C*) Sequence alignments of ICs from chordates (human, rat, zebrafish, and shark) and from nonchordates (octopus, fruit fly, CT, and yeast) using the MAFFT alignment program (58–60). Highlighted residues show predicted (*light blue*) and experimentally determined (*dark blue*) H2 regions. Identical (*asterisk*), strongly similar (*colon*), and weakly similar (*period*) residues are shown at the bottom of each alignment. To see this figure in color, go online.

observations suggest that although a fully folded H2 does not contribute to binding of p150$^{Glued}$, a nascent H2 does play a role, although it is unclear what this role is in any of these species.

In this study, we set out to address three questions about the interaction between the IC and p150$^{Glued}$: Is the structural propensity of the H2 region species dependent? Is there direct interaction between H2 and p150$^{Glued}$? And where on the CC1B region does IC bind? Using a helical propensity prediction algorithm, we observed that the length and location of H2 is conserved in a wide range of species (Fig. 2). To explore the function and contributions of H2 in p150$^{Glued}$ binding, we focused on proteins from the model organism *Chaetomium thermophilum* (CT), a thermophilic fungus that grows at temperatures as high as 60°C (22). In addition to its thermostability, we selected this system because, as seen in Fig. 2 *A*, Agadir predicts only a very low level of helicity for the H2 region. ITC, together with circular dichroism (CD) and NMR, was used to establish the importance of a nascent H2 helix in the interaction between dynein and dynactin and to define the locus of IC binding on the p150$^{Glued}$ CC1B region.

## MATERIALS AND METHODS

### Helix prediction of ICs using Agadir

Sequences for the dynein IC from a range of species were obtained from the UniProt protein database (23). *Rattus norvegicus* (rat, UniProt: Q62871-3), *Drosophila melanogaster* (fruit fly, UniProt: Q24246-11), and *Saccharomyces cerevisiae* (yeast, UniProt: P40960-1) were selected because of the availability of biophysical data from previous studies (8,13,21). Additionally, six other species were selected to provide phylogenetic diversity: *Homo sapiens* (human, UniProt: O14576-2), *Danio rerio* (zebrafish, UniProt: A1A5Y4-1), *Callorhinchus milii* (Australian ghostshark, UniProt: V9KAN3-1), *Octopus bimaculoides* (Californian two-spot octopus, UniProt: A0A0L8HM30-1), *Caenorhabditis elegans* (nematode, UniProt: O45087-1), and CT (thermophilic fungus, UniProt: G0SCF1-1). The first 100 amino acids of the IC from each species were scored using the Agadir algorithm, which outputs a prediction for percent helicity at the residue level (24–28).

### Protein expression and purification

The CT IC$_{88}$ (residues 1–88), IC$_{35}$ (residues 1–35), p150$_{ABC}$ (residues 478–680), p150$_{AB}$ (residues 478–622), p150$_{BC}$ (residues 542–680), p150$_{A}$ (residues 478–541), p150$_{B}$ (residues 542–622), and p150$_{C}$ (residues 623–680) constructs were prepared by PCR and cloned into a pET-24d vector with an N-terminal His$_6$ tag using the Gibson Assembly protocol (29,30). In addition, an IC$_{88}$ construct with the 11 amino acid residues of the H2 region replaced by 11 residues from the H2 region of rat IC was generated by the same method. The gene for CT IC$_{37–88}$ (residues 37–88) was synthesized and inserted into a pET-24a(+) vector by GeneScript (Piscataway, NJ). For all constructs, an internal tobacco etch virus (TEV) protease cleavage site was included to allow the removal of the His$_6$ tag. IC$_{37–88}$ was engineered so that after cleavage with TEV protease, no non-native residues were present, whereas for the other constructs, a non-native GAHM sequence remained at the N-terminus of the proteins after cleavage. DNA sequences were verified by Sanger sequencing. The recombinant plasmids were transformed into Rosetta (DE3) *E. coli* cells (Merck KGaA, Darmstadt, Germany) for protein expression.

Bacterial cultures for expression of unlabeled proteins were grown in lysogeny broth at 37°C to an optical density (A$_{600}$) of 0.6–0.8, whereas cultures for expression of isotopically labeled ($^{15}$N or $^{15}$N/$^{13}$C) proteins were grown in MJ9 minimal media (31) at 37°C to an A$_{600}$ of 0.8. Protein synthesis was then induced with 0.4 mM isopropyl-$\beta$-D-1-thiogalactopyranoside, and expression proceeded overnight at 26°C. Cells were harvested from the cultures by centrifugation at 2500 rcf, lysed by sonication, and then centrifuged a second time at 24,000 rcf to clarify the lysate. The resulting supernatant was purified by immobilized metal affinity chromatography using TALON Metal Affinity Resin (Takara Bio USA, Mountain View, CA), after which the His$_6$ tag was cleaved using TEV protease. Proteins were further purified using a Superdex 75 (Cytiva, Marlborough, MA) size-exclusion chromatography (SEC) column or, for IC$_{37–88}$, a HiPrep 16/60 Sephacryl S-100 HR SEC column (Cytiva).

Protein concentrations were determined by measuring the absorbance at 280 nm for proteins with nonzero molar extinction coefficients at this wavelength. Molar extinction coefficients measured at 280 nm were computed for p150$_{ABC}$, p150$_{AB}$, p150$_{BC}$, p150$_{A}$, and p150$_{C}$, using the ProtParam tool on the ExPASy website (32). The concentrations of proteins that lack aromatic residues, including IC$_{88}$, IC$_{35}$, IC$_{37–88}$, and p150$_{B}$, were determined using a Bradford assay (Bio-Rad, Hercules, CA), absorbance at 205 nm, and/or by comparing their band sizes on SDS polyacrylamide gels against standards. Molar extinction coefficients for absorbance measurements at 205 nm were computed with a Protein Calculator tool (33). All purified proteins were stored at 4°C with a protease inhibitor mixture of pepstatin A and phenylmethanesulfonyl fluoride and used within 1 week. Gels run before the use of the samples showed no evidence of proteolysis when stained with Coomassie blue, and NMR confirmed that samples were stable over the period of at least a week because additional peaks from proteolyzed fragments were not observed over this time period.

### CD measurements

CD measurements were made using a JASCO (Easton, MD) J-720 CD spectropolarimeter. Samples consisted of protein at concentrations of 5–20 $\mu$M in 10 mM sodium phosphate buffer (pH 7.5). All experiments were done using a 300 $\mu$L cuvette with a path length of 0.1 cm. The data shown are the average of three scans. Thermal unfolding data were collected in increments of 5°C over a temperature range of 5–50°C. Fractional helicity estimations based on the CD measurements were made using the following equation:

$$Fractional\ Helicity = \frac{(\theta_{222}^{exp} + 3000\ \mathrm{deg\ cm^2\ dmol^{-1}})}{-36,500\ \mathrm{deg\ cm^2\ dmol^{-1}}}, \quad (1)$$

where $\theta_{222}^{exp}$ is the experimentally observed mean residue ellipticity (MRE) at 222 nm (34).

### ITC

ITC experiments were conducted using a MicroCal VP-ITC microcalorimeter (Malvern Panalytical, Malvern, UK). All experiments were performed at 25°C, with the protein samples in a buffer containing 50 mM sodium phosphate (pH 7.5), 50 mM sodium chloride, and 1 mM sodium azide. Samples were degassed at 25°C for 15 min before loading. Each experiment was started with a 2 $\mu$L injection and followed by 27–33 injections of 10 $\mu$L. In every experiment, IC constructs at a concentration of 200–300 $\mu$M were titrated into p150$^{Glued}$ constructs at a concentration of 20–30 $\mu$M.

### Analytical ultracentrifugation

To form the IC$_{88}$-p150$_{ABC}$ complex, p150$_{ABC}$ was combined with an excess of IC$_{88}$, after which the complex was purified using SEC on a Superdex 200 column (GE Healthcare Life Sciences, Pittsburgh, PA). The buffer conditions for the final purification were 20 mM Tris-HCl, 50 mM NaCl, 1 mM sodium azide, 5 mM DTT (pH 7.4). Samples were checked for purity and a lack of proteolysis using SDS-PAGE.

Sedimentation velocity analytical ultracentrifugation (SV-AUC) experiments with p150$_{ABC}$ at a concentration of 80 $\mu$M and the IC$_{88}$-p150$_{ABC}$ complex at a concentration of 60 $\mu$M were performed with a Beckman Optima XL-A analytical ultracentrifuge (Beckman Coulter Life Sciences, Indianapolis, IN). The buffer conditions were the same as the final purification, except that the 5 mM DTT was replaced with 5 mM tris(2-carboxyethyl)phosphine. The complex was loaded into a standard double-sectored cell with quartz windows and a 1.2 cm pathlength and centrifuged at 42,000 rpm at 20°C using an An-60 Ti analytical rotor (Beckman Coulter). 300 scans were acquired at 280 nm with no interscan delay. Data were fitted to a c(*s*) distribution using the software SEDFIT (35).

## NMR measurements and analysis

NMR samples were buffer exchanged into a 50 mM sodium chloride, 50 mM phosphate (pH 6.5) buffer that included 5% $D_2O$, 1 mM sodium azide, 0.2 mM 2,2-dimethylsilapentane-5-sulfonic acid, and 1× cOmplete protease inhibitor cocktail (Roche, Basel, Switzerland). All NMR spectra were collected with a sample temperature of 15°C using a Bruker (Billerica, MA) Avance III HD 800 MHz spectrometer with a TCI cryoprobe or, for IC$_{37–88}$, a Bruker Avance NEO 600 MHz spectrometer with a room-temperature TXI probe. Band-selective excitation short transient variants of TROSY-based triple resonance sequences (HNCO, HN(CA)CO, HNCACB, HN(CO)CACB) were used for backbone assignment (36), whereas side-chain assignments were performed using $^{15}$N-TOCSY-HSQC and HCCH TOCSY experiments.

For IC$_{88}$, transverse ($R_2$) and longitudinal ($R_1$) $^{15}$N relaxation rates, as well as $\{^1\text{H}\}$-$^{15}$N heteronuclear nuclear Overhauser effect enhancements (NOEs), were measured for amide nitrogens using TROSY-based experiments (37). The longitudinal relaxation measurement used delay times of 20, 60, 100, 200, 400, 600, 800, and 1200 ms; the experiments with 60 and 600 ms delay times were collected in triplicate. The transverse relaxation measurement repeated a 17 ms CPMG block 1, 2, 4, 8, 10, 12, 14, and 16 times; the experiments with 2 and 12 repetitions of the CPMG block were collected in triplicate. For the HetNOE experiment, $^1$H spins were saturated by 120° radiofrequency pulses every 5 ms during the 6 s relaxation delay before the initial $^{15}$N pulse. For all three dynamics experiments, the data were collected in an interleaved fashion.

$^{15}$N-separated NOESY-TROSY (120 ms mixing time) and $^{13}$C-separated NOESY-HSQC (200 ms mixing time) experiments were used to generate NOE-based distance constraints for IC$_{88}$. $^3J_{HNH\alpha}$-values for IC$_{88}$ were estimated by measuring the intensity ratio between the cross-peaks and the diagonal peaks in a three-dimensional (3D) HNHA experiment. The accessibility of IC$_{88}$ amide protons to exchange with the solvent was determined from peak volumes in a Fast-HSQC spectrum with a 50 ms CLEANEX-PM mixing period (38).

Residual dipolar couplings ($^1D_{HN}$) for IC$_{88}$ were determined by first measuring scalar couplings ($^1J_{HN}$) for an isotropic $^{15}$N IC$_{88}$ sample by the difference in position between peaks in TROSY and semi-TROSY spectra. These spectra were then repeated for an aligned sample to measure the sum of the residual and dipolar couplings ($^1J_{HN} + {}^1D_{HN}$). IC$_{88}$ was aligned by soaking it into a cylindrical 6% polyacrylamide gel that was then compressed from a 5.4 mm diameter to a 4.2 mm diameter using a gel press (New Era, Vineland, NJ), resulting in a gel sample that was stretched along the axis of the magnet field. Under these conditions, the splitting of the HOD peak in the $^2$H spectrum was 6.4 Hz.

For measuring chemical exchange between different structural forms of p150$_B$, a series of TROSY-based $^{15}$N ZZ-exchange experiments (39) were acquired with mixing times of 10, 50, 100, 200, and 500 ms.

NMR data were processed using TopSpin 4.0.7 (Bruker) and NMRPipe (40). For 3D experiments that employed nonuniform sampling, the spectra were reconstructed using SCRUB (41). Peak assignment and relaxation analysis were performed using CCPN Analysis 2.5.0 (42).

## NMR titrations

NMR titrations were performed by combining $^{15}$N-labeled IC$_{88}$ with unlabeled p150$_{ABC}$, p150$_{AB}$, p150$_{BC}$, p150$_A$, p150$_B$, and p150$_C$. A small stoichiometric excess (1:1.2) was used when titrating with p150$_{AB}$, p150$_{BC}$, and p150$_C$; the other combinations used a 1:1 ratio of binding partners. The IC$_{88}$ concentration was in the range of 100–150 $\mu$M for these samples.

## Structure calculation

Structures were calculated using ARIA 2.3 (43), using chemical shift-matched NOESY peaks lists, backbone dihedral angle restraints generated using TALOS-N (44), residual dipolar couplings ($^1D_{HN}$), and $^3J_{HNH\alpha}$ couplings as inputs. $^1D_{HN}$ values were only used for the most rigid part of the SAH region (residues 5–28). The alignment tensor was determined to have an axial component ($D_a$) of $-28.1$ Hz and a rhombicity ($R$) of 0.066 by using REDCAT (45) and a preliminary structure of the SAH region that was calculated without RDCs. Approximately 25% of the NOE restraints (all of which corresponded to intraresidue or sequential contacts) were manually assigned, the remainder were automatically assigned starting from chemical shift-matched restraint lists using the ARIA protocol. The high degree of manual NOE restraint assignment was needed to achieve good convergence in the structure calculations because of the poor chemical shift dispersion for many of the amino acid types. Without manual assignments, the high degree of degeneracy in the side-chain chemical shifts occasionally resulted in erroneous structures in which IC$_{88}$ was bent into a hairpin because of the ARIA protocol assigning some NOE restraints as long-distance interactions when they could be more parsimoniously interpreted as short-range restraints. The standard ARIA protocol was used except that 30 structures were calculated in each iteration (other than for iteration 8, in which 50 structures were calculated), the log-harmonic potential was enabled (46), the spin diffusion correction was enabled, and the number of simulated annealing steps were tripled in the two cooling periods (cool1 and cool2). The 20 structures with the lowest energy from iteration 8 were selected for water refinement and used to generate the structural ensemble.

Statistics describing the restraints used for the calculations and the resulting ensemble are provided in Table S2. Chemical shifts and restraints were deposited at the Biological Magnetic Resonance Bank (BMRB: 30633), and atomic coordinates were deposited in the Worldwide Protein Data Bank (PDB: 6PQT).

# RESULTS

## Predicted H2 structural propensity varies among species

To determine how the level of helicity varies among species, we used the Agadir algorithm (24–28) for helical prediction. Because the Agadir algorithm does not account for long-range or tertiary interactions, it is most useful for predicting helicity in short peptides as well as helices in otherwise intrinsically disordered regions such as the N-terminus of the IC. Predictions for the selected species illustrate a highly conserved helix in the SAH region, whereas the H2 region varies in its helical propensity (Fig. 2, *A* and *B*). Human, rat, and zebrafish ICs are predicted to have a completely helical H2 region, whereas the H2 regions from Australian ghostshark and *C. elegans* are predicted to have intermediate helicity. For octopus, CT, *Drosophila*, and yeast, only a very small amount of helicity (i.e., a nascent/transient helix) is predicted for the H2 region, a result that on its own is not necessarily meaningful because these values are smaller than the standard deviation for Agadir helix predictions ($\pm 6\%$). Nevertheless, NMR secondary chemical shift data for rat, *Drosophila*, and yeast are consistent with the Agadir predictions (Fig. 2 *B*) for both the SAH and H2 regions, suggesting that in the case of ICs, Agadir is able to accurately predict the location of helical regions and their degree of helicity.

In chordates, the H2 region is highly conserved, with a high degree of helicity, whereas the H2 region in non-chordates shows little conservation in either sequence or position (Fig. 2 *C*). Interestingly, despite the lack of conservation in sequence, a nascent H2 helix is predicted in all species analyzed here, suggesting that this feature is in some way functionally important.

## Structural characterization of the IC from CT

We selected CT IC as a promising system for determining the importance of this region and its disorder in interactions with p150$^{\mathrm{Glued}}$. We designed a CT IC construct that included only the SAH region (IC$_{35}$, containing residues 1–35) and another construct that included the SAH, the potential H2, and the intrinsically disordered linker regions (IC$_{88}$, containing residues 1–88; Fig. 1 *B*).

The double minima observed at 222 and 208 nm (34) in CD spectra of IC$_{35}$ (Fig. 3 *A*) are consistent with completely helical structure. In contrast, the shift of the first minimum to 205 nm in the CD spectra of IC$_{88}$ (Fig. 3 *B*) suggests that it contains a mixture of helical and disordered regions. Analysis of the MRE at 222 nm indicates that IC$_{35}$ and IC$_{88}$ have fractional helicities of $\sim$80 and 35%, respectively, results that are consistent with the Agadir predictions (Fig. 2 *A*).

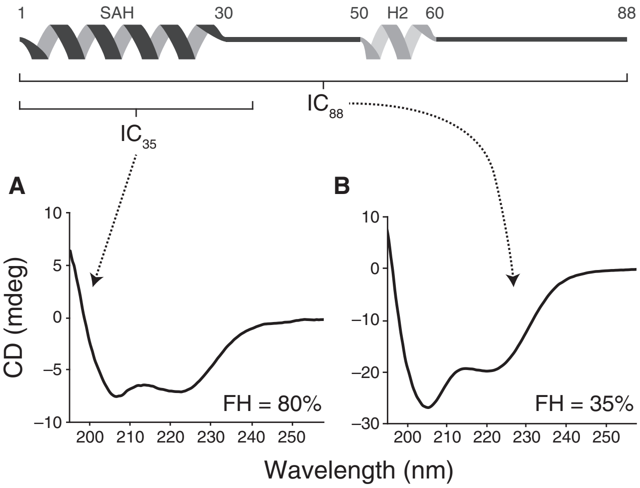


FIGURE 3 CD spectra of IC$_{35}$ and IC$_{88}$ at 15°C. The CD spectrum of IC$_{35}$ (*A*) shows double minima at 222 and 208 nm consistent with a completely helical structure, whereas the shift of the first minimum for IC$_{88}$ (*B*) to 205 nm suggests that its structure consists of a mixture of helical and disordered regions. FH indicates the fractional helicity estimated at 222 nm using Eq. 1.

We used NMR spectroscopy to confirm the location of the SAH and H2 regions for IC$_{88}$, to establish the degree of helical propensity of these regions, and to generate a structure of the ordered regions. Complete backbone and side-chain resonance assignments for IC$_{88}$ (Fig. 4 *A*) show evidence for both $\alpha$-helical and random-coil regions. Secondary chemical shift analysis (Fig. 4 *C*) clearly indicates an $\alpha$-helical structure for residues 1–30 (the SAH region) and a nascent $\alpha$-helix for residues 50–64 (the H2 region). Other residues with secondary chemical shift values close to zero are intrinsically disordered.

Three-bond $H_N$-$H_\alpha$ scalar coupling constants ($^{3}J_{\mathrm{HN\text{-}H}\alpha}$) for IC$_{88}$ (Fig. 4 *D*) are in the range of 4–6 Hz for residues 1–30, which implies that this region has an $\alpha$-helical secondary structure. The remaining residues have coupling values that are, for the most part, between 6 and 8 Hz, which is consistent with these residues lacking a fixed secondary structure. Likewise, measurements of one-bond H-N residual dipolar couplings ($^{1}D_{\mathrm{HN}}$, Fig. 4 *E*) show a cluster of positive values for residues 1–30 consistent with this region forming a stable $\alpha$-helix and do not show values indicative of structural propensity elsewhere. Both the $^{3}J_{\mathrm{HN\text{-}H}\alpha}$ and the $^{1}D_{\mathrm{HN}}$ coupling values support our prediction that the H2 region forms a nascent, rather than a fully formed, $\alpha$-helix.

Using a combination of NOESY, chemical shift-based backbone dihedral angle restraints, residual dipolar coupling restraints ($^{1}D_{\mathrm{HN}}$), and scalar coupling restraints ($^{3}J_{\mathrm{HNH}\alpha}$, which constrains the backbone dihedral angles), we generated the structural ensemble shown in Fig. 4 *G* (see also Table S2). As expected for a protein with no tertiary structure, the NOE contact map from the final cycle of the ARIA protocol (Fig. 4 *F*) shows that most of the NOE restraints correspond to short-range intraresidue or sequential ($i \pm 1$) contacts. For the SAH region, the pattern of medium-range $i \pm 3$ and $i \pm 4$ contacts for residues 5–32 is due to contacts between protons that are one turn of helix apart from one another and is a pattern expected for an $\alpha$-helical region. A similar, but weaker, pattern of medium-range contacts is seen for some residues in the H2 region.

The structures that make up the ensemble shown in Fig. 4 *G* were aligned based on residues 5–25 (within the SAH region) and show consistency in the SAH length. The portion of the H2 region that is regular enough to be rendered as a helix varies considerably from structure to structure, a result consistent with a nascent, rather than a stable, helix. As expected, the ensemble does not reveal any evidence of tertiary structure, consistent with the residues outside of the SAH and H2 regions being intrinsically disordered.

For proteins studied at high magnetic fields, longitudinal ($R_1$) $^{15}$N relaxation rates decrease and transverse ($R_2$) $^{15}$N relaxation rates increase for amides in more structured

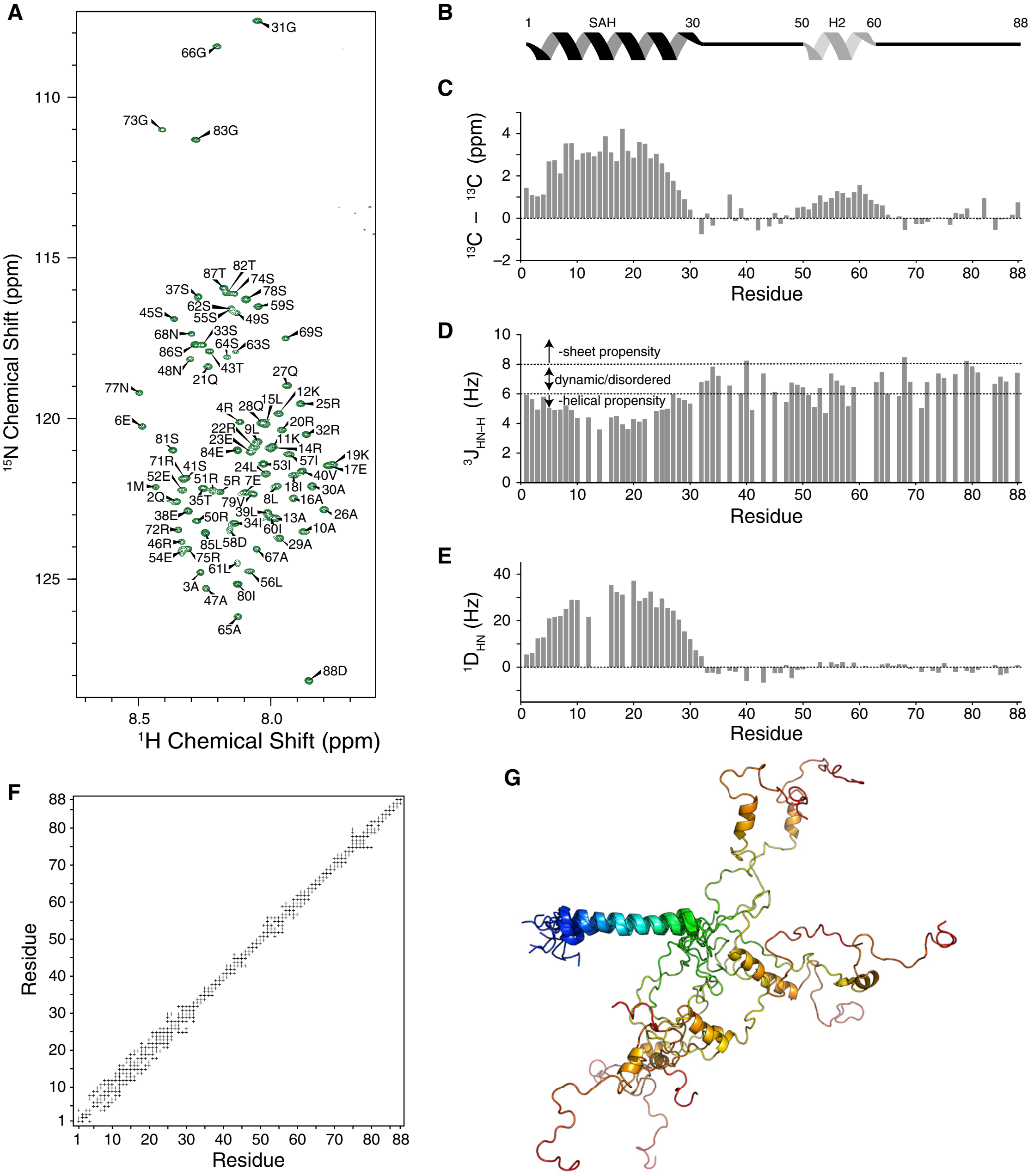


FIGURE 4 NMR assignments, secondary chemical shifts, couplings, and structural ensemble for $IC_{88}$. (*A*) $^{15}N$-$^{1}H$ TROSY-HSQC spectrum of $IC_{88}$ acquired at 800 MHz for $^{1}H$ showing assignments for the amide resonances. (*B*) Schematic structure for $IC_{88}$ showing the location of the $\alpha$-helical SAH region and the nascent $\alpha$-helix H2 region. (*C*) Secondary chemical shifts for $IC_{88}$; clusters of residues with positive values correspond to $\alpha$-helical regions, whereas values close to zero correspond to intrinsically disordered regions. (*D*) Three-bond $H_N$-$H_\alpha$ scalar coupling constants for $IC_{88}$ calculated using cross-peak/diagonal peak intensity ratios from a 3D HNHA experiment. The predominance of coupling values between 6 and 8 Hz for residues 30–88 is consistent with those residues being dynamic or disordered. (*E*) One-bond H-N residual dipolar couplings measured using an $IC_{88}$ sample in a stretched gel. The cluster of positive values for residues 1–30 is consistent with this region forming a stable $\alpha$-helix. (*F*) NOE contact map for $IC_{88}$ after using ARIA to assign NOE peaks. Most of the NOE contacts are either intraresidue or sequential ($i \pm 1$). The pattern of medium-range contacts ($i \pm 3$ and $i \pm 4$) observed in the SAH region (residues 1–30) is consistent with $\alpha$-helical secondary structure. (*G*) $IC_{88}$ structural ensemble calculated using a combination of NOE restraints, scalar coupling-based torsion angle restraints, residual dipolar couplings, and chemical shift-based torsion angle restraints. The 20 lowest-energy structures from 50 structures calculated in the final iteration of the ARIA protocol were then refined in explicit water and used to generate the ensemble. The N-terminus is colored blue, whereas the C-terminus (residue 88) is colored red. For purposes of clarity, only 10 of the 20 structures used to define the structural ensemble are shown here, with residues 5–25 aligned. To see this figure in color, go online.

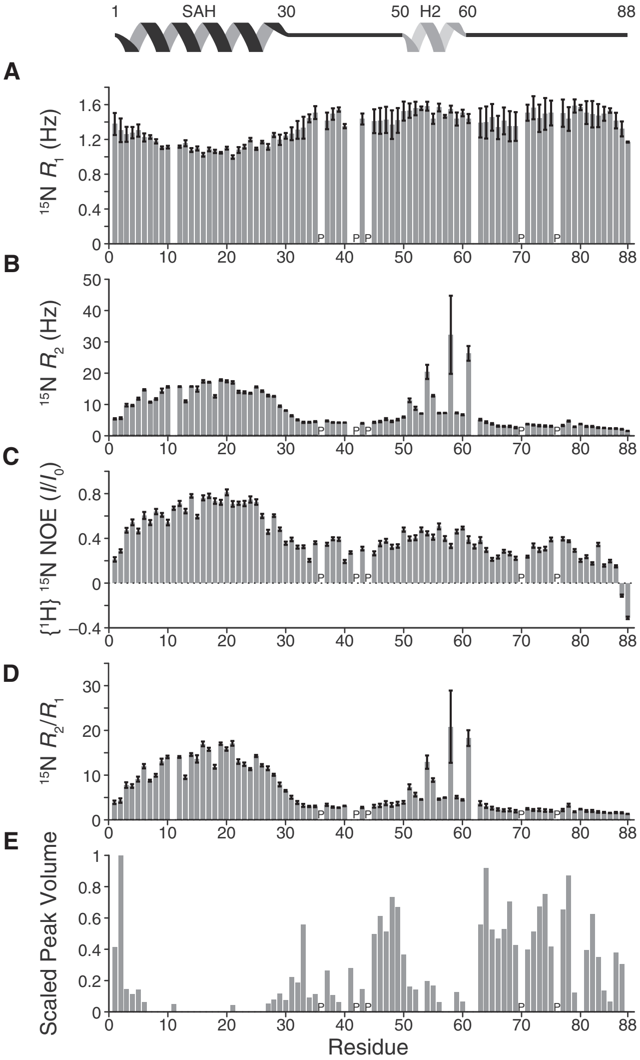


FIGURE 5 NMR relaxation and chemical exchange data for $IC_{88}$. $IC_{88}$ NMR relaxation data acquired at 800 MHz for $^1$H include (*A*) $^{15}$N longitudinal relaxation rates, (*B*) $^{15}$N transverse relaxation rates, and (*C*) {$^1$H} $^{15}$N heteronuclear NOE enhancements. The location of proline residues, which do not provide a signal in these experiments, are indicated by "P." The ratio of the transverse and the longitudinal relaxation rates ($R_2/R_1$) shown in (*D*) is expected to be close to unity for portions of $IC_{88}$ that have a short correlation time (i.e., regions that are highly dynamic), whereas larger values for this ratio are observed for more rigid regions. (*E*) Chemical exchange data for $IC_{88}$ measured using CLEANEX-PM; positive values indicate amide protons that exchange with the solvent, whereas zero or near-zero values indicate amide protons that are protected from exchange. The stretches of near-zero values for residues 3–30 and 50–60 are consistent with $\alpha$-helical secondary structures in these regions. For (*A*) and (*B*) the error bars are the standard errors of the $R_1$ and $R_2$ values derived from fitting the data; for (*C*) the error bars are from propagating the root-mean-square deviations of the noise floors of the saturated and control spectra; for (*D*) the error bars are derived from propagating the errors in the $R_1$ and $R_2$ values.

regions of proteins because of the increase in the effective correlation time. For $IC_{88}$, the decrease in $R_1$ (Fig. 5 *A*) and the increase in $R_2$ (Fig. 5 *B*) for residues 1–30 are indicative of this region having a high structural propensity relative to the rest of the protein. For residues 50–60, a smaller increase in $R_2$ (Fig. 5 *B*) is observed compared with residues 1–30. The large $R_2$-values for 54E, 58D, and 61L (all of which are on the same face of the H2 helix) are most likely due to conformational exchange on the micro- to millisecond timescale. {$^1$H}-$^{15}$N heteronuclear NOE values (Fig. 5 *C*) are greatest for residues 1–30 and are somewhat increased for residues 50–60 in comparison with the more dynamic parts of the protein. The information conveyed by the $R_1$- and $R_2$-values is highlighted by the ratio of these rates (Fig. 5 *D*), as large values indicate less flexible (more structured) regions and values approaching unity indicate dynamic regions. Taken together, the NMR dynamics data support a stable $\alpha$-helix for the SAH region and a nascent helix for the H2 region.

Measurement of chemical exchange between the solvent and amide protons made using CLEANEX-PM mixing provides insight into the accessibility of amide protons: amide protons involved in stable hydrogen bonds exchange more slowly with the solvent and have peaks of lower intensity than protons that are not protected from exchange. The stretches of near-zero values for residues 3–30 and 51–62 in the chemical exchange data shown in Fig. 5 *E* are consistent with these regions being involved in $\alpha$-helices, and the nonzero values for amides in other parts of the protein indicate that these regions are relatively disordered.

## Structural characterization of CT p150$^{Glued}$ CC1B

Studies of mammalian p150$^{Glued}$ have shown that the site at which the IC binds is located within the CC1B region (Fig. 1 *C*; (9)). Alignment of human, *Drosophila*, yeast, and CT p150$^{Glued}$ sequences and analysis of the coiled-coil propensity of these sequences (Fig. S1) was used to identify residues 478–680 of CT p150$^{Glued}$ as the CC1B region, which in the following sections we will refer to as $p150_{ABC}$ (Fig. 1 *D*). Given that the region of the IC that binds p150$^{Glued}$ has fewer than 80 amino acids, we presumed that only a portion of the ~200 amino acid CC1B region is actually involved in IC binding. To test this hypothesis, we engineered a series of shorter constructs that contained only portions of the p150$^{Glued}$ CC1B region. Analysis of $p150_{ABC}$ using COILS (47) with a window width of 21 residues revealed that the coiled-coil contains two short stretches of decreased coiled-coil propensity (Fig. 1 *D*). This prediction was used to subdivide $p150_{ABC}$ into three smaller regions (A, B, and C), which formed the basis of the $p150_{AB}$, $p150_{BC}$, $p150_{A}$, $p150_{B}$, and $p150_{C}$ constructs.

CD spectra for the CT p150$^{Glued}$ constructs (Fig. 6 *A*) show that the full-length $p150_{ABC}$ construct is largely $\alpha$-helical but that some of the shorter constructs have less secondary structure. Analysis of the MRE at 222 nm reveals that $p150_{ABC}$, $p150_{AB}$, and $p150_{A}$ are largely helical with fractional helicities of 80, 90, and 70%, respectively. The remaining constructs have a greater mixture of disorder and helix, with a fractional helicity of ~60% for $p150_{BC}$ and

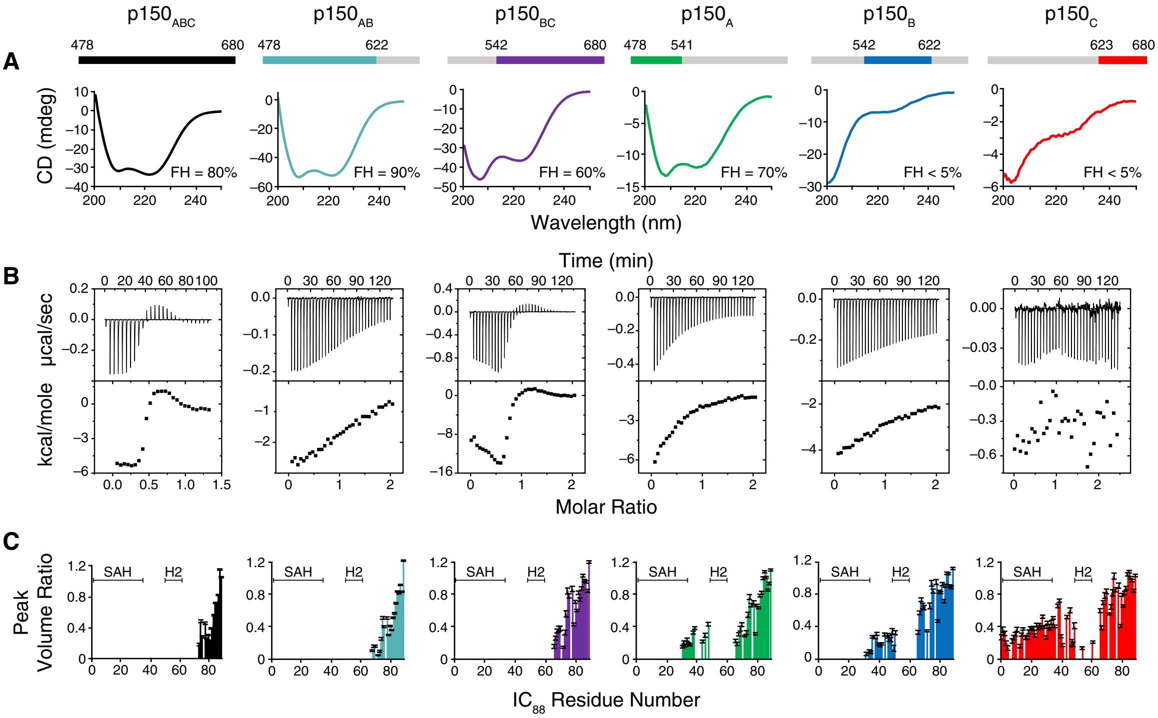


FIGURE 6 Identification of minimal sequence requirements for folding and binding using CD, ITC, and NMR titrations. (*A*) CD spectra at 25°C for $p150_{ABC}$, $p150_{AB}$, $p150_{BC}$, $p150_{A}$, $p150_{B}$, and $p150_{C}$; FH indicates the fractional helicity estimated at 222 nm using Eq. 1. The colored bars above the CD spectra graphically represent the portion of the CC1B region (residues 478–680) used for each construct. (*B*) ITC thermograms (*top*) and binding isotherms (*bottom*) from titrations of $p150_{ABC}$, $p150_{AB}$, $p150_{BC}$, $p150_{A}$, $p150_{B}$, and $p150_{C}$ with $IC_{88}$. (*C*) $^{1}H$-$^{15}N$ spectra peak volume ratios from titrations of $^{15}N$-labeled $IC_{88}$ with unlabeled $p150_{ABC}$, $p150_{AB}$, $p150_{BC}$, $p150_{A}$, $p150_{B}$, and $p150_{C}$. All data were acquired at 800 MHz for $^{1}H$. Samples for $p150_{ABC}$, $p150_{A}$, and $p150_{C}$ were at an IC to ligand molar ratio of ~1:1. Samples for $p150_{AB}$, $p150_{BC}$, and $p150_{B}$ were at an IC to ligand molar ratio of 1:1.2. To see this figure in color, go online.

less than 5% (i.e., mostly disordered) for $p150_{B}$ and $p150_{C}$. These results roughly match the predicted coiled-coil propensities for these regions of CC1B (Fig. 1 *D*), which suggests that the "A" region of $p150^{Glued}$ helps to stabilize the coiled-coil dimer. Therefore, constructs lacking this region tend to fray ($p150_{BC}$) and, if shortened even more, almost completely unravel ($p150_{B}$ and $p150_{C}$).

We expressed $^{15}N$- and/or $^{13}C/^{15}N$-labeled $p150_{AB}$, $p150_{B}$, and $p150_{BC}$ samples for NMR spectroscopy. $^{1}H$-$^{15}N$ correlation spectra of $p150_{AB}$ and $p150_{BC}$ (Fig. S2, *left* and *right spectra*) showed far fewer peaks than what we would expect based on the number of residues. One possible explanation is that the $p150_{AB}$ and $p150_{BC}$ constructs form long coiled-coils that tumble anisotropically. Because the N-H bond vectors for coiled-coils are nearly parallel to the long axis of the coiled-coil structure, the effective correlation time for nuclei in the folded regions of these proteins is relatively long, resulting in rapid relaxation and peak broadening (48). For both $p150_{AB}$ and $p150_{BC}$, the peaks observed in the spectra are assigned to amides at the dynamic N- and C-termini of these constructs. For $p150_{AB}$, amide peaks for only the first ~4 and last ~20 amino acids are observable, whereas in the case of $p150_{BC}$, a larger number of peaks (the first ~25 and the last ~15 amino acids) appear in the $^{1}H$-$^{15}N$ spectrum, a result that correlates with the greater disorder and lower helicity observed in the CD spectrum of $p150_{BC}$ (see Fig. 6 *A*).

Peaks from residues 480–600 for the $p150_{AB}$ construct and 570–665 for the $p150_{BC}$ construct are not readily observed, implying that these regions form the coiled-coil structure. As these regions overlap, this suggests that the reason that we see peaks for the C-terminal amides of $p150_{AB}$ and for the N-terminal amides of $p150_{BC}$ is because the coiled-coil structure has been destabilized at the termini of the constructs and not because these residues are intrinsically disordered in the full-length construct. This implies that in the $p150_{ABC}$ construct we would only observe peaks from the disordered ~4 amino acids at the N-terminus and ~15 amino acids at the C-terminus.

Attempts were made to further study the $p150_{BC}$ construct by generating a $^{2}H/^{13}C/^{15}N$-labeled sample. Although many additional peaks with greater chemical shift dispersion (data not shown) were observed in the $^{1}H$-$^{15}N$ TROSY spectrum for this deuterated sample, these peaks were relatively broad and not observable in 3D backbone assignment experiments. Because $p150_{BC}$ is structurally unstable at higher temperatures, increasing the sample temperature to narrow the peaks was not successful.

NMR spectra of the shorter p150$_B$ construct were readily assigned using TROSY-based 3D backbone experiments. Analysis of the backbone chemical shifts indicated that all of p150$_B$ was disordered rather than $\alpha$-helical, a result that is consistent with the degree of disorder predicted from its CD spectrum (Fig. 6 *A*). Interestingly, a second set of peaks for p150$_B$ was observed with somewhat greater chemical shift dispersion (labeled in *red*, Fig. S2, *center*), suggesting that this minor conformation has more helical propensity. $^{15}$N ZZ-exchange spectra established the timescale for interconversion between these two forms to be on the order of 100 ms (data not shown). This more structured minor conformer of p150$_B$ explains why an interaction was observed between p150$_B$ and IC$_{88}$ in the NMR titration and ITC data (see following section) despite the major conformer of p150$_B$ being disordered.

## Characterization of the IC-p150$^{Glued}$ interaction

Previous ITC studies on rat IC and p150$^{Glued}$ (13) reported a dissociation constant ($K_d$) on the order of 10 $\mu$M (Table S1), whereas our results for the CT versions of these proteins revealed binding that is apparently stronger and substantially more complex than in the mammalian system (Fig. 7 *A*). In comparison, the thermogram for IC$_{35}$ binding to p150$_{ABC}$ (Fig. 7 *A*, *middle*) showed that removing H2 and the disordered linker regions surrounding it (residues 30–50 and 60–88) drastically diminished the second step. These differences suggest that some combination of the H2 and linker regions is important for strong interactions with p150$^{Glued}$. Fitting results for the ITC curves in Fig. 7 *A* are provided in Table S1, but we caution that the numerical values are model dependent and should be treated with the usual caveats for interpreting multiphasic ITC fits (49).

We also characterized binding using SV-AUC. SV-AUC experiments clearly show a peak for the p150$_{ABC}$-IC$_{88}$ complex that is significantly shifted from the peak for free p150$_{ABC}$ (Fig. 7 *C*), consistent with the tight binding of the complex observed with ITC as well as the existence of only a single oligomeric state.

To probe whether the contribution to the IC-p150$^{Glued}$ interaction by H2 is sequence specific, we generated a chimeric IC$_{88}$ construct (Fig. 7 *B*), referred to in the following as IC$_{88C}$, in which 11 amino acids of the CT H2 sequence were replaced by 11 (out of 15) amino acids from the rat H2 sequence. Although in rat IC, the H2 region is an ordered helix, this helix is destabilized in the chimeric construct; the CD spectrum for IC$_{88C}$ is nearly identical to the spectrum for the native IC$_{88}$ construct (data not shown), indicating that the level of helicity for the H2 region is the same for both IC$_{88}$ and IC$_{88C}$. As shown to the right of Fig. 7 *A*, the thermogram for IC$_{88C}$ binding to p150$_{ABC}$ is almost identical to the thermogram for the native IC$_{88}$ sequence (Fig. 7 *A*, *left*), a result that implies that the interaction between H2 and p150$^{Glued}$ may depend on the charge distribution or the helical propensity of H2 rather than its specific sequence. Although this binding behavior could also be due to the disordered linker regions flanking H2, our NMR data strongly suggest that it is H2, and not the linkers, that plays the more significant role (as explained in the following section and shown in Fig. 7 *D*).

ITC was also used to broadly determine the regions necessary for binding between IC$_{88}$ and the fragments of p150$_{ABC}$ (Fig. 6 *B*). p150$_{BC}$ is the only construct that bound IC$_{88}$ in a similar two-step manner as p150$_{ABC}$, whereas for p150$_{AB}$ and p150$_A$, only weak binding was observed (Table S1). Further isolation of the region of p150$_{BC}$ involved in binding IC was not successful because shortening the sequence to form the p150$_B$ and p150$_C$ constructs completely destabilized their coiled-coil structures, leading to weak (p150$_B$) or negligible (p150$_C$) binding at the concentrations used for ITC. The persistence of a weak interaction between IC$_{88}$ and p150$_B$ despite the disorder of the p150$_B$ construct is most likely due to the presence of the minor, partially folded conformer that we observed in the NMR spectra (Fig. S2, *middle*).

## H2 of the IC directly interacts with p150$^{Glued}$

NMR spectra of $^{15}$N-labeled IC$_{88}$ with excess of the p150$^{Glued}$ constructs identify the regions of CT IC involved in binding (Fig. 6 *C*). Peak volume ratios close to unity indicate regions of IC$_{88}$ that are mostly unaffected by the addition of the p150$^{Glued}$ constructs, whereas decreased or zero values indicate regions that are either directly involved in binding or are adjacent to regions that bind. Upon addition of p150$_{ABC}$, p150$_{AB}$, and p150$_{BC}$, peaks from IC$_{88}$ amides in both the SAH and H2 regions, as well as the disordered linker between them, completely disappear in the NMR spectra. Addition of the smaller (and more weakly binding) p150$_A$ and p150$_B$ constructs show similar disappearances of peaks, except for the interhelix linker region, which shows only reduced peak intensity. For samples that we tested with a substoichiometric ratio of the binding partners (i.e., an excess of IC$_{88}$), the disappearance of the peaks was intermediate between the unbound and fully bound samples, indicating that the binding process occurs on a timescale that is fast compared with the NMR experiment (data not shown). Furthermore, although they have different binding affinities, the protein concentrations in the NMR samples are significantly above the $K_d$ for all of them, and therefore, we would not expect to see a difference in the spectra of the bound proteins. Intriguingly, whereas ITC showed no measurable interaction between p150$_C$ and IC$_{88}$, and CD showed that p150$_C$ is almost completely disordered, the higher protein concentration used for the NMR experiments revealed a weak interaction between p150$_C$ and IC$_{88}$, resulting in the disappearance of peaks from the H2 region, but not the SAH region. This result suggests that it is the H2 region, and not the neighboring linkers, that interacts with

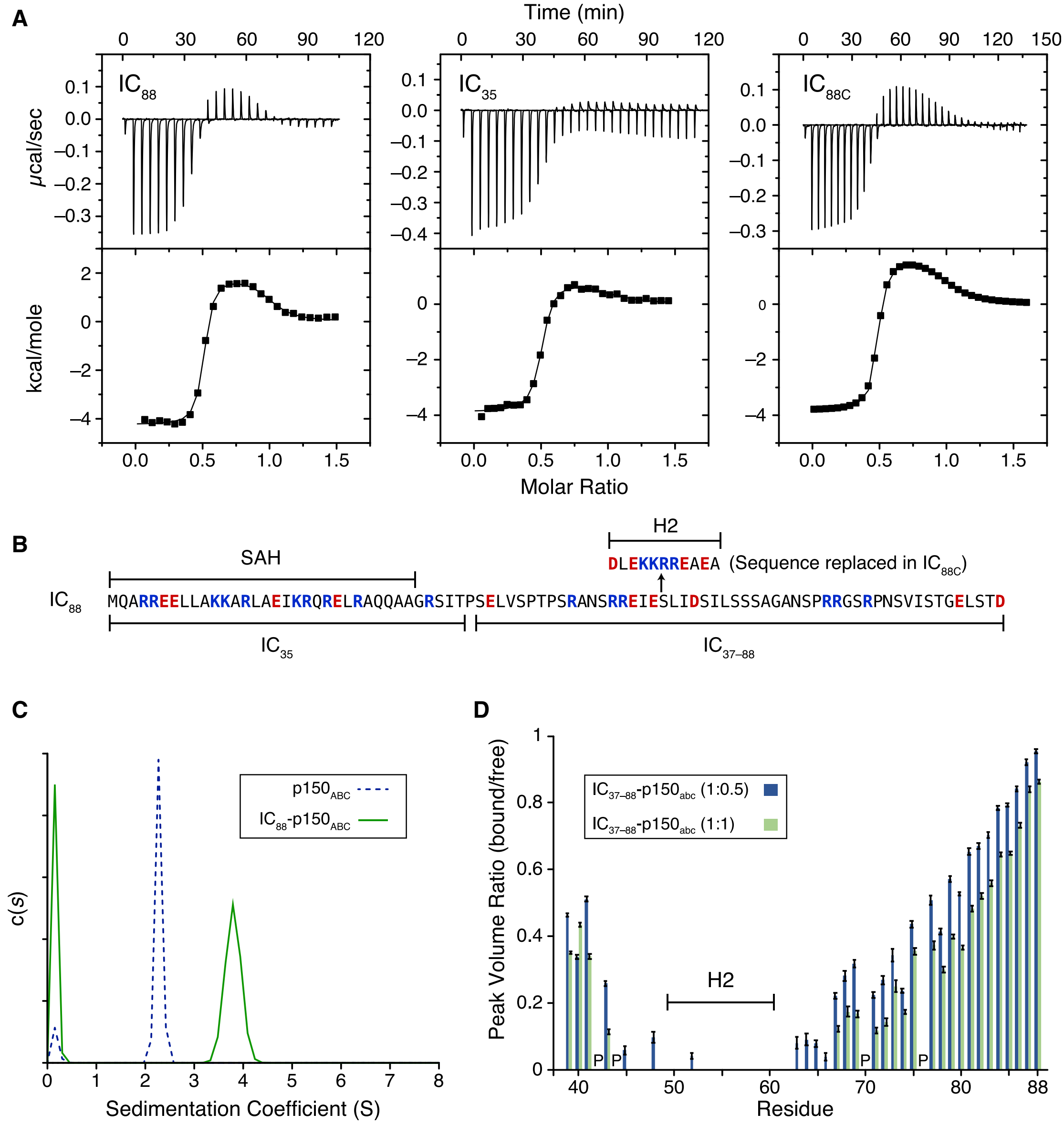


FIGURE 7 Interactions of CT IC constructs with p150$_{ABC}$. (*A*) Representative thermograms (*top*) and binding isotherms (*bottom*) from ITC titrations of IC$_{88}$ (*left*), IC$_{35}$ (*middle*), and IC$_{88C}$ (*right*) with p150$_{ABC}$ collected at 25°C (pH 7.5). (*B*) The sequence for IC$_{88}$ illustrating which amino acids in the H2 region are replaced in the IC$_{88C}$ construct. The IC$_{35}$ construct consists of the N-terminal 35 amino acids of IC$_{88}$, and the IC$_{37-88}$ construct consists of the C-terminal 52 amino acids of IC$_{88}$. Basic residues (K, R) and acidic residues (D, E) are indicated in blue and red, respectively. (*C*) SV-AUC of free p150$_{ABC}$ (80 $\mu$M, *blue dashed line*) and the SEC-purified IC$_{88}$-p150$_{ABC}$ complex (60 $\mu$M, *green solid line*) shows that IC$_{88}$-p150$_{ABC}$ forms a stable complex with an apparent single oligomeric state. No data are collected for IC$_{88}$ for comparison because it has no absorbance at 280 nm. (*D*) $^{1}$H-$^{15}$N spectra peak volume ratios from titrations of 240 $\mu$M $^{15}$N-labeled IC$_{37-88}$ with unlabeled p150$_{ABC}$ acquired at 600 MHz for $^{1}$H. Samples were prepared with a substoichiometric (1:0.5, *blue*) or a stoichiometric (1:1, *green*) ratio of IC to p150$_{ABC}$. The location of the four prolines in IC$_{37-88}$ are indicated by "P." To see this figure in color, go online.

p150$^{Glued}$ and that this interaction is direct rather than mediated by the SAH region. Whether the H2 region of IC$_{88}$ interacts with unfolded p150$_{C}$ or with a more folded minor conformation populated at NMR concentrations cannot be established with the data that we have at present, but the fact that peaks for the H2 region disappear in all of the NMR titrations suggests that it plays an important role in binding p150$^{Glued}$.

To confirm that the H2 region does actually interact directly with p150$^{Glued}$ and not in an indirect manner dependent on the SAH region, we investigated a CT IC construct consisting of only residues 37–88 (IC$_{37-88}$). This construct contains the H2 region and the neighboring linker regions but does not contain the SAH region. Except for residues near the N-terminus, the peaks in the spectra for this construct were at identical chemical shifts as for IC$_{88}$, indicating that the structure of the H2 and linker regions was unchanged relative to the larger construct. An NMR titration of $^{15}$N-labeled $IC_{37–88}$ with unlabeled $p150_{ABC}$ (Fig. 7 *D*) shows a complete loss of signal for the H2 region and that signal for the linker regions diminishes depending on how close the linker residues are to the H2 region. This confirms that the H2 region directly interacts with $p150_{ABC}$ and strongly suggests that reduced peak volumes for residues in the linker regions (as seen in Figs. 6 *C* and 7 *D*) are a result of the reduced mobility of these residues due to their proximity to regions that bind $p150_{ABC}$ rather than due to a direct interaction between these residues and $p150_{ABC}$.

To elucidate the importance of H2 to the stability of the $p150^{Glued}$-IC complex, we performed temperature dependence experiments on $IC_{35}$, $IC_{88}$, and $p150_{ABC}$ as well as complexes of these proteins (Fig. 8). The CD spectra for individual proteins show a rapid decrease in helicity as the temperature is increased (222 nm becomes less negative; Fig. 8, *A–D*). On the other hand, complexes of $p150_{ABC}$ with $IC_{35}$ or $IC_{88}$ have increased thermal stability (Fig. 8, *E* and *F*), indicating that these proteins are co-stabilizing, with the $IC_{88}$-$p150_{ABC}$ complex, which contains the H2 region of IC, showing higher stability. A comparison of the temperature-dependent 222 nm signal between the complexes and free proteins (Fig. 8 *A*) clearly illustrates that the $IC_{88}$-$p150_{ABC}$ complex has the highest stability. Although these CD data alone cannot rule out a role for the disordered linker regions of $IC_{88}$ in stabilizing the complex, together with the NMR data, they strongly suggest that the H2 region promotes the formation of a more stable IC-$p150^{Glued}$ complex.

## DISCUSSION

The 300 amino acid N-terminal domain of the IC is primarily disordered except for a short SAH within residues 1–40, which is the primary binding site to dynactin $p150^{Glued}$. Linked to the SAH by a short, disordered linker is helix H2, which in rat IC is fully ordered but does not contribute to binding of $p150^{Glued}$ (13), whereas in yeast and *Drosophila* ICs (8,21), H2 is disordered, and its removal drastically diminishes IC binding to $p150^{Glued}$. In this work, we use the IC from CT as a tractable model for ICs with disordered H2 regions to establish the importance of H2 in IC-$p150^{Glued}$ interactions.

### The structural propensity of H2 varies among species

Across a wide range of eukaryotes, the N-terminal region of IC has a well-defined helical region (SAH) and a secondary helix (H2) that varies in structural propensity. Although the

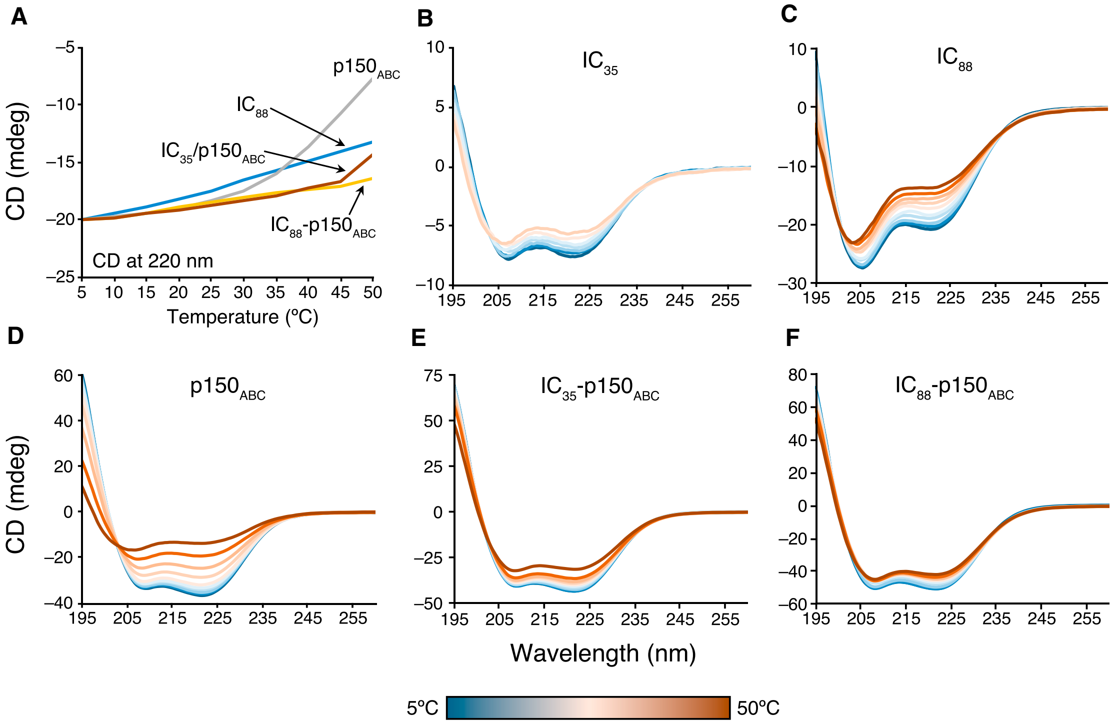


FIGURE 8 Complex-dependent secondary structure stability detected using temperature-dependent CD. (*A*) CD signal (in millidegrees, mdeg) as a function of temperature at 220 nm for $IC_{88}$, $p150_{ABC}$, $IC_{88}$ bound with $p150_{ABC}$, and $IC_{35}$ bound with $p150_{ABC}$. The curves on this graph are scaled so that they align at low temperatures. (*B–F*) Temperature-dependent CD spectra of (*B*) $IC_{35}$, (*C*) $IC_{88}$, (*D*) $p150_{ABC}$, (*E*) $IC_{35}$-$p150_{ABC}$, and (*F*) $IC_{88}$-$p150_{ABC}$ collected at temperatures between 5 and 50°C. To see this figure in color, go online.

location, sequence, and size of H2 varies among non-chordate species, structural propensities for this region based on predictions and NMR data are consistent with a nascent helix. Along with this, binding to p150$^{Glued}$ in rat is largely unaffected by the removal of the folded H2 region (13), whereas binding to p150$^{Glued}$ drastically weakens in *Drosophila*, yeast (8,21), and, as shown here, CT when the nascent H2 region is removed. Binding to p150$_{ABC}$ is significantly stronger for IC$_{88}$ than it is for the IC$_{35}$ construct, which only contains the SAH region (Fig. 7 *A*), suggesting that disorder in H2 is species dependent and that this disorder is correlated with its contribution to binding affinity.

## H2 is involved in a direct interaction with p150$^{Glued}$

Whereas previous work with *Drosophila* and yeast ICs indicated that H2 is somehow important for tighter binding between the IC and p150$^{Glued}$, it was never clear whether H2 was involved in direct interactions with p150$^{Glued}$ or packed against the SAH region (8,21), as shown in the binding model of Fig. 9 *A*. The existence of a p150$^{Glued}$ construct (p150$_C$) that binds the H2 region, but not the SAH region, provides for the first time, to our knowledge, evidence that in CT IC, the H2 region binds directly to p150$^{Glued}$, leading us to propose a new model for IC-p150$^{Glued}$ binding (Fig. 9 *B*).

Interestingly, the H2 region is affected in titrations with all six p150$^{Glued}$ constructs (Fig. 6 *C*), despite regions "A," "B," and "C" having no obvious sequence homology. Likewise, replacing the CT H2 region with the rat H2 region does not diminish binding to p150$^{Glued}$ (Fig. 7 *A*, *right*). Taken together, these results indicate that the interaction between H2 and p150$^{Glued}$ is not sequence specific. Most likely, this nonspecific interaction is electrostatic in nature because over 40% of both the CC1B region of p150$^{Glued}$ and the H2 region of CT and rat IC consists of charged residues (D, E, L, K). The dominance of electrostatic interactions in driving CT IC-p150$^{Glued}$ binding is consistent with the observation that mammalian IC-p150$^{Glued}$ complex formation is sensitive to salt, temperature, and pH (9). Interestingly, despite the nonspecific nature of the interaction between the H2 region and p150$^{Glued}$, the H2 region is not completely promiscuous in binding to coiled coils; when we titrated $^{15}$N-labeled IC$_{88}$ with Swa$_{DIMER}$ (205–275), a stable 71-residue parallel homodimeric coiled-coil with a similar percentage of charged residues (50), no change in the IC$_{88}$ NMR spectrum was observed (data not shown). This shows that there is some degree of specificity to the "nonspecific" interaction between the H2 region of CT IC and the various p150$^{Glued}$ constructs that we have studied.

One possible model for IC-p150$^{Glued}$ binding is that the H2 region is involved in scanning the CC1B region via relatively nonspecific electrostatic interactions (Fig. 9 *C*) until the SAH region locks into a more specific interaction with p150$^{Glued}$ (Fig. 9 *B*) in a manner analogous to how DNA binding proteins slide or hop along the DNA double helix until locking onto a specific binding site (51,52). We speculate that the nascent H2 helix provides conformational flexibility that supports this function, whereas other organisms (chordates) have developed or co-opted an alternative mechanism for promoting p150$^{Glued}$ binding that does not depend on the H2 region having a low structural propensity.

## Regions "B" and "C" of p150$^{Glued}$ are the main locus of IC binding

The majority of the binding interface is located within the "B" and "C" regions, as shown by the ability of the p150$_{BC}$ construct to replicate almost the entirety of the

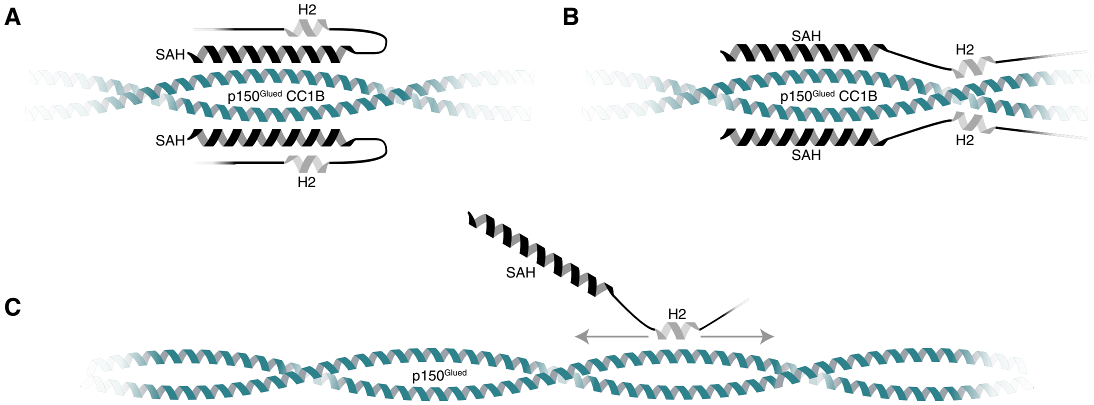


FIGURE 9 Models for the binding interaction between IC and p150$^{Glued}$. (*A*) Current model for IC-p150$^{Glued}$ binding, in which only the SAH region interacts directly with the p150$^{Glued}$ CC1B region (13,21). (*B*) New model based on our NMR results for CT IC-p150$^{Glued}$ binding, in which both the SAH and H2 regions directly interact with p150$^{Glued}$. For clarity, only a small portion of the ~200-amino-acid-long p150$^{Glued}$ CC1B region is shown. (*C*) The observation of nonspecific electrostatic interactions between H2 and p150$^{Glued}$ CC1B suggest a speculative model in which H2 is able to interact nonspecifically with the CC1B region of p150$^{Glued}$ until the SAH region locks to a specific binding site, resulting in the configuration modeled in (*B*). To see this figure in color, go online.

IC-p150$_{ABC}$ thermogram (Fig. 6 *B*). The main difference, an initial increase in heat magnitudes seen in the p150$_{BC}$ thermogram, may represent the stabilization of the coiled-coil structure of p150$_{BC}$ in the final complex. We hypothesize that the "A" region of p150$^{Glued}$ is essential for stabilizing the coiled-coil structure of the CC1B region, as the CD data show that it is the only one of the shorter p150$^{Glued}$ constructs that is helical by itself, a result that is consistent with the coiled-coil propensity predictions for the regions of p150$^{Glued}$ shown in Fig. 1 *D*.

The IC binding site on p150$^{Glued}$ is unlikely to include the entire CC1B region, but evidence points to a majority of the region being required for strong binding in CT. The "BC" region is sufficient to replicate the binding of the full construct, whereas p150$_A$ and p150$_{AB}$ only show weak binding with CT IC even though they are fully folded. This suggests that although it is likely that there are a number of nonspecific electrostatic interactions that occur before formation of a stable complex, the B and C regions contain the majority of the binding sequence that is essential for formation of a tight IC-CC1B complex.

## CONCLUSIONS

The N-terminal region of the CT IC contains a helical SAH region, consistent with all ICs studied to date, as well as a nascent helix in the H2 region. Our predictions for the structural propensity of H2, along with experimental structural characterization of ICs, suggests that the extent of helicity is species dependent and that low helical propensity is correlated with its importance in promoting tight complex formation with the CC1B region of dynactin p150$^{Glued}$. Our results suggest that the CT IC binding site on p150$^{Glued}$ is within regions "B" and "C" (and most likely in the range of residues 550–650, which roughly corresponds to residues 410–510 in human p150$^{Glued}$). The intriguing observation that the H2 region in CT interacts with all subsections of CC1B suggests that the disorder in H2 may form initial nonspecific contacts with p150$^{Glued}$ before the SAH region locks into a specific interaction with CC1B. Organisms with folded H2 regions of IC, however, could have evolved alternative mechanisms to initiate and stabilize IC-p150$^{Glued}$ binding. Indeed, for rat IC, which has a folded H2, its interaction with p150$^{Glued}$ is regulated by phosphorylation (13,53), making it tantalizing to suggest that disorder in H2 underlies regulation of binding in the absence of phosphorylation.

One intriguing aspect of the CT IC interactions with p150$^{Glued}$ is the uncharacteristically high affinity observed when compared with mammalian IC-p150$^{Glued}$ interactions. In the latter, even when in the full complex, dynactin and dynein required an adaptor such as BicD2, Rab11-FIP3, Spindly, or Hook3 to enhance the stability of the complex that dynein forms with dynactin, leading to highly processive motility (54). In contrast, yeast dynein shows directional motion on its own, and its processivity is enhanced upon binding dynactin (55). Based on the very tight binding observed in CT, we propose that adaptors such as BicD2, which add another layer of regulation in the mammalian system, may not be needed when the disordered H2 is involved in a direct interaction with p150$^{Glued}$.

## SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.07.023.

## AUTHOR CONTRIBUTIONS

N.M.L., S.S., N.E.J., and E.B. designed the study. N.M.L., S.S., N.E.J., and K.J. performed the research and analyzed the data. N.M.L., S.S., and E.B. wrote the manuscript with input from all authors. E.B. oversaw the study.

## ACKNOWLEDGMENTS

We thank Patrick Reardon for his assistance with collecting the NMR and AUC data.

We acknowledge support from the National Science Foundation (Award 1617019). The Oregon State University NMR Facility is funded in part by the National Institutes of Health (HEI Grant 1S10OD018518) and by the M. J. Murdock Charitable Trust (Grant 2014162). The Bruker Avance NEO 600 NMR spectrometer was purchased with support from the National Science Foundation (Award 1917696) and the M. J. Murdock Charitable Trust (Grant 201811283).

## SUPPORTING CITATIONS

References (56,57) can be found in the Supporting Material.

## REFERENCES

1. Kardon, J. R., and R. D. Vale. 2009. Regulators of the cytoplasmic dynein motor. *Nat. Rev. Mol. Cell Biol.* 10:854–865.
2. Vallee, R. B., R. J. McKenney, and K. M. Ori-McKenney. 2012. Multiple modes of cytoplasmic dynein regulation. *Nat. Cell Biol.* 14:224–230.
3. Reck-Peterson, S. L., W. B. Redwine, …, A. P. Carter. 2018. The cytoplasmic dynein transport machinery and its many cargoes. *Nat. Rev. Mol. Cell Biol.* 19:382–398.
4. Kini, A. R., and C. A. Collins. 2001. Modulation of cytoplasmic dynein ATPase activity by the accessory subunits. *Cell Motil. Cytoskeleton.* 48:52–60.
5. Makokha, M., M. Hare, …, E. Barbar. 2002. Interactions of cytoplasmic dynein light chains Tctex-1 and LC8 with the intermediate chain IC74. *Biochemistry.* 41:4302–4311.
6. Nyarko, A., and E. Barbar. 2011. Light chain-dependent self-association of dynein intermediate chain. *J. Biol. Chem.* 286:1556–1566.
7. Schroer, T. A. 2004. Dynactin. *Annu. Rev. Cell Dev. Biol.* 20:759–779.
8. Morgan, J. L., Y. Song, and E. Barbar. 2011. Structural dynamics and multiregion interactions in dynein-dynactin recognition. *J. Biol. Chem.* 286:39349–39359.
9. Siglin, A. E., S. Sun, …, J. C. Williams. 2013. Dynein and dynactin leverage their bivalent character to form a high-affinity interaction. *PLoS One.* 8:e59453.

10. Tripathy, S. K., S. J. Weil, …, S. P. Gross. 2014. Autoregulatory mechanism for dynactin control of processive and diffusive dynein transport. *Nat. Cell Biol.* 16:1192–1201.
11. Moore, J. K., D. Sept, and J. A. Cooper. 2009. Neurodegeneration mutations in dynactin impair dynein-dependent nuclear migration. *Proc. Natl. Acad. Sci. USA.* 106:5147–5152.
12. Eschbach, J., and L. Dupuis. 2011. Cytoplasmic dynein in neurodegeneration. *Pharmacol. Ther.* 130:348–363.
13. Jie, J., F. Löhr, and E. Barbar. 2017. Dynein binding of competitive regulators dynactin and NudE involves novel interplay between phosphorylation site and disordered spliced linkers. *Structure.* 25:421–433.
14. Urnavicius, L., K. Zhang, …, A. P. Carter. 2015. The structure of the dynactin complex and its interaction with dynein. *Science.* 347:1441–1446.
15. Urnavicius, L., C. K. Lau, …, A. P. Carter. 2018. Cryo-EM shows how dynactin recruits two dyneins for faster movement. *Nature.* 554:202–206.
16. Benison, G., A. Nyarko, and E. Barbar. 2006. Heteronuclear NMR identifies a nascent helix in intrinsically disordered dynein intermediate chain: implications for folding and dimerization. *J. Mol. Biol.* 362:1082–1093.
17. Tompa, P., and M. Fuxreiter. 2008. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* 33:2–8.
18. Clark, S. A., N. Jespersen, …, E. Barbar. 2015. Multivalent IDP assemblies: unique properties of LC8-associated, IDP duplex scaffolds. *FEBS Lett.* 589:2543–2551.
19. Lee, I.-G., M. A. Olenick, …, R. Dominguez. 2018. A conserved interaction of the dynein light intermediate chain with dynein-dynactin effectors necessary for processivity. *Nat. Commun.* 9:986.
20. Celestino, R., M. A. Henen, …, B. Vögeli. 2019. A transient helix in the disordered region of dynein light intermediate chain links the motor to structurally diverse adaptors for cargo transport. *PLoS Biol.* 17:e3000100.
21. Jie, J., F. Löhr, and E. Barbar. 2015. Interactions of yeast dynein with dynein light chain and dynactin: general implications for intrinsically disordered duplex scaffolds in multiprotein assemblies. *J. Biol. Chem.* 290:23863–23874.
22. Kellner, N., J. Schwarz, …, E. Hurt. 2016. Developing genetic tools to exploit Chaetomium thermophilum for biochemical analyses of eukaryotic macromolecular assemblies. *Sci. Rep.* 6:20937.
23. UniProt Consortium. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47:D506–D515.
24. Muñoz, V., and L. Serrano. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol.* 1:399–409.
25. Muñoz, V., and L. Serrano. 1995. Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol.* 245:275–296.
26. Muñoz, V., and L. Serrano. 1995. Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *J. Mol. Biol.* 245:297–308.
27. Muñoz, V., and L. Serrano. 1997. Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm-Bragg and Lifson-Roig formalisms. *Biopolymers.* 41:495–509.
28. Lacroix, E., A. R. Viguera, and L. Serrano. 1998. Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol.* 284:173–191.
29. Gibson, D. G., L. Young, …, H. O. Smith. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods.* 6:343–345.
30. Gibson, D. G., J. I. Glass, …, J. C. Venter. 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science.* 329:52–56.
31. Jansson, M., Y. C. Li, …, B. Nilsson. 1996. High-level production of uniformly $^{15}$N- and $^{13}$C-enriched fusion proteins in *Escherichia coli*. *J. Biomol. NMR.* 7:131–141.
32. Gasteiger, E., C. Hoogland, …, A. Bairoch. 2005. Protein identification and analysis tools on the ExPASy server. *In* The Proteomics Protocols Handbook. J. M. Walker, ed. Humana Press, pp. 571–607.
33. Anthis, N. J., and G. M. Clore. 2013. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci.* 22:851–858.
34. Wei, Y., A. A. Thyparambil, and R. A. Latour. 2014. Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230nm. *Biochim. Biophys. Acta.* 1844:2331–2337.
35. Schuck, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* 78:1606–1619.
36. Solyom, Z., M. Schwarten, …, B. Brutscher. 2013. BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J. Biomol. NMR.* 55:311–321.
37. Zhu, G., Y. Xia, …, K. H. Sze. 2000. Protein dynamics measurements by TROSY-based NMR experiments. *J. Magn. Reson.* 143:423–426.
38. Hwang, T. L., P. C. van Zijl, and S. Mori. 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.
39. Li, Y., N. L. Altorelli, …, A. G. Palmer, III. 2013. Mechanism of E-cadherin dimerization probed by NMR relaxation dispersion. *Proc. Natl. Acad. Sci. USA.* 110:16462–16467.
40. Delaglio, F., S. Grzesiek, …, A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 6:277–293.
41. Coggins, B. E., J. W. Werner-Allen, …, P. Zhou. 2012. Rapid protein global fold determination using ultrasparse sampling, high-dynamic range artifact suppression, and time-shared NOESY. *J. Am. Chem. Soc.* 134:18619–18630.
42. Vranken, W. F., W. Boucher, …, E. D. Laue. 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins.* 59:687–696.
43. Rieping, W., M. Habeck, …, M. Nilges. 2007. ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics.* 23:381–382.
44. Shen, Y., and A. Bax. 2013. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR.* 56:227–241.
45. Valafar, H., and J. H. Prestegard. 2004. REDCAT: a residual dipolar coupling analysis tool. *J. Magn. Reson.* 167:228–241.
46. Nilges, M., A. Bernard, …, W. Rieping. 2008. Accurate NMR structures through minimization of an extended hybrid energy. *Structure.* 16:1305–1312.
47. Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. *Science.* 252:1162–1164.
48. MacKay, J. P., G. L. Shaw, and G. F. King. 1996. Backbone dynamics of the c-Jun leucine zipper: $^{15}$N NMR relaxation studies. *Biochemistry.* 35:4867–4877.
49. Brautigam, C. A. 2015. Fitting two- and three-site binding models to isothermal titration calorimetric data. *Methods.* 76:124–136.
50. Kidane, A. I., Y. Song, …, E. Barbar. 2013. Structural features of LC8-induced self-association of swallow. *Biochemistry.* 52:6011–6020.

51. Halford, S. E., and J. F. Marko. 2004. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* 32:3040–3052.
52. Tafvizi, A., L. A. Mirny, and A. M. van Oijen. 2011. Dancing on DNA: kinetic aspects of search processes on DNA. *ChemPhysChem.* 12:1481–1489.
53. Vaughan, P. S., J. D. Leszyk, and K. T. Vaughan. 2001. Cytoplasmic dynein intermediate chain phosphorylation regulates binding to dynactin. *J. Biol. Chem.* 276:26171–26179.
54. McKenney, R. J., W. Huynh, …, R. D. Vale. 2014. Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes. *Science.* 345:337–341.
55. Kardon, J. R., S. L. Reck-Peterson, and R. D. Vale. 2009. Regulation of the processivity and intracellular localization of *Saccharomyces cerevisiae* dynein by dynactin. *Proc. Natl. Acad. Sci. USA.* 106:5669–5674.
56. Laskowski, R. A., J. A. Rullmannn, …, J. M. Thornton. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR.* 8:477–486.
57. Chen, V. B., W. B. Arendall, III, …, D. C. Richardson. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66:12–21.
58. Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.
59. McWilliam, H., W. Li, …, R. Lopez. 2013. Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* 41:W597–W600.
60. Li, W., A. Cowley, …, R. Lopez. 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 43:W580–W584.