



HAL
open science

When Order Meets Disorder: Modeling and Function of the Protein Interface in Fuzzy Complexes

Sophie Sacquin-Mora, Chantal Prevost

► **To cite this version:**

Sophie Sacquin-Mora, Chantal Prevost. When Order Meets Disorder: Modeling and Function of the Protein Interface in Fuzzy Complexes. *Biomolecules*, 2021, 11 (10), pp.1529. 10.3390/biom11101529 . hal-03389754

HAL Id: hal-03389754

<https://hal.science/hal-03389754>

Submitted on 21 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Review

When Order Meets Disorder: Modeling and Function of the Protein Interface in Fuzzy Complexes

Sophie Sacquin-Mora ^{1,2} and Chantal Prévost ^{1,2,*}

¹ CNRS, Laboratoire de Biochimie Théorique, UPR9080, Université de Paris, 13 Rue Pierre et Marie Curie, 75005 Paris, France; sacquin@ibpc.fr

² Institut de Biologie Physico-Chimique, Fondation Edmond de Rothschild, PSL Research University, 75006 Paris, France

* Correspondence: chantal.prevost@ibpc.fr

Abstract: The degree of proteins structural organization ranges from highly structured, compact folding to intrinsic disorder, where each degree of self-organization corresponds to specific functions: well-organized structural motifs in enzymes offer a proper environment for precisely positioned functional groups to participate in catalytic reactions; at the other end of the self-organization spectrum, intrinsically disordered proteins act as binding hubs via the formation of multiple, transient and often non-specific interactions. This review focusses on cases where structurally organized proteins or domains associate with highly disordered protein chains, leading to the formation of interfaces with varying degrees of fuzziness. We present a review of the computational methods developed to provide us with information on such fuzzy interfaces, and how they integrate experimental information. The discussion focusses on two specific cases, microtubules and homologous recombination nucleoprotein filaments, where a network of intrinsically disordered tails exerts regulatory function in recruiting partner macromolecules, proteins or DNA and tuning the atomic level association. Notably, we show how computational approaches such as molecular dynamics simulations can bring new knowledge to help bridging the gap between experimental analysis, that mostly concerns ensemble properties, and the behavior of individual disordered protein chains that contribute to regulation functions.

Keywords: intrinsically disordered proteins; molecular modeling; fuzzy complexes; disorder and function



Citation: Sacquin-Mora, S.; Prévost, C. When Order Meets Disorder: Modeling and Function of the Protein Interface in Fuzzy Complexes. *Biomolecules* **2021**, *11*, 1529. <https://doi.org/10.3390/biom11101529>

Academic Editors: Nathalie Sibille, Sonia Longhi and Carine Van Heijenoort

Received: 14 September 2021

Accepted: 12 October 2021

Published: 16 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the resolution of the first protein crystallographic structures some sixty years ago, and following Anfinsen's dogma [1], the assumption that protein function requires a well-defined structure has been a cornerstone of protein science. However, at the turn of the century the biological importance of intrinsically disordered proteins (IDPs), and of intrinsically disordered regions (IDRs) in proteins, became increasingly clear [2]. What was once described as the Dark Proteome [3,4], since disordered segments remained invisible in most structural approaches, progressively turned into a central element of protein activity [5,6], taking an important part in numerous cellular processes [7,8], and thus showing that function does not always rely on a well-defined structure. Nowadays, around 10% of the 10,000 structures deposited annually in the PDB comprise long disorder regions of at least 30 residues [9], providing structural biologists with a wealth of experimental data to work on.

Accounting for disorder in protein assemblies was also a slow process [10]. One of the first scenarios considered, coupled folding to binding, would limit the conformational diversity of the interacting partners to their unbound form. It took another decade to finally apprehend the concept of fuzzy complexes, where one or both partners in the protein interaction can present some structural ambiguity [11]. In our current knowledge

of disordered proteins, their binding modes span a continuum, ranging from disorder-to-order transitions, with a well-defined bound state, to disordered binding, with an also disordered bound state [12,13] (see Figure 1).

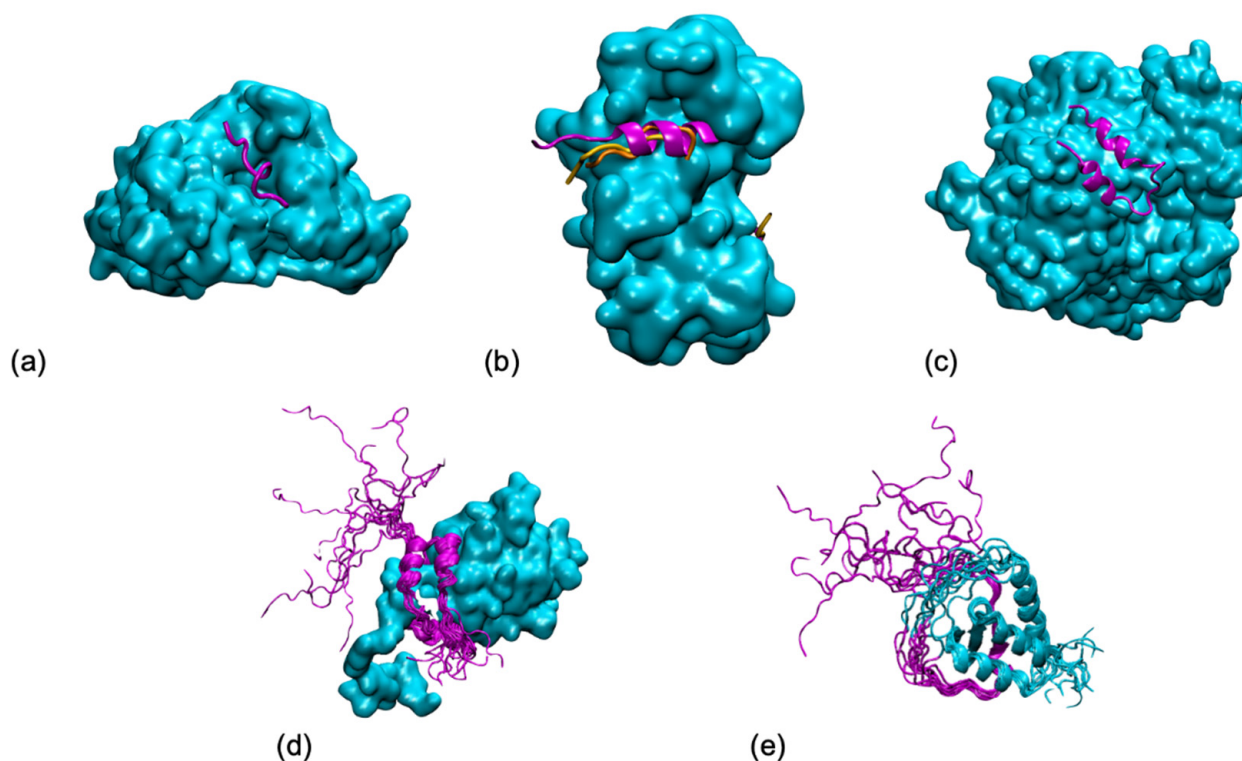


Figure 1. Five shades of disorder in protein/protein interactions (examples taken from (Horvath et al., PLoS Comput Biol 2020, 16, e1007864)): (a) Folding upon binding of an antigen (in magenta) from *P. falciparum* on an antibody (in cyan) from (pdb 4qxt). (b) Polymorphism of a ribosomal kinase (magenta and oranges) upon binding to S100B (in cyan) (pdb 5csf, 5csi, 5csj). (c) Conditional folding, the folding of the N-terminal tail (in magenta) from yeast ribonucleotide reductase (in cyan) depends on the interaction partner (pdb 1zyz). (d) Fuzziness, the p150 subunit of the eukaryotic initiation factor 4F (in magenta) wraps around the translation initiation factor 4E, but its N-terminal tail remains disordered (pdb 1rf8). (e) Disorder, both partners remain mostly disordered in the AF4-AF9 complex (pdb 2lm0). All graphical representations have been made using the VMD software.

Meanwhile, our understanding of the functional importance of fuzziness in protein interactions has been steadily increasing. Protein structural heterogeneity enables interactions with multiple partners, either simultaneously or consecutively [13], and weakens the sequence constraint on specificity [14–16]. For example, disordered histone tails serve as hubs, regulating chromatin accessibility and playing a central part in the nucleosome stability [17].

The first part of the present review presents recent developments in computational methods designed to investigate protein interfaces that include disorder. In particular, we discuss the integration of experimental information provided by various techniques. In its second part, the review focuses on two specific cases, microtubules, and homologous recombination nucleoprotein filaments, where a network of intrinsically disordered tails exerts regulatory function in recruiting partner macromolecules, proteins, or DNA and tuning the atomic level association. Notably, we show how computational approaches such as molecular dynamics (MD) simulations can bring new knowledge to help bridging the gap between experimental analysis, that mostly concerns ensemble properties, and the behavior of individual disordered protein chains that contribute to regulation functions.

2. Modeling Tools for Fuzzy Complexes

2.1. All-Atom Force Fields

The systematic comparison of classic all atom force fields commonly used for molecular dynamics simulations for modeling disordered systems highlighted several problems [18]. The first issue met with older force fields was the over-stabilization of secondary structures elements, α -helices and β -sheets, thus making the observation of the unfolded states that are characteristic of IDPs difficult [19]. This problem was addressed by optimizing the backbone torsion parameters against experimental NMR data [20–22]. Note that when working on the reparametrization of a force field, one must also pay some special attention to the training data that are used. For example, including data from coil fragments will help improving the reparametrization of the dihedral parameters [22–24]. This strategy was notably applied in the Amber ff03* and ff99SB* [25], CHARMM22* [26], OPLS-AA/M [27] and OPLS3 [28] force fields. Using a training set comprising both folded and unfolded proteins is of particular importance if one wants to model biological systems where order and disorder coexist, and this approach was used when developing the ff03CMAP force field [29].

The protein–water interaction is another central issue when developing force-fields for IDPs, as they do not have the hydrophobic core with many buried nonpolar residues that is usually found in folded proteins. As a consequence, all-atom force fields would lead to the over-stabilization of the collapsed molten globule state compared to the extended state [30,31]. The refinement of the protein–water interaction can be done via an adjustment of the Lennard-Jones potential parameter [21]. This led to the development of the TIP4P-D water model, which is better suited for the extended shape of IDPs [32]. Interestingly, this permitted to solve another problem encountered with older force fields, namely the over-stabilizing of protein–protein interactions, which could lead to protein aggregation [33]. Despite the success achieved through these developments, some issues remain. For example, while folded, globular proteins tend to unfold upon heating, IDPs have been shown to present some temperature-induced partial folding or the formation of secondary structures [34–36], and this effect still has to be accurately modeled [37].

The next step for improving the accuracy of IDP-specific force fields might lie in a better description of the electrostatic and hydrogen-bonding interactions, since these polar interactions play a key role in the IDPs structural behavior. This should be handled by polarizable force fields, and many efforts have been made in that direction over the last two decades [38–41]. Polarization has been implemented in all the traditional force fields, notably with AMOEBA [42], using fluctuating charge models [43], the Drude oscillator model [44], or induced dipoles [45]. These force fields remain computationally costly, but should greatly benefit from the increase in computational efficiency provided by GPUs.

2.2. Alternate Protein and Solvent Models

Coarse-grain protein models allow to push further the accessible length and time-scales of the simulations by reducing the number of degrees of freedom that have to be considered during the simulation. This can be particularly useful when investigating long timescale processes such as crowding. A classic coarse-grain approach is the use of G \ddot{o} -like (or native centric) models to investigate coupled folding-binding events [46,47]. Multi-state G \ddot{o} -like models have also been developed to study IDPs that can bind to different partners [48]. To describe fuzzy complexes, where no folding event is associated to binding, several alternative coarse-grain models are available, which were modified to be used for IDPs [49–52]. Among them we can mention AWSEM [53], PLUM [54], OPEP [55], UNRES [56], and SYRAH [57]. Analytical approaches derived from polymer physics somehow represent an ultimate stage of coarse-graining. These can be used to describe IDPs properties [58], but the remaining challenge is to relate these properties, and notably the phase behavior, to the IDP sequence [59,60].

Even without taking into account polarization, explicit solvent remain expensive from the computational point of view. An alternative is to use an implicit solvent model, where

the solvation term will only depend on the protein coordinates. A classic implicit solvation method for folded systems is the Generalized Born model [61,62]. However, it presents the same issues as the traditional all atom force fields, namely an over-stabilization of the secondary structure elements and an over-collapse of the disordered states [63–65]. The problem can be addressed by basing the solvation term on the experimental solvation free energies of functional group and weighting it as a function of the group solvent exposure. This approach was used in the EEF1 [66] and ABSINTH [67,68] models. One should however keep in mind that using an implicit solvent model also means that one no longer has access to the detail of the solvent molecules individual behavior at the protein/water interface, and in particular to the water-mediated hydrogen bonds.

2.3. Algorithms

While classical molecular dynamics simulations remain a first choice tool for modeling protein assemblies, the efficient sampling of the rugged conformational landscape of IDPs is a costly process as it requires the crossing of many energy barriers. In particular, coupled folding and binding of IDPs to their partners are still out of reach due to the large number of degrees of freedom as well as the extensive conformational transitions involved in the process [69]. As a consequence, numerous enhanced sampling methods have been developed that will accelerate the exploration of a disordered system's energy surface [70].

A first strategy is to add potential energy terms that will help overcome energy barriers along the simulation and improve the conformational space sampling. This is the case of metadynamics [71,72] and multi-canonical MD [73], which were used to investigate the coupled folding and binding process in the α -MoRE-MeV-XD complex [74] and in the RAP74-FCP1 complex [75] respectively. In umbrella sampling simulations, which were used to study the formation of the c-myb-KIX complex [76], the added potential takes the shape of a harmonic constraint. One should also consider the accelerated MD approach [77], which was used to perform microsecond-long simulations showing the partial folding of the GCN4 activation domain upon binding its coactivator [78].

Another option, that can be combined with the previous one, is to use multiple replicas from parallel trajectories which will be exchanged along time [79]. The replicas will differ by temperature in T-REMD, or by the introduction of a bias in the potential (BEMD). These approaches were used for modeling the binding of the ArkA IDP with a SH3 domain [80], and of p53 on MDM2 [81].

One can also apply constraints on a subset of degrees of liberty to accelerate the sampling. Bui and McCammon used targeted MD to investigate the conformational transitions undergone by fasciculin upon binding to acetylcholinesterase [82]. In the GNEIMO approach, the high frequency bond and angle vibrations are frozen, which enables the simulation of long timescale transitions that are inaccessible with classical MD [83].

However, the question remains whether all these enhanced sampling methods can capture realistic dynamics as well as correct ensemble properties. Comparison of the conformational dynamics obtained at different timescales (from picoseconds to tens of nanoseconds) by experimental approaches and MD simulations still presents discrepancies [84].

Conformational ensembles for IDPs and fuzzy complexes can also be generated with the *Flexible-Meccano* tool, which builds multiple copies of a polypeptide chain by random sampling of the backbone dihedral angles [85], or MoMA, a Robotics- and Artificial Intelligence-based approach initially developed to sample flexible loops but that can be used for open chains as well [86]. This approach was combined with SAXS data to investigate the structure of an intramolecular fuzzy complex in the Src family kinases [87].

2.4. Integrating Experimental Data

Conventional experimental methods, such as X-ray crystallography, SAXS, NMR, FRET or CryoEM, are not sufficient on their own to determine the conformational ensemble

that characterizes a fuzzy complex, as they will only provide mean values and a global structural signal for the system. However, they can still bring in some precious information regarding secondary structure contents, side chain orientations and the dynamics and lifetime of local residue contacts. These can be used for the pruning of a conformational ensemble generated by an unbiased simulation [20,88]. Alternatively, experimental data will be used as a set of constraints and a starting point for the modeling of molecular assemblies involving ordered and disordered proteins [13,89]. The resulting ensembles can be found in the Protein Ensemble Database (PED, <https://proteinensemble.org/>) (accessed on 11 October 2021), an open access repository for the deposition of structural ensembles, including IDPs [90]. In addition, the FuzDB database (<http://protdyn-database.org/>) (accessed on 11 October 2021) specifically focuses on fuzzy complexes [91]. It was used to develop the FuzPred method (http://protdyn-fuzpred.org) (accessed on 11 October 2021), which predicts the binding mode of disordered proteins based on their amino acid sequences and without prior knowledge of the interaction partners [92,93]. SAXS and FRET can also provide us with information regarding the size of IDPs as measured by their radius of gyration (R_g), which can be used for the training of IDP force fields [39], or for confronting MD simulations results [94,95].

Over the past decades, integrative approaches have also proved a valuable tool for deciphering protein interactions that involve one, or more, disordered partner [96,97]. For example, NMR and all-atom MD are a classic combination to study protein assemblies, with NMR parameters being used to set up the starting structures for the simulations [98]. Solvent paramagnetic relaxation enhancement (sPRE), which uses NMR with the addition of soluble paramagnetic molecules, will provide quantitative information regarding surface accessibility at atomic resolution. This data can be used to map solvent-exposed regions in protein assemblies and allows the detection of transient interactions in fuzzy complexes [99]. Tsytlonok et al. investigated the conformational dynamics of the complex formed by the IDP p27 and Cdk2/cyclin A [100]. They combined single molecule FRET and REMD to gain further insight in a multistep binding mechanism that involves conformational selection followed by local induced folding of the p27 partner. As mentioned earlier, SAXS gives us information on the shape of biomolecular objects over a wide range of sizes, and also on their oligomerization state. The fact that this technique can handle polydisperse systems makes it particularly useful when working on IDPs and numerous ensemble modeling tools based on SAXS data have been developed [101]. The meta-inference approach developed in the Vendruscolo laboratory permits to simultaneously determine the structure and dynamics of macromolecular systems from cryo-electron microscopy density maps [102,103]. This was applied by Brotzakis et al. to determine the conformational ensemble and the dynamics of the tau-microtubule complex [104], based on the Cryo-EM determined structure of this macromolecular assembly [105].

Finally, one should mention the growing use of artificial intelligence and machine learning (AI/ML) approaches for characterizing conformational ensembles in disordered systems, and integrating experimental data with simulations [106]. For example, Ramanathan et al. used a ML approach to investigate the disorder to order transition in viral proteins binding on host pro-apoptotic proteins [107]. Machine learning can also be used for the refinement of force fields parameters, by adapting these to reproduce experimental SAXS scattering profiles [108].

2.5. Measuring and Comparing Disorder

The traditional metrics that were developed to analyze the structure of folded proteins, such as RMSD, are no longer relevant when working on IDPs, and comparing conformational ensembles of IDPs requires the development of specific tools. Lazar et al. proposed to use distance-based metrics relying on the median and the standard deviation of inter-residue distance distributions [109]. This approach is of particular interest for partially folded proteins comprising both a structured domain and IDRs, as it enables to directly identify the protein fragments that present structural similarity. The Local Compaction

Plots (LCP) [110], which show the intramolecular distance between residues separated by a fixed span along the primary sequence, represent another interesting tool for analyzing MD trajectories, as they highlight disordered and folded region in the protein, while still showing its conformational diversity along time.

3. Functional Role of the Fuzzy Interface in the Cell

A growing body of reported observation on fuzzy interfaces depicts a continuum of association properties that range from quasi non-selective, liquid-like interactions to highly specific interactions, resulting from already mentioned folding-upon-binding mechanisms. Liquid-like association aims at ensuring proximity between the partner macromolecules and mostly involves electrostatic or polar interactions. Disordered proteins are a major component of membraneless cellular compartments, where they participate in liquid–liquid phase separation while avoiding aggregation, via the formation of dynamic, multivalent interactions [111,112]. Interestingly, high level of disorder in fuzzy interfaces are not necessarily associated to low affinity: in the complex between the human proteins histone H1 and its nuclear chaperone prothymosin- α , large opposite net charges have been shown to confer picomolar affinity to the association in spite of the absence of defined binding sites [113].

In this section, we examine intermediate situations where disordered proteins or segments present ubiquitous motifs that can transiently associate to defined binding sites on folded protein partners. We more specifically address cases where both structurally organized and disordered regions coexist in the same protein. Typical examples are proteins that present disordered C- or N-terminal tails, largely represented among DNA-binding or DNA-processing proteins. The disordered tails in these proteins generally present a net charge. Positive tails can assist the efficiency of DNA search for specific sequences by proteins such as transcription factors. The tails can non-selectively bind to DNA and promote inter-segment cross talks in a “monkey-bar”-type mechanism [114,115]. When they bear a net negative charge, the tails can compete with DNA for binding sites [116,117] or they can bind one or more protein partners. For example, the tetrameric SSB protein that binds DNA single strands, an essential contributor of DNA replication, recombination, and repair in bacteria, functions as a recruitment platform where its four negatively charged C-terminal tails can simultaneously bind one or more proteins, thus favoring the transfer of bound DNA to these partner proteins [118,119]. Competition and recruitment mechanisms are also common in self-associating proteins that present disordered tails, such as tubulin or fibrinogen [120,121]. In those cases, the strongly charged tails actively contribute to the binding of partner subunits using fly-casting types of mechanisms, but do not participate in the protein–protein interface once the assembly is formed. In the case of tubulin associating into microtubules, the tails form molecular brushes around the microtubule lattice and participate in active or passive diffusion of proteins along the microtubule protofilaments [122].

The delicate balance between binding and unbinding provides the disordered terminal tails affinity tuning functions: the tails have been shown to modulate binding behaviors in response to changes in salt concentration or composition. The tail properties are also very sensitive to changes in the distribution of their charges resulting from post-translational modifications, as well as associated excluded volume modifications [121]. In what follows, we will concentrate on two examples taken from our former or present studies where charged disordered tails interact with the folded core regions of the protein they belong to or with a lattice of these protein cores. We will discuss particular physical properties, frustration, and steric adaptability, that may enable the tails to appropriately respond to changes in their environment, and also how the interplay between the folded and the disordered protein regions helps participate in the binding modulation.

3.1. Interactions between the C-Terminal Tails of α,β -Tubulin Dimers and the Tubulin Core

Tubulin proteins exist in the cell as dimers of α - and β -tubulin, two closely related proteins whose sequences essentially differ at the level of their disordered C-terminal tails;

both tails bear a net negative charge but they differ in length and amino-acid composition. α,β -tubulin dimers are the building blocks of microtubules (MT), the largest components of the cytoskeleton, that form highways for intracellular trafficking as well as separating chromosomes during meiosis. Modeling and NMR studies have shown that in tubulin dimers, both α - and β -tails can interact with the structurally organized region of the protein dimer (the core region) in spite of the core surface potential being mainly negative [95]. The tails are also known to contribute to the formation of microtubules by favoring the proper uptake of new tubulin dimers within the tubular architecture: alternative association forms of tubulin could be observed in the absence of tails [123]. It is therefore likely that the MT tubulin tails interact with free tubulin dimers during the assembly process, thus orienting the dimers toward the desired binding geometry. This association however needs to be transient, since a large fraction of the tails (notably the longer β -tails) are released during the process and become free to interact with microtubule-binding proteins (MAPs) [121]. Similar process has been observed by AFM when fibrin proteins assemble into fibrinogen [120], while the C-terminal tails of RecA proteins have been shown to be involved in their association process into filaments [124]. These observations indicate that the ability of the disordered protein tails to bind the protein core surface but also to unbind from it is key to their function.

How exactly the tails influence auto-assembly remains to be established. Theoretical simulation of the tubulin tails binding to their associated dimeric protein cores enabled to gain insights on this question [95]. Notably, while the surface spanned by the tails during atomic molecular dynamics simulations was found compatible with ensemble observations obtained by AFM (radius of gyration), the simulation enabled proposing a finer characterization of the spatial and temporal distribution of the tails, based on specifically developed metrics using the position of the tail center of mass, together with time analysis of the contacts between tails and protein cores. This analysis revealed the presence of a handful of specific tail-binding spots, or anchors, distributed on the tubulin surface and presenting reduced surface areas. The tails develop versatile interactions with these binding spots, mostly based on electrostatic complementarity [95,121]. Interestingly, we observed that negatively charged amino-acid patches distributed along the whole β -tail (see Figure 2) can individually bind separate binding spots, and that adjacent negative patches can slide within a given anchor and exchange their binding interactions. Binding different sites on the core surface does not seem to be cooperative but rather self-exclusive, one reason being that several negative patches on the tail may not be able to simultaneously access spatially separated anchors. Another factor arises from the electrostatic potential around the tubulin dimer. Indeed, we found that the electrostatic potential partitions the space available to the tails into electronegative regions, that are strongly repulsive for the most part of the tail length, and electropositive funnels that strongly attract the negative tail patches. This situation creates tension and frustration in the bound tails, part of which needs to reside in an unfavorable, repulsive region to allow contacts to form on the tubulin surface [95]. Frustration, a tradeoff between conflicting forces within their interatomic contact network environment [125], has been identified as a critical property of IDPs or IDRs binding to their protein targets [13]. Because of unsolved conflicts at such interfaces, added to the multiplicity of binding sites, the disordered regions are prone to switching to alternate binding geometries. We propose that the concept of frustration extends to long-range interactions such as the response of protein tail conformations to the potential energy created by the protein core, coupled to the physical attachment between the tail and the protein core. Long range frustration may constitute a powerful driving force to facilitate the tail unbinding from its core protein. It is also easily tunable via changes in the salt concentration or modification of the charge distribution in the tail via post-translational modifications (PTM). Indeed, recent work from Bigman and Levy showed that PTMs tune the binding ability of the tails to the MT, a function that is also partly linked to their exclusion volume properties [121]. Recent simulations of the hepatitis B virus (HBV) Core protein, that exhibits a 33-residue long, positively charged and intrinsically

disordered C-terminal tail, suggests the existence of long range frustration in the binding of the negatively charged extremity of the tail to the positively charged extremity of the HBV capsid spike, with very sparse interactions between the rest of the tail and the external surface of the spikes [126].

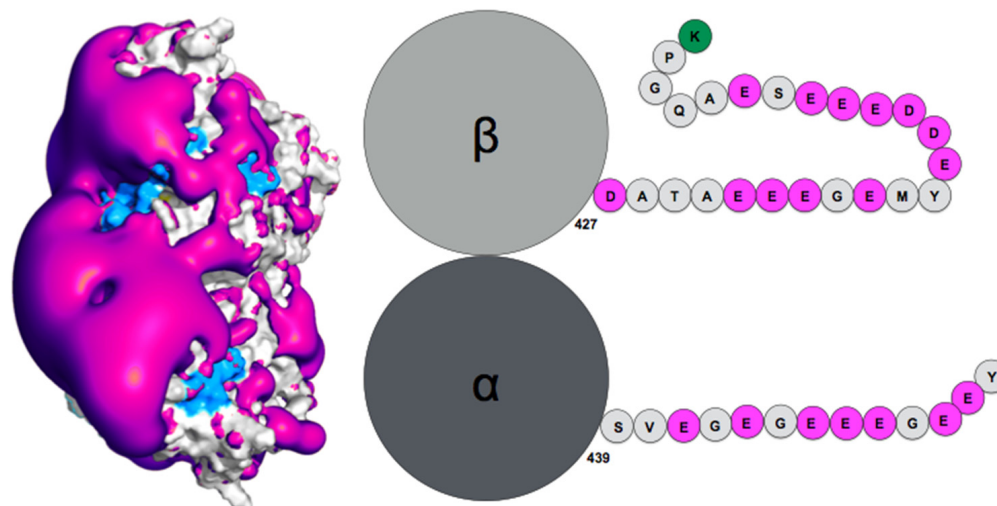


Figure 2. (Left) Negative surface electrostatic potential (-1 kT, magenta) of the α I/ β III isotype tubulin body without tails. The anchor residues involved in interactions with the disordered tail during molecular dynamics simulations are shown in blue; the representation is based on data published in (Laurin et al., *Biochemistry* 2017, 56, 1746); (right) schematic sequence of the α I/ β III isotype of tubulin, the acidic amino acids are highlighted in magenta and the basic terminal residue in green.

3.2. Role of the RecA Protein C-Terminal Tails in Homologous Recombination

Homologous recombination permits the faithful repair of DNA double strand breaks in the genome, by recruiting intact genomic DNA (dsDNA) with sequence similar to the damaged DNA and using that DNA to restore the lost sequence continuity. To this aim, the dsDNA complementary strand is captured by a single strand (ssDNA) from the damaged DNA, in a process called strand exchange that occurs within filaments of recombinase proteins (RecA in bacteria) [127]. Alike many proteins that process DNA, *E. coli* RecA proteins present a disordered terminal tail, here a 25-amino acids, negatively charged C-terminal tail with seven acidic amino-acids. The C-terminal tail was shown to participate in the regulation of various stages of the recombination process: the filament self-assembly, the intake of the dsDNA into the filament, and the yield of strand exchange. While all those stages can take place in the absence of the tail or with partly deleted tails, the tail has been shown to mediate the response of the process to changes in pH or in magnesium concentration [124,128,129]. Specifically, full-length tails slow the RecA self-association process but promote the formation of longer and more stable filaments on ssDNA [124]. During the search, the dsDNA intake is also slowed in the presence of the tail, but this effect is reduced by adding 2 mM free Mg^{2+} ions. This observation has been related to the fact that the searched dsDNA non-specifically binds to the filament gateway [130], a region that crosses the filament groove and involves basic amino-acids from the C-terminal domains. The acidic C-terminal tails may restrict the access of the dsDNA to the gateway via electrostatic repulsion or physical steric hindrance and the added magnesium ions may reduce the electrostatic repulsion and possibly induce the formation of secondary structures in the tails, which would confine the tails in a smaller volume.

How the disordered tail influences the strand exchange process is more puzzling. In the presence of the disordered tails, addition of 5 mM magnesium ions maximizes the formation of the strand exchange product; the magnesium concentration has no effect if the tail has been deleted, indicating that the tail is fully involved in the process. It has been proposed that in condition of low magnesium concentration, the tail may compete with the incorporated dsDNA for binding to the filament secondary binding site (site II) [128,129]. In that hypothesis, the tail may stimulate dsDNA binding to site II by disengaging from that site following changes in magnesium concentration [128]; alternatively, the tail may assist unbinding of non-homologous incorporated dsDNA, thus accelerating the search process [124]. However, site II is buried in the filament interior whereas the tail extremities are situated at the periphery of the filament [131,132] (Figure 3). In order for its acidic residues to reach site II in the filament interior, the tail would need to adopt a stretched conformation along the C-terminal domain toward the filament interior. Our first exploration of the tail structural dynamics by molecular dynamics simulations did not show such behavior [129]. Instead, all seven tails of a simulated filament, made of seven RecA monomers bound to a 21-nt ssDNA, remained at the exterior of the filament during the course of two 200 ns simulation with no added or 2 mM magnesium ions. The tails partly formed helical folds and partly lied on the external surface of the filament, sometimes spanning over two consecutive monomers, but they did not penetrate into the filament interior. Recently, we further explored the tail dynamics by taking into account the perturbation induced in the filament structure by the hydrolysis of ATP molecules situated at the interface between monomers. Indeed, experimental observation of the influence of the tails on the strand exchange process was performed in conditions of ATP hydrolysis, using ATP regeneration system. Our recent modeling studies indicate that the response of the filament to ATP hydrolysis may involve important modifications in the spatial partitioning of the filament groove, which may modify the tail accessibility to the filament interior. We used our published model of a 2-turn (12 monomers) filament [133] where the central RecA-RecA interface was modified for an ADP interface (the RecA-RecA binding geometries differ whether the cofactor is ATP or ADP) as a starting point for two 100-ns molecular dynamics simulations, one with no added magnesium ions and one with 5 mM magnesium concentration. Interestingly during the simulation with 5 mM magnesium, the tail associated to the central monomer with modified interface spontaneously penetrated in the filament interior and reached the secondary DNA binding site, showing that this proposed behavior is indeed topologically and energetically possible within filament architectures associated to ATP hydrolysis (Figure 3). These preliminary simulations need to be replicated and call for further investigation in order to draw any reliable conclusions on the effects of the magnesium concentration; notably, force fields adapted not only to different levels of structural disorder but that also correctly capture magnesium ion interactions need to be tested in order to confirm the reported observations. Magnesium ions can individually mediate interactions between negative charges but can also as an ensemble contribute to weaken salt-bridge interactions, therefore contributing to order-disorder transitions. Molecular dynamics is a tool of choice for disentangling individual from ensemble effects of the magnesium ions, provided that the interactions are correctly accounted for. Our present MD observations are too preliminary to conclude about the exact role of the magnesium, nevertheless they point to topological and steric factors as additional factors for the tails to exert their control function.

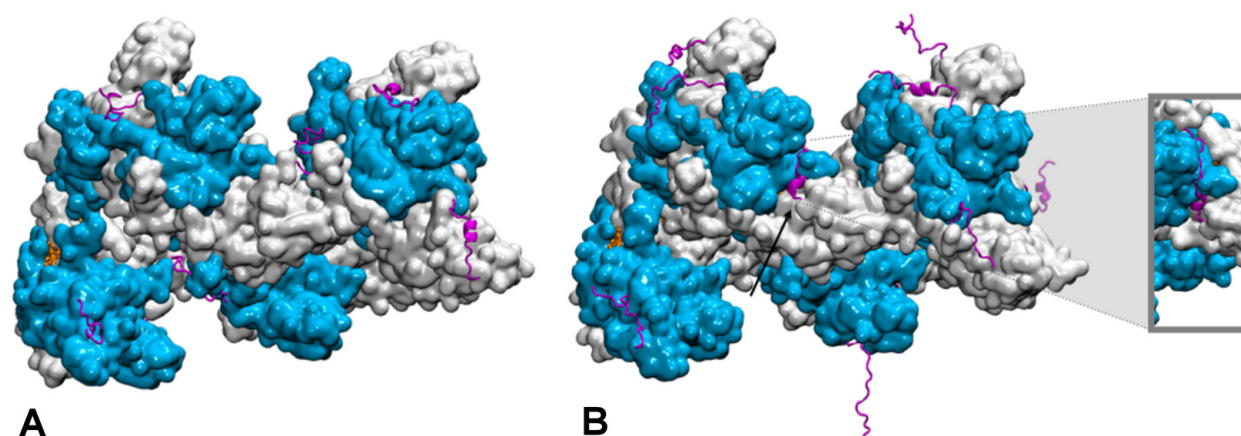


Figure 3. Conformational dynamics of the C-terminal tails of a two-turn, twelve monomer RecA-ssDNA filament with modified central interface, after 100 ns of molecular dynamics simulation. Successive RecA proteins are alternatively colored cyan and white. The tails (magenta, cartoon representation) explore different regions of the conformational space in terms of folding—partial α -helical folds or extended conformation—and binding to the protein core surface. (A) Simulation with no added salt; the tails mostly bind the core protein surface; (B) simulation with 5 mM Mg^{2+} ; some tails remain far from the surface, the tail from the central monomer (black arrow) penetrates inside the filament and reaches the filament site II, within 8 Å of the basic residue cluster of the neighboring monomer. The insert shows a view of the penetrating tail after 30° rotation around the filament axis. Simulations conditions are described in (Kim et al., *Nucleic Acids Res.* 2018, 46, 2548).

4. Conclusions

As the importance of IDPs and IDRs biological functions is now fully acknowledged, the development of specific numerical tools that permit to properly model complex assemblies where order meets disorder is reaching a point where new information can be obtained out of numerical simulation studies. This requires that experimental information, which often consists of statistical values obtained from ensemble conformations, but can also include indirect information on the IDP or IDR response to physicochemical perturbations, is integrated to the model. A large panel of new functions is now within reach, which opens the way for exciting new explorations.

Author Contributions: S.S.-M. and C.P. contributed in the conception, the bibliography search, the figures realization and the manuscript writing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the “Initiative d’Excellence” program from the French State (Grant “DYNAMO”, ANR-11-LABX-0011-01). Simulations were performed using HPC resources from GENCI-CINES (2016- [x2016077438] and 2018-[A0040707438]).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The preliminary data presented in this study are available on request from the corresponding author.

Acknowledgments: CP thanks Masayuki Takahashi for insightful discussions on RecA C-terminal tails.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230. [[CrossRef](#)]
2. Wright, P.E.; Dyson, H.J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331. [[CrossRef](#)] [[PubMed](#)]
3. Perdigo, N.; Heinrich, J.; Stolte, C.; Sabir, K.S.; Buckley, M.J.; Tabor, B.; Signal, B.; Gloss, B.S.; Hammang, C.J.; Rost, B.; et al. Unexpected features of the dark proteome. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15898–15903. [[CrossRef](#)] [[PubMed](#)]

4. Bhowmick, A.; Brookes, D.H.; Yost, S.R.; Dyson, H.J.; Forman-Kay, J.D.; Gunter, D.; Head-Gordon, M.; Hura, G.L.; Pande, V.S.; Wemmer, D.E.; et al. Finding Our Way in the Dark Proteome. *J. Am. Chem. Soc.* **2016**, *138*, 9730–9742. [[CrossRef](#)]
5. Uversky, V.N. The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome. *J. Biomed. Biotechnol.* **2010**, *2010*, 568068. [[CrossRef](#)]
6. Tompa, P. Unstructural biology coming of age. *Curr. Opin. Struct. Biol.* **2011**, *21*, 419–425. [[CrossRef](#)] [[PubMed](#)]
7. Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N.; Obradovic, Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J. Proteome Res.* **2007**, *6*, 1882–1898. [[CrossRef](#)]
8. Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29. [[CrossRef](#)]
9. Monzon, A.M.; Necci, M.; Quaglia, F.; Walsh, I.; Zanotti, G.; Piovesan, D.; Tosatto, S.C.E. Experimentally Determined Long Intrinsically Disordered Protein Regions Are Now Abundant in the Protein Data Bank. *Int. J. Mol. Sci.* **2020**, *21*, 4496. [[CrossRef](#)]
10. Fuxreiter, M. Fuzziness in Protein Interactions—A Historical Perspective. *J. Mol. Biol.* **2018**, *430*, 2278–2287. [[CrossRef](#)]
11. Tompa, P.; Fuxreiter, M. Fuzzy complexes: Polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **2008**, *33*, 2–8. [[CrossRef](#)]
12. Fuxreiter, M. Classifying the Binding Modes of Disordered Proteins. *Int. J. Mol. Sci.* **2020**, *21*, 8615. [[CrossRef](#)]
13. Freiburger, M.I.; Wolynes, P.G.; Ferreira, D.U.; Fuxreiter, M. Frustration in Fuzzy Protein Complexes Leads to Interaction Versatility. *J. Phys. Chem. B* **2021**, *125*, 2513–2520. [[CrossRef](#)]
14. Ross, E.D.; Edsles, H.K.; Terry, M.J.; Wickner, R.B. Primary sequence independence for prion formation. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12825–12830. [[CrossRef](#)] [[PubMed](#)]
15. Lu, X.; Hamkalo, B.; Parseghian, M.H.; Hansen, J.C. Chromatin condensing functions of the linker histone C-terminal domain are mediated by specific amino acid composition and intrinsic protein disorder. *Biochemistry* **2009**, *48*, 164–172. [[CrossRef](#)] [[PubMed](#)]
16. Fuxreiter, M.; Tompa, P. Fuzzy complexes: A more stochastic view of protein function. *Adv. Exp. Med. Biol.* **2012**, *725*, 1–14. [[CrossRef](#)] [[PubMed](#)]
17. Peng, Y.; Li, S.; Landsman, D.; Panchenko, A.R. Histone tails as signaling antennas of chromatin. *Curr. Opin. Struct. Biol.* **2021**, *67*, 153–160. [[CrossRef](#)]
18. Rauscher, S.; Gapsys, V.; Gajda, M.J.; Zweckstetter, M.; de Groot, B.L.; Grubmüller, H. Structural Ensembles of Intrinsically Disordered Proteins Depend Strongly on Force Field: A Comparison to Experiment. *J. Chem. Theory Comput.* **2015**, *11*, 5513–5524. [[CrossRef](#)]
19. Best, R.B.; Buchete, N.V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys. J.* **2008**, *95*, L07–L09. [[CrossRef](#)]
20. Best, R.B. Computational and theoretical advances in studies of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2017**, *42*, 147–154. [[CrossRef](#)]
21. Robustelli, P.; Piana, S.; Shaw, D.E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4758–E4766. [[CrossRef](#)]
22. Mu, J.; Liu, H.; Zhang, J.; Luo, R.; Chen, H.F. Recent Force Field Strategies for Intrinsically Disordered Proteins. *J. Chem. Inf. Model.* **2021**, *61*, 1037–1047. [[CrossRef](#)] [[PubMed](#)]
23. Best, R.B.; Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015. [[CrossRef](#)] [[PubMed](#)]
24. Yu, L.; Li, D.W.; Bruschweiler, R. Balanced Amino-Acid-Specific Molecular Dynamics Force Field for the Realistic Simulation of Both Folded and Disordered Proteins. *J. Chem. Theory Comput.* **2020**, *16*, 1311–1318. [[CrossRef](#)] [[PubMed](#)]
25. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958. [[CrossRef](#)]
26. Piana, S.; Lindorff-Larsen, K.; Shaw, D.E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **2011**, *100*, L47–L49. [[CrossRef](#)]
27. Robertson, M.J.; Tirado-Rives, J.; Jorgensen, W.L. Improved Peptide and Protein Torsional Energetics with the OPLSAA Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 3499–3509. [[CrossRef](#)]
28. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J.Y.; Wang, L.; Lupyan, D.; Dahlgren, M.K.; Knight, J.L.; et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281–296. [[CrossRef](#)]
29. Zhang, Y.; Liu, H.; Yang, S.; Luo, R.; Chen, H.F. Well-Balanced Force Field ff03CMAP for Folded and Disordered Proteins. *J. Chem. Theory Comput.* **2019**, *15*, 6769–6780. [[CrossRef](#)]
30. Nettels, D.; Müller-Spath, S.; Kuster, F.; Hofmann, H.; Haenni, D.; Ruegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; et al. Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 20740–20745. [[CrossRef](#)]
31. Piana, S.; Klepeis, J.L.; Shaw, D.E. Assessing the accuracy of physical models used in protein-folding simulations: Quantitative evidence from long molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2014**, *24*, 98–105. [[CrossRef](#)] [[PubMed](#)]
32. Piana, S.; Donchev, A.G.; Robustelli, P.; Shaw, D.E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113–5123. [[CrossRef](#)]

33. Abriata, L.A.; Dal Peraro, M. Assessment of transferable forcefields for protein simulations attests improved description of disordered states and secondary structure propensities, and hints at multi-protein systems as the next challenge for optimization. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 2626–2636. [[CrossRef](#)] [[PubMed](#)]
34. Uversky, V.N. Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* **2009**, *28*, 305–325. [[CrossRef](#)] [[PubMed](#)]
35. Kjaergaard, M.; Norholm, A.B.; Hendus-Altenburger, R.; Pedersen, S.F.; Poulsen, F.M.; Kragelund, B.B. Temperature-dependent structural changes in intrinsically disordered proteins: Formation of alpha-helices or loss of polyproline II? *Protein Sci.* **2010**, *19*, 1555–1564. [[CrossRef](#)]
36. Wuttke, R.; Hofmann, H.; Nettels, D.; Borgia, M.B.; Mittal, J.; Best, R.B.; Schuler, B. Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5213–5218. [[CrossRef](#)]
37. Jephthah, S.; Staby, L.; Kragelund, B.B.; Skepo, M. Temperature Dependence of Intrinsically Disordered Proteins in Simulations: What are We Missing? *J. Chem. Theory Comput.* **2019**, *15*, 2672–2683. [[CrossRef](#)]
38. Kaminski, G.A.; Stern, H.A.; Berne, B.J.; Friesner, R.A.; Cao, Y.X.; Murphy, R.B.; Zhou, R.; Halgren, T.A. Development of a polarizable force field for proteins via ab initio quantum chemistry: First generation model and gas phase tests. *J. Comput. Chem.* **2002**, *23*, 1515–1531. [[CrossRef](#)]
39. Huang, J.; MacKerell, A.D., Jr. Force field development and simulations of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2018**, *48*, 40–48. [[CrossRef](#)]
40. Wang, A.; Zhang, Z.; Li, G. Higher Accuracy Achieved in the Simulations of Protein Structure Refinement, Protein Folding, and Intrinsically Disordered Proteins Using Polarizable Force Fields. *J. Phys. Chem. Lett.* **2018**, *9*, 7110–7116. [[CrossRef](#)]
41. Inakollu, V.S.; Geerke, D.P.; Rowley, C.N.; Yu, H. Polarizable force fields: What do they add in biomolecular simulations? *Curr. Opin. Struct. Biol.* **2020**, *61*, 182–190. [[CrossRef](#)] [[PubMed](#)]
42. Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J.W.; Ren, P. The Polarizable Atomic Multipole-based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063. [[CrossRef](#)] [[PubMed](#)]
43. Patel, S.; Mackerell, A.D., Jr.; Brooks, C.L., III. CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comput. Chem.* **2004**, *25*, 1504–1514. [[CrossRef](#)] [[PubMed](#)]
44. Lopes, P.E.M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A.D. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9*, 5430–5449. [[CrossRef](#)]
45. Cieplak, P.; Caldwell, J.; Kollman, P. Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: Aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *J. Comput. Chem.* **2001**, *22*, 1048–1057. [[CrossRef](#)]
46. Shoemaker, B.A.; Portman, J.J.; Wolynes, P.G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8868–8873. [[CrossRef](#)]
47. Rogers, J.M.; Oleinikovas, V.; Shammass, S.L.; Wong, C.T.; De Sancho, D.; Baker, C.M.; Clarke, J. Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15420–15425. [[CrossRef](#)]
48. Knott, M.; Best, R.B. Discriminating binding mechanisms of an intrinsically disordered protein via a multi-state coarse-grained model. *J. Chem. Phys.* **2014**, *140*, 175102. [[CrossRef](#)]
49. Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A.E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.* **2016**, *116*, 7898–7936. [[CrossRef](#)]
50. Cragnell, C.; Rieloff, E.; Skepo, M. Utilizing Coarse-Grained Modeling and Monte Carlo Simulations to Evaluate the Conformational Ensemble of Intrinsically Disordered Proteins and Regions. *J. Mol. Biol.* **2018**, *430*, 2478–2492. [[CrossRef](#)]
51. Baul, U.; Chakraborty, D.; Mugnai, M.L.; Straub, J.E.; Thirumalai, D. Sequence Effects on Size, Shape, and Structural Heterogeneity in Intrinsically Disordered Proteins. *J. Phys. Chem. B* **2019**, *123*, 3462–3474. [[CrossRef](#)] [[PubMed](#)]
52. Shea, J.E.; Best, R.B.; Mittal, J. Physics-based computational and theoretical approaches to intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2021**, *67*, 219–225. [[CrossRef](#)] [[PubMed](#)]
53. Wu, H.; Wolynes, P.G.; Papoian, G.A. AWSEM-IDP: A Coarse-Grained Force Field for Intrinsically Disordered Proteins. *J. Phys. Chem. B* **2018**, *122*, 11115–11125. [[CrossRef](#)]
54. Rutter, G.O.; Brown, A.H.; Quigley, D.; Walsh, T.R.; Allen, M.P. Testing the transferability of a coarse-grained model to intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31741–31749. [[CrossRef](#)]
55. Nguyen, P.H.; Derreumaux, P. Structures of the intrinsically disordered Abeta, tau and alpha-synuclein proteins in aqueous solution from computer simulations. *Biophys. Chem.* **2020**, *264*, 106421. [[CrossRef](#)]
56. Sieradzan, A.K.; Niadzvedtski, A.; Scheraga, H.A.; Liwo, A. Revised Backbone-Virtual-Bond-Angle Potentials to Treat the l- and d-Amino Acid Residues in the Coarse-Grained United Residue (UNRES) Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 2194–2203. [[CrossRef](#)]
57. Klein, F.; Barrera, E.E.; Pantano, S. Assessing SIRAH's Capability to Simulate Intrinsically Disordered Proteins and Peptides. *J. Chem. Theory Comput.* **2021**, *17*, 599–604. [[CrossRef](#)]

58. Schuler, B.; Soranno, A.; Hofmann, H.; Nettels, D. Single-Molecule FRET Spectroscopy and the Polymer Physics of Unfolded and Intrinsically Disordered Proteins. *Annu. Rev. Biophys.* **2016**, *45*, 207–231. [[CrossRef](#)]
59. Lin, Y.H.; Forman-Kay, J.D.; Chan, H.S. Sequence-Specific Polyampholyte Phase Separation in Membraneless Organelles. *Phys. Rev. Lett.* **2016**, *117*, 178101. [[CrossRef](#)] [[PubMed](#)]
60. Dignon, G.L.; Zheng, W.; Kim, Y.C.; Best, R.B.; Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* **2018**, *14*, e1005941. [[CrossRef](#)] [[PubMed](#)]
61. Still, W.C.; Tempczyk, A.; Hawley, R.C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129. [[CrossRef](#)]
62. Kleinjung, J.; Fraternali, F. Design and application of implicit solvent models in biomolecular simulations. *Curr. Opin. Struct. Biol.* **2014**, *25*, 126–134. [[CrossRef](#)] [[PubMed](#)]
63. Bottaro, S.; Lindorff-Larsen, K.; Best, R.B. Variational Optimization of an All-Atom Implicit Solvent Force Field to Match Explicit Solvent Simulation Data. *J. Chem. Theory Comput.* **2013**, *9*, 5641–5652. [[CrossRef](#)]
64. Lee, K.H.; Chen, J. Optimization of the GBMV2 implicit solvent force field for accurate simulation of protein conformational equilibria. *J. Comput. Chem.* **2017**, *38*, 1332–1341. [[CrossRef](#)]
65. Das, P.; Matysiak, S.; Mittal, J. Looking at the Disordered Proteins through the Computational Microscope. *ACS Cent. Sci.* **2018**, *4*, 534–542. [[CrossRef](#)]
66. Lazaridis, T.; Karplus, M. Effective energy function for proteins in solution. *Proteins* **1999**, *35*, 133–152. [[CrossRef](#)]
67. Vitalis, A.; Pappu, R.V. ABSINTH: A new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* **2009**, *30*, 673–699. [[CrossRef](#)]
68. Choi, J.M.; Pappu, R.V. Improvements to the ABSINTH Force Field for Proteins Based on Experimentally Derived Amino Acid Specific Backbone Conformational Statistics. *J. Chem. Theory Comput.* **2019**, *15*, 1367–1382. [[CrossRef](#)] [[PubMed](#)]
69. Mollica, L.; Bessa, L.M.; Hanouille, X.; Jensen, M.R.; Blackledge, M.; Schneider, R. Binding Mechanisms of Intrinsically Disordered Proteins: Theory, Simulation, and Experiment. *Front. Mol. Biosci.* **2016**, *3*, 52. [[CrossRef](#)] [[PubMed](#)]
70. Ikebe, J.; Umezawa, K.; Higo, J. Enhanced sampling simulations to construct free-energy landscape of protein-partner substrate interaction. *Biophys. Rev.* **2016**, *8*, 45–62. [[CrossRef](#)] [[PubMed](#)]
71. Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566. [[CrossRef](#)]
72. Laio, A.; Gervasio, F.L. Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* **2008**, *71*, 126601. [[CrossRef](#)]
73. Higo, J.; Umezawa, K. Free-energy landscape of intrinsically disordered proteins investigated by all-atom multicanonical molecular dynamics. *Adv. Exp. Med. Biol.* **2014**, *805*, 331–351. [[CrossRef](#)] [[PubMed](#)]
74. Han, M.; Xu, J.; Ren, Y.; Li, J. Simulation of coupled folding and binding of an intrinsically disordered protein in explicit solvent with metadynamics. *J. Mol. Graph. Model.* **2016**, *68*, 114–127. [[CrossRef](#)]
75. Wostenberg, C.; Kumar, S.; Noid, W.G.; Showalter, S.A. Atomistic simulations reveal structural disorder in the RAP74-FCP1 complex. *J. Phys. Chem. B* **2011**, *115*, 13731–13739. [[CrossRef](#)] [[PubMed](#)]
76. Ithuralde, R.E.; Roitberg, A.E.; Turjanski, A.G. Structured and Unstructured Binding of an Intrinsically Disordered Protein as Revealed by Atomistic Simulations. *J. Am. Chem. Soc.* **2016**, *138*, 8742–8751. [[CrossRef](#)] [[PubMed](#)]
77. Pierce, L.C.T.; Salomon-Ferrer, R.; De Oliveira, C.A.F.; McCammon, J.A.; Walker, R.C. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 2997–3002. [[CrossRef](#)]
78. Scholes, N.S.; Weinzierl, R.O. Molecular Dynamics of “Fuzzy” Transcriptional Activator-Coactivator Interactions. *PLoS Comput. Biol.* **2016**, *12*, e1004935. [[CrossRef](#)]
79. Sugita, Y.; Kamiya, M.; Oshima, H.; Re, S. Replica-Exchange Methods for Biomolecular Simulations. *Methods Mol. Biol.* **2019**, *2022*, 155–177. [[CrossRef](#)]
80. Gerlach, G.J.; Carrozz, R.; Stix, R.; Stollar, E.J.; Ball, K.A. A disordered encounter complex is central to the yeast Abp1p SH3 domain binding pathway. *PLoS Comput. Biol.* **2020**, *16*, e1007815. [[CrossRef](#)]
81. Zou, R.; Zhou, Y.; Wang, Y.; Kuang, G.; Agren, H.; Wu, J.; Tu, Y. Free Energy Profile and Kinetics of Coupled Folding and Binding of the Intrinsically Disordered Protein p53 with MDM2. *J. Chem. Inf. Model.* **2020**, *60*, 1551–1558. [[CrossRef](#)] [[PubMed](#)]
82. Bui, J.M.; McCammon, J.A. Protein complex formation by acetylcholinesterase and the neurotoxin fasciculin-2 appears to involve an induced-fit mechanism. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15451–15456. [[CrossRef](#)] [[PubMed](#)]
83. Gangupomu, V.K.; Wagner, J.R.; Park, I.H.; Jain, A.; Vaidehi, N. Mapping conformational dynamics of proteins using torsional dynamics simulations. *Biophys. J.* **2013**, *104*, 1999–2008. [[CrossRef](#)] [[PubMed](#)]
84. Rezaei-Ghaleh, N.; Parigi, G.; Soranno, A.; Holla, A.; Becker, S.; Schuler, B.; Luchinat, C.; Zweckstetter, M. Local and Global Dynamics in Intrinsically Disordered Synuclein. *Angewandte Chemie Int. Ed.* **2018**, *57*, 15262–15266. [[CrossRef](#)] [[PubMed](#)]
85. Ozenne, V.; Bauer, F.; Salmon, L.; Huang, J.R.; Jensen, M.R.; Segard, S.; Bernado, P.; Charavay, C.; Blackledge, M. Flexible-meccano: A tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* **2012**, *28*, 1463–1470. [[CrossRef](#)]
86. Barozet, A.; Molloy, K.; Vaisset, M.; Simeon, T.; Cortes, J. A reinforcement-learning-based approach to enhance exhaustive protein loop sampling. *Bioinformatics* **2020**, *36*, 1099–1106. [[CrossRef](#)]
87. Arbesu, M.; Maffei, M.; Cordeiro, T.N.; Teixeira, J.M.; Perez, Y.; Bernado, P.; Roche, S.; Pons, M. The Unique Domain Forms a Fuzzy Intramolecular Complex in Src Family Kinases. *Structure* **2017**, *25*, 630–640. [[CrossRef](#)]

88. Brookes, D.H.; Head-Gordon, T. Experimental Inferential Structure Determination of Ensembles for Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2016**, *138*, 4530–4538. [[CrossRef](#)]
89. Boomsma, W.; Ferkinghoff-Borg, J.; Lindorff-Larsen, K. Combining experiments and simulations using the maximum entropy principle. *PLoS Comput. Biol.* **2014**, *10*, e1003406. [[CrossRef](#)]
90. Lazar, T.; Martinez-Perez, E.; Quaglia, F.; Hatos, A.; Chemes, L.B.; Iserte, J.A.; Mendez, N.A.; Garrone, N.A.; Saldano, T.E.; Marchetti, J.; et al. PED in 2021: A major update of the protein ensemble database for intrinsically disordered proteins. *Nucleic Acids Res.* **2021**, *49*, D404–D411. [[CrossRef](#)]
91. Miskei, M.; Antal, C.; Fuxreiter, M. FuzDB: Database of fuzzy complexes, a tool to develop stochastic structure-function relationships for protein complexes and higher-order assemblies. *Nucleic Acids Res.* **2017**, *45*, D228–D235. [[CrossRef](#)] [[PubMed](#)]
92. Horvath, A.; Miskei, M.; Ambrus, V.; Vendruscolo, M.; Fuxreiter, M. Sequence-based prediction of protein binding mode landscapes. *PLoS Comput. Biol.* **2020**, *16*, e1007864. [[CrossRef](#)] [[PubMed](#)]
93. Miskei, M.; Horvath, A.; Vendruscolo, M.; Fuxreiter, M. Sequence-Based Prediction of Fuzzy Protein Interactions. *J. Mol. Biol.* **2020**, *432*, 2289–2303. [[CrossRef](#)] [[PubMed](#)]
94. Asakawa, H.; Ikegami, K.; Setou, M.; Watanabe, N.; Tsukada, M.; Fukuma, T. Submolecular-scale imaging of alpha-helices and C-terminal domains of tubulins by frequency modulation atomic force microscopy in liquid. *Biophys. J.* **2011**, *101*, 1270–1276. [[CrossRef](#)]
95. Laurin, Y.; Eyer, J.; Robert, C.H.; Prevost, C.; Sacquin-Mora, S. Mobility and Core-Protein Binding Patterns of Disordered C-Terminal Tails in beta-Tubulin Isoforms. *Biochemistry* **2017**, *56*, 1746–1756. [[CrossRef](#)]
96. Koukos, P.I.; Bonvin, A. Integrative Modelling of Biomolecular Complexes. *J. Mol. Biol.* **2020**, *432*, 2861–2881. [[CrossRef](#)]
97. Yang, S.; Bernado, P. Integrative Biophysics: Protein Interaction and Disorder. *J. Mol. Biol.* **2020**, *432*, 2843–2845. [[CrossRef](#)]
98. Dudas, E.F.; Palfy, G.; Menyhard, D.K.; Sebak, F.; Ecsedi, P.; Nyitray, L.; Bodor, A. Tumor-Suppressor p53TAD(1-60) Forms a Fuzzy Complex with Metastasis-Associated S100A4: Structural Insights and Dynamics by an NMR/MD Approach. *ChemBioChem* **2020**, *21*, 3087–3095. [[CrossRef](#)]
99. Spreitzer, E.; Usluer, S.; Madl, T. Probing Surfaces in Dynamic Protein Interactions. *J. Mol. Biol.* **2020**, *432*, 2949–2972. [[CrossRef](#)]
100. Tsytlonok, M.; Hemmen, K.; Hamilton, G.; Kolimi, N.; Felekyan, S.; Seidel, C.A.M.; Tompa, P.; Sanabria, H. Specific Conformational Dynamics and Expansion Underpin a Multi-Step Mechanism for Specific Binding of p27 with Cdk2/Cyclin A. *J. Mol. Biol.* **2020**, *432*, 2998–3017. [[CrossRef](#)] [[PubMed](#)]
101. Grawert, T.W.; Svergun, D.I. Structural Modeling Using Solution Small-Angle X-ray Scattering (SAXS). *J. Mol. Biol.* **2020**, *432*, 3078–3092. [[CrossRef](#)] [[PubMed](#)]
102. Bonomi, M.; Camilloni, C.; Vendruscolo, M. Metadynamic metainference: Enhanced sampling of the metainference ensemble using metadynamics. *Sci. Rep.* **2016**, *6*, 31232. [[CrossRef](#)]
103. Bonomi, M.; Pellarin, R.; Vendruscolo, M. Simultaneous Determination of Protein Structure and Dynamics Using Cryo-Electron Microscopy. *Biophys. J.* **2018**, *114*, 1604–1613. [[CrossRef](#)] [[PubMed](#)]
104. Brotzakis, Z.F.; Lindstedt, P.R.; Taylor, R.; Bernardes, G.J.L.; Vendruscolo, M. A Structural Ensemble of a Tau-Microtubule Complex Reveals Regulatory Tau Phosphorylation and Acetylation Mechanisms. *bioRxiv* **2020**. [[CrossRef](#)]
105. Kellogg, E.H.; Hejab, N.M.A.; Poepsel, S.; Downing, K.H.; DiMaio, F.; Nogales, E. Near-atomic model of microtubule-tau interactions. *Science* **2018**, *360*, 1242–1246. [[CrossRef](#)]
106. Ramanathan, A.; Ma, H.; Parvatikar, A.; Chennubhotla, S.C. Artificial intelligence techniques for integrative structural biology of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2021**, *66*, 216–224. [[CrossRef](#)]
107. Ramanathan, A.; Parvatikar, A.; Chennubhotla, S.C.; Mei, Y.; Sinha, S.C. Transient Unfolding and Long-Range Interactions in Viral BCL2 M11 Enable Binding to the BECN1 BH3 Domain. *Biomolecules* **2020**, *10*, 1308. [[CrossRef](#)]
108. Demerdash, O.; Shrestha, U.R.; Petridis, L.; Smith, J.C.; Mitchell, J.C.; Ramanathan, A. Using Small-Angle Scattering Data and Parametric Machine Learning to Optimize Force Field Parameters for Intrinsically Disordered Proteins. *Front. Mol. Biosci.* **2019**, *6*, 64. [[CrossRef](#)]
109. Lazar, T.; Guharoy, M.; Vranken, W.; Rauscher, S.; Wodak, S.J.; Tompa, P. Distance-Based Metrics for Comparing Conformational Ensembles of Intrinsically Disordered Proteins. *Biophys. J.* **2020**, *118*, 2952–2965. [[CrossRef](#)] [[PubMed](#)]
110. Weinzierl, R.O.J. Molecular Dynamics Simulations of Human FOXO3 Reveal Intrinsically Disordered Regions Spread Spatially by Intramolecular Electrostatic Repulsion. *Biomolecules* **2021**, *11*, 856. [[CrossRef](#)]
111. Cuevas-Velazquez, C.L.; Dinneny, J.R. Organization out of disorder: Liquid-liquid phase separation in plants. *Curr. Opin. Plant Biol.* **2018**, *45*, 68–74. [[CrossRef](#)] [[PubMed](#)]
112. Darling, A.L.; Liu, Y.; Oldfield, C.J.; Uversky, V.N. Intrinsically Disordered Proteome of Human Membrane-Less Organelles. *Proteomics* **2018**, *18*, e1700193. [[CrossRef](#)] [[PubMed](#)]
113. Borgia, A.; Borgia, M.B.; Bugge, K.; Kissling, V.M.; Heidarsson, P.O.; Fernandes, C.B.; Sottini, A.; Soranno, A.; Buholzer, K.J.; Nettels, D.; et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018**, *555*, 61–66. [[CrossRef](#)] [[PubMed](#)]
114. Khazanov, N.; Levy, Y. Sliding of p53 along DNA can be modulated by its oligomeric state and by cross-talks between its constituent domains. *J. Mol. Biol.* **2011**, *408*, 335–355. [[CrossRef](#)] [[PubMed](#)]
115. Vuzman, D.; Levy, Y. Intrinsically disordered regions as affinity tuners in protein-DNA interactions. *Mol. Biosyst.* **2012**, *8*, 47–57. [[CrossRef](#)] [[PubMed](#)]

116. Shishmarev, D.; Wang, Y.; Mason, C.E.; Su, X.C.; Oakley, A.J.; Graham, B.; Huber, T.; Dixon, N.E.; Otting, G. Intramolecular binding mode of the C-terminus of Escherichia coli single-stranded DNA binding protein determined by nuclear magnetic resonance spectroscopy. *Nucleic Acids Res.* **2014**, *42*, 2750–2757. [[CrossRef](#)]
117. Mondal, A.; Bhattacharjee, A. Mechanism of Dynamic Binding of Replication Protein A to ssDNA. *J. Chem. Inf. Model.* **2020**, *60*, 5057–5069. [[CrossRef](#)]
118. Shereda, R.D.; Bernstein, D.A.; Keck, J.L. A central role for SSB in Escherichia coli RecQ DNA helicase function. *J. Biol. Chem.* **2007**, *282*, 19247–19258. [[CrossRef](#)]
119. Marceau, A.H.; Bahng, S.; Massoni, S.C.; George, N.P.; Sandler, S.J.; Marians, K.J.; Keck, J.L. Structure of the SSB-DNA polymerase III interface and its role in DNA replication. *EMBO J.* **2011**, *30*, 4236–4247. [[CrossRef](#)]
120. Protopopova, A.D.; Litvinov, R.I.; Galanakis, D.K.; Nagaswami, C.; Barinov, N.A.; Mukhitov, A.R.; Klinov, D.V.; Weisel, J.W. Morphometric characterization of fibrinogen's alphaC regions and their role in fibrin self-assembly and molecular organization. *Nanoscale* **2017**, *9*, 13707–13716. [[CrossRef](#)]
121. Bigman, L.S.; Levy, Y. Modulating Microtubules: A Molecular Perspective on the Effects of Tail Modifications. *J. Mol. Biol.* **2021**, *433*, 166988. [[CrossRef](#)] [[PubMed](#)]
122. Bigman, L.S.; Levy, Y. Tubulin tails and their modifications regulate protein diffusion on microtubules. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 8876–8883. [[CrossRef](#)] [[PubMed](#)]
123. Bhattacharyya, B.; Sackett, D.L.; Wolff, J. Tubulin, hybrid dimers, and tubulin S. Stepwise charge reduction and polymerization. *J. Biol. Chem.* **1985**, *260*, 10208–10216. [[CrossRef](#)]
124. Fan, H.F.; Su, S. The regulation mechanism of the C-terminus of RecA proteins during DNA strand-exchange process. *Biophys. J.* **2021**, *120*, 3166–3179. [[CrossRef](#)]
125. Ferreira, D.U.; Komives, E.A.; Wolynes, P.G. Frustration in biomolecules. *Q. Rev. Biophys.* **2014**, *47*, 285–363. [[CrossRef](#)]
126. Carvaillo, J.-C. *From Assembly Unit to Capsid: In Silico Application to Norovirus and Hepatitis B Virus*; Université Paris-Saclay: Gif-sur-Yvette, France, 2021.
127. Bell, J.C.; Kowalczykowski, S.C. RecA: Regulation and Mechanism of a Molecular Search Engine. *Trends Biochem. Sci.* **2016**, *41*, 491–507. [[CrossRef](#)]
128. Lusetti, S.L.; Shaw, J.J.; Cox, M.M. Magnesium ion-dependent activation of the RecA protein involves the C terminus. *J. Biol. Chem.* **2003**, *278*, 16381–16388. [[CrossRef](#)]
129. Kim, R.; Kanamaru, S.; Mikawa, T.; Prevost, C.; Ishii, K.; Ito, K.; Uchiyama, S.; Oda, M.; Iwasaki, H.; Kim, S.K.; et al. RecA requires two molecules of Mg²⁺ ions for its optimal strand exchange activity in vitro. *Nucleic Acids Res.* **2018**, *46*, 2548–2559. [[CrossRef](#)]
130. Kurumizaka, H.; Aihara, H.; Ikawa, S.; Kashima, T.; Bazemore, L.R.; Kawasaki, K.; Sarai, A.; Radding, C.M.; Shibata, T. A possible role of the C-terminal domain of the RecA protein. A gateway model for double-stranded DNA binding. *J. Biol. Chem.* **1996**, *271*, 33515–33524. [[CrossRef](#)] [[PubMed](#)]
131. Chen, Z.; Yang, H.; Pavletich, N.P. Mechanism of homologous recombination from the RecA–ssDNA/dsDNA structures. *Nature* **2008**, *453*, 489–494. [[CrossRef](#)]
132. Yang, D.; Boyer, B.; Prevost, C.; Danilowicz, C.; Prentiss, M. Integrating multi-scale data on homologous recombination into a new recognition mechanism based on simulations of the RecA–ssDNA/dsDNA structure. *Nucleic Acids Res.* **2015**, *43*, 10251–10263. [[CrossRef](#)] [[PubMed](#)]
133. Boyer, B.; Danilowicz, C.; Prentiss, M.; Prevost, C. Weaving DNA strands: Structural insight on ATP hydrolysis in RecA-induced homologous recombination. *Nucleic Acids Res.* **2019**, *47*, 7798–7808. [[CrossRef](#)] [[PubMed](#)]