



Article

Balanced Force Field *ff03CMAP* Improving the Dynamics Conformation Sampling of Phosphorylation Site

Bozitao Zhong^{1,†} , Ge Song^{1,†} and Hai-Feng Chen^{1,2,*} 

¹ State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, Department of Bioinformatics and Biostatistics, National Experimental Teaching Center for Life Sciences and Biotechnology, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

² Shanghai Center for Bioinformation Technology, Shanghai 200240, China

* Correspondence: haifengchen@sjtu.edu.cn; Tel./Fax: +86-21-34204073

† These authors contributed equally to this work.

Abstract: Phosphorylation plays a key role in plant biology, such as the accumulation of plant cells to form the observed proteome. Statistical analysis found that many phosphorylation sites are located in disordered regions. However, current force fields are mainly trained for structural proteins, which might not have the capacity to perfectly capture the dynamic conformation of the phosphorylated proteins. Therefore, we evaluated the performance of *ff03CMAP*, a balanced force field between structural and disordered proteins, for the sampling of the phosphorylated proteins. The test results of 11 different phosphorylated systems, including dipeptides, disordered proteins, folded proteins, and their complex, indicate that the *ff03CMAP* force field can better sample the conformations of phosphorylation sites for disordered proteins and disordered regions than *ff03*. For the solvent model, the results strongly suggest that the *ff03CMAP* force field with the *TIP4PD* water model is the best combination for the conformer sampling. Additional tests of *CHARMM36m* and *FB18* force fields on two phosphorylated systems suggest that the overall performance of *ff03CMAP* is similar to that of *FB18* and better than that of *CHARMM36m*. These results can help other researchers to choose suitable force field and solvent models to investigate the dynamic properties of phosphorylation proteins.

Keywords: phosphorylation; force field; *ff03CMAP*; disordered proteins



Citation: Zhong, B.; Song, G.; Chen, H.-F. Balanced Force Field *ff03CMAP* Improving the Dynamics Conformation Sampling of Phosphorylation Site. *Int. J. Mol. Sci.* **2022**, *23*, 11285. <https://doi.org/10.3390/ijms231911285>

Academic Editors: Christo Z. Christov and Tatyana Karabancheva-Christova

Received: 12 July 2022

Accepted: 11 September 2022

Published: 25 September 2022

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



Copyright: © 2022 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

Post-translational modification (PTM) is common in nature [1]. It refers to the process by which a chemical group or small protein is covalently bound to an amino acid or protein. PTM plays a key role in mediating cellular signaling [2], regulation [3], and metabolic activities [4]. Therefore, the study of the mechanistic role of post-translational modifications has become a popular research field. Until now, more than 5000 types of PTM have been identified. Among all these types of PTM, the most common and pervasive one is phosphorylation [1]. It is estimated that one-third of all proteins in eukaryotic cells, including human cells, undergo phosphorylation and dephosphorylation mediated by roughly 500 putative kinases and 150 phosphatases. Phosphorylation modifications occur mainly on serine, threonine, and tyrosine, while some other residues can sometimes undergo O- and N-phosphorylation.

Reversible phosphorylation plays an important function in the cell. The reversible phosphorylation process is involved in signal transduction pathways, amplification, integration during the cell cycle, and reactions to external effectors. Moreover, specific phosphorylation controls protein turnover [5], localization (e.g., nuclear transportation [6]), chromatin remodeling [7], and the assembly of protein complexes [8,9]. In these biological processes, phosphorylation often functions by affecting residue–residue interactions.

Compared with the 20 types of canonical residues, phosphorylated residues have novel physicochemical properties and negative charges to form electrostatic interactions. For example, the salt bridge between phosphoserine and lysine can stabilize the helical structure, especially when it is on the N-terminal end of the helix [10].

Previous statistical research found that many phosphorylation sites are located in disordered regions, which are preferred substrates of protein kinases [1]. Intrinsically disordered proteins (or intrinsically disordered regions) refer to peptides (or fragments) that lack a well-defined three-dimensional structure in solution under physiological conditions [11]. Phosphorylation in disordered proteins and regions is involved in the transition between disordered and folded states, changes in the association state, and the activation or deactivation of a protein [12–15]. For example, in the Ets-1 domain, phosphorylation can trigger disordered-to-folded transitions [16]. Recent research found that phosphorylation can also induce folding of the disordered translational inhibitor 4E-BP2 [14,17].

To detect phosphorylated protein structures and their regulation mechanisms, there are a series of experimental and computational methods that can be used. X-ray, NMR (Nuclear Magnetic Resonance), SAXS (Nuclear Magnetic Resonance) and FRET (Förster resonance energy transfer) can be used to detect phosphorylated protein structures [18]. The experimental measurements include atom coordinates, chemical shifts, scalar couplings, and so on. However, these experimental methods only reflect the average structural insights and have difficulty in tracing the motion of proteins, thus have the limitation of revealing mechanisms at the atomistic level.

Phosphorylation also plays a key role in plant biology [19]. In plant cells, many proteins undergo post-translational modifications and form the mature proteoforms that accumulate to form the observed proteome [20]. Plant PTM Viewer, a PTM database for plants, contains 437,404 PTMs matching 103,992 proteins from plants, of which 74.7% of modifications are phosphorylations [21]. In Arabidopsis, at least 7603 nonredundant proteins have been experimentally identified as phosphorylated at 42,649 different sites [19]. Phosphorylation can function as a molecular switch in regulation, and many signaling pathways work through phosphorylation. A well-studied example is the MAP kinase network, consisting of a hierarchy in which MAP triple-kinases phosphorylate and activate MAPK kinases, which in turn phosphorylate and activate MAPKs at conserved Thr and Tyr residues [22].

Using molecular dynamics (MD) to study the dynamical properties is an important complement to experimental methods [12]. Recently, many studies have used molecular dynamics simulations to study structural change via phosphorylation. Parameters for major types of phosphorylation (phosphorylated serine, threonine, and tyrosine) can be found in most well-known force fields. In the AMBER force field series, the ff9x family contains phosphorylated parameters from Homeyer et al. [23]. The AMBER ff03 family has the post-translational modification parameter *Forcefield_PTm* from Duan et al. [24]. Another popular force field family, CHARMM, includes phosphorylated parameters for version 27, 36 and 36m [25,26]. The Gromos force field family also includes two sets of modified residue parameters [27–29]. These force fields have already been used in modified protein structural research.

The accuracy of chosen force-field and phosphorylation parameters is the key for better dynamic simulation results. Unfortunately, recent studies carried out on phosphorylated residue parameters showed that the simulation result did not fit well with experimental data for dipeptides or disordered proteins. Vymetal et al. simulated the dipeptide structures of the top three phosphorylated amino acids with three commonly used force fields with their corresponding parameters for phosphorylation, the ff99SBildn and ff03 in AMBER, and CHARMM36m. Results showed that dihedral angle distributions and the trends of the secondary structure changes for modifications were inconsistent [30]. Both the ratios in the α -helical, β -sheet and disordered regions, and the distribution of dihedral angles φ/ψ , showed significant inconsistency among the three force fields. For J-coupling constants, the results showed that the CHARMM36m force field had a greater difference from the

experimental value compared to the other two force fields. Furthermore, all three force fields were not accurate enough. In another study, simulation of the SN15 peptides was carried out with *ff99SBildn* and *CHARMM36m*. The results of *CHARMM36m* were worse than those of *ff99SBildn* for SAXS values. However, the results of the AMBER force field were not perfect where the proportion of α -helix is higher than the actual conformation [31].

Considering that most of the popular major force fields (*ff99SBildn* [32], *ff03* [33] and *ff14SB* [34]) are more suitable for folded proteins, while phosphorylation sites are more likely to be located in disordered regions, we conjecture that the balanced force field *ff03CMAP* [35] might outperform other force fields in the simulation of phosphorylated proteins. Compared with the *ff03* force field, *ff03CMAP* adds a correction map (CMAP) parameter to adjust the dihedral angle potential energy, which affects the conformation of the proteins in simulation, achieving improvements in the simulation of disordered proteins without affecting the accuracy of folded proteins [35]. The CMAP method has been widely used to improve the accuracy of force fields [36,37], especially for disordered proteins [38–44]. The solvent model also affects protein conformations in simulations, therefore we also tested two water models of *TIP4P-Ew* [45] and *TIP4P-D* [46] with the two force fields to show a combination effect on phosphorylated protein simulation. To fully cover all kinds of phosphorylated proteins, we chose test systems ranging from dipeptides, disordered proteins, folded proteins, and disordered-folded protein complexes. This diverse test set can adequately cover most types of phosphorylated proteins. The results indicate that the combination between *ff03CMAP* and *TIP4PD* can better sample the conformations of phosphorylation sites for disordered proteins or disordered regions than *ff03*.

2. Material and Methods

2.1. Phosphorylated Protein Structural Propensity

At first, statistical analysis was carried out for all phosphorylated proteins in the PDB database [47]. This analysis was performed to determine the propensity of modification sites, which can help us to better understand the characteristics of the PTM protein structure.

In this statistical analysis, all proteins with phosphorylated serine, threonine, and tyrosine were collected. For each type of phosphorylation, the amount of phosphorylated protein and phosphorylated sites were calculated. We also compared the statistics with former research on the Swiss-Prot database [1]. We used the DSSP program [48,49] to assign the secondary structure of proteins and phosphorylation sites. In the DSSP prediction results, H, G and I were classified as helices, B, E, T, S were considered as sheets, and blank was considered a disordered region. We calculated the protein secondary structure propensity by calculating the ratio between disordered residues and all residues. Differences between the secondary structure composition in protein and modification sites can present the structural tendency of modification sites.

2.2. Molecular Dynamics Simulation

After understanding the secondary structure tendency of post-translational modification sites, molecular dynamics simulations were conducted on various selected phosphorylated modified proteins, including dipeptides, disordered proteins, folded proteins and complexes of disordered and folded proteins. The test set covered all typical types of phosphorylated proteins. In simulation tests, we used 2 different force fields with 2 different solvent models to compare their performances on phosphorylated proteins. For force fields, we chose the widely-used *ff03* force field and the *ff03CMAP* force field, which has a well-balanced performance between folded and disordered proteins. For solvent models, *TIP4P-Ew* and *TIP4P-D* were chosen. In previous *ff03CMAP* simulation tests, *TIP4P-Ew* performed well in folded proteins while *TIP4P-D* performed better in disordered proteins. With the phosphorylated residue parameters from *Forcefield_PTM* [24], we can compare these 4 different methods in simulating modified peptides/proteins to determine the best force field and solvent model set in phosphorylated protein simulations. In addition, to obtain a broader view of the performance of *ff03* and *ff03CMAP*, we included the tests for two other force fields combined with their recommended water models: *CHARMM36m* [50]

(with *TIP3P*) and *FB18* [51] (with *TIP3PFB*). The simulation systems and conditions are gathered in Table 1.

Table 1. All simulated phosphorylated systems and simulation conditions.

System	Description	Modification Type	Length	Initial Structure	Ions	Number of Waters	Simulation Time (ns)	Number of Tested Force Fields	Number of Trajectories
Dipeptide									
GpSG [52]	Phosphoserine dipeptide	pSer	3	Extended	2 Na ⁺	962	500	4	1
GpTG [52]	Phosphothreonine dipeptide	pThr	3	Extended	2 Na ⁺	811	500	4	1
GpYG [52]	Phosphotyrosine dipeptide	pTyr	3	Extended	2 Na ⁺	1089	500	4	1
Disordered Protein									
TF [53]	Cytoplasmic Tail of Tissue Factor	2*pSer	19	2CEF	2 Na ⁺	2511	100	6	5
Vg [54]	Vitellogenin	pSer	35	2LID	7 Na ⁺	6316	100	4	5
Folded Protein									
β 3 [55]	β 3 cytoplasmic tail	pSer	24	2LKJ	1 Na ⁺	3303	100	4	5
α M [56]	α M cytoplasmic tail	2*pTyr	47	2LJE	2 Na ⁺	3828	100	4	5
Ets1 [57]	Ets1	pSer, pThr	111	2KMD	9 Na ⁺	7644	1000	6	5
p-Ubiquitin [58]	Phosphorylated ubiquitin	pSer	76	5XK4	2 Na ⁺	3613	100	4	1
Complex									
SH2 [59]	SH2 domain of Csk in complex with a phosphopeptide from Cbp	pTyr	137	2RSY	7 Na ⁺	8267	100	4	5
TrkB [60]	FRS2a PTB domain with neurotrophin receptor TrkB	pTyr	138	2MFQ	7 Na ⁺	8756	100	4	5

Initial structures of the G-X-G systems were built in the LEaP module in the AMBER 18 suite [61] and other structures were taken from the PDB database. All simulation systems were neutralized and solvated in octahedron boxes of *TIP4P-Ew* or *TIP4P-D* solvent models with 10 Å to the edge. We checked that the water boxes were big enough to keep the minimum distance between periodic images greater than the cutoff value, as shown in Figure S1. All bonds involving hydrogen atoms were constrained with the SHAKE algorithm [62]. The van der Waals interactions and short-range electrostatic interactions were truncated with 8 Å cutoff. The Particle-Mesh Ewald (PME) algorithm was used to calculate long-range electrostatic interactions [63]. Initial structures were relaxed by 3000 steps of minimization with the steepest descent method, followed by 3000 steps of minimization with the conjugate gradient algorithm. The minimized structures were then heated up to 298.15 K over 50 ps at constant volume, followed by another heating with the same temperature over 100 ps at a constant pressure of 1 bar, and equilibrated over 100 ps at the constant pressure. The production simulations were performed in the pmemd.cuda [64] with simulation time corresponding to Table 1. The simulation temperature was set to room temperature. As postulated in previous research [65,66], ionic concentration had significant impacts on the stability of solute's intramolecular salt bridges and in turn changed the conformation ensemble. In this study, sodium ions were set to neutralized protein charges.

2.3. Evaluation Metrics

To compare the performance of various force fields and solvent models, we used a series of quantification metrics to evaluate simulation results. For dynamical properties, RMSD, RMSF and radius of gyration (Rg) were used for evaluation, and all of these metrics were calculated by CPPTRAJ in AmberTools [67]. For protein structural properties, chemical shifts, secondary structures and contact maps were used in comparison. Chemical shifts were calculated by SPARTA+ [68] for C α , C β , C, N, H α and HN atom types [68]. The detailed experimental data are listed in Table S1. Secondary structure and contact maps were also calculated by CPPTRAJ in AmberTools [67]. The PyMOL molecular visualization system was used to present structures for all simulated proteins [69].

In addition, the quantification metric force field was used for the benchmark for all methods. The force field score was calculated as in Equation (1) [70]

$$\text{FF}_{\text{score}} = \frac{1}{N} \sum_{i=1}^{i=N} \frac{\text{FF}_{\text{rmsd}}}{\text{rmsd}_{\text{norm}}} \quad (1)$$

where N is the number of classes of experimental measurements, FF_{rmsd} is the RMSD of the i th class for simulated and experimental values, and $\text{rmsd}_{\text{norm}}$ is the lowest RMSD of i th class in all force fields. Based on the force field score, which considers the overall properties of a protein, we crafted another metric called the modification site force field score to evaluate the experimental values corresponding to post-translational modification site. The modification site force field score was defined as in Equation (2)

$$\text{ModificationFF}_{\text{score}} = \frac{1}{N} \sum_{i=1}^{i=N} \frac{\text{ModificationFF}_{\text{rmsd}}}{\text{rmsd}_{\text{norm}}} \quad (2)$$

where $\text{ModificationFF}_{\text{rmsd}}$ is the RMSD of the i th class for simulated and experimental values in modified residues, and $\text{rmsd}_{\text{norm}}$ is the lowest RMSD of i th class among all force fields in modified residues. According to this metric, FFscore and ModificationFFscore are always larger than or equal to 1, and 1 is the best score theoretically.

3. Results

3.1. Structural Statistics of Phosphorylated Proteins in PDB Database

As we compared the secondary structure propensity of phosphorylation modifications in PDB, we found distinct secondary structure propensity of the modification sites. Phosphorylated proteins are widely distributed in both disordered and folded proteins. Most of the proteins had a folded conformation and only a small number of proteins were disordered (>50% disordered region) (shown in Figure 1A). In all three kinds of phosphorylated proteins, secondary structures were distributed similarly among the helix, sheet and coil. The helix and sheet take up around 40%, and the coil usually takes up the smallest part at 22~23% (Figure 1B). On the contrary, the secondary structure distribution of phosphorylated sites was significantly different. Most of the phosphorylation sites occurred in disordered regions; 52.18% of phosphotyrosine, 73.44% of phosphoserine, and 89.31% of phosphothreonine were located in disordered regions (Figure 1C). Considering the low percentage of coil structure in the overall structure, there is a clear tendency for phosphorylation to occur in disordered regions.

To sum up, phosphorylated proteins in PDB had a balanced distribution on major secondary structures, but phosphorylation sites tended to be located in disordered regions. Therefore, the precise sampling for disordered regions should be important. The findings suggested that a balanced force field for phosphorylated proteins' simulation might achieve better accuracy in proteins with both folded and disordered regions.

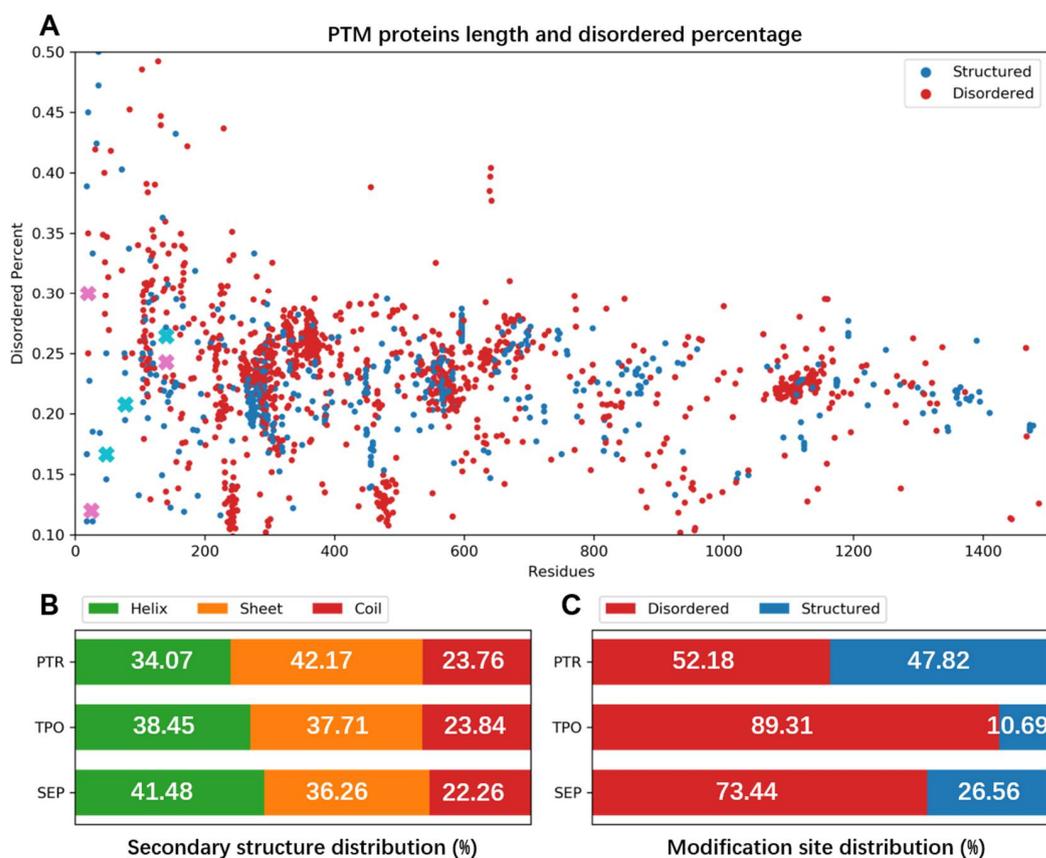


Figure 1. Structural analysis of phosphorylated proteins in PDB. Distributions of phosphorylated proteins' disorder percentage and protein length (A), total protein secondary structure (B) and modification site secondary structure (C). In (A), each point represents a phosphorylated protein in the PDB database, and its color shows whether it had a phosphorylated residue in a disordered region (red) or not (blue). Selected simulation systems are marked by a cross in pink (modification in disordered region) and cyan (modification in folded region). In (B), the white numbers show the three kinds of secondary structures' percentage in all residues in modified proteins. In (C), we present the two kinds (helix and sheet are considered as structured) of secondary structures' percentage in all three kinds of phosphorylation site. (B,C) show that the phosphorylation is usually located in disordered regions.

3.2. Convergence of Simulation System

In order to evaluate the convergence, five independent trajectories of 100 ns each were obtained for each system. Convergence of the simulation needs to be considered first because it reflects the adequacy of the simulation time. We calculated the RMSE between the secondary chemical shifts of experimental value and the cumulative average value over time during the simulation. The results show that the simulation of the representative Ets1 system converges under *ff03CMAP*, *ff03*, *FB18*, and *CHARMM36m* force fields after 100 ns (shown in Figure 2). At the same time, we also compared the performance of sampling between multiple short trajectories (5*100 ns) and a single long trajectory (1000 ns). The results indicate that multiple short trajectories have better sampling ability than a single long trajectory (shown in Figure S2).

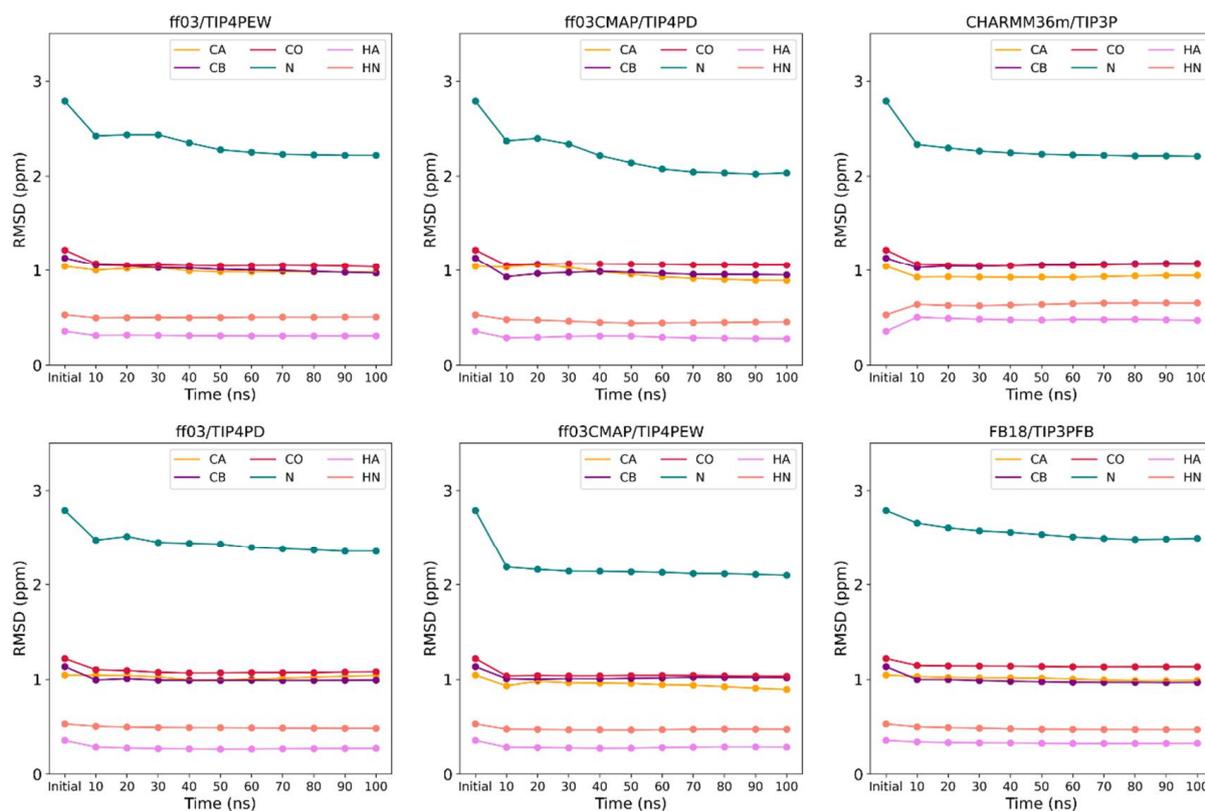


Figure 2. Time-dependent RMS error of chemical shift for Ets1 between experiment data and simulated ones with six force fields.

3.3. Phosphorylated Dipeptides

The simulations of phosphoserine, phosphothreonine and phosphotyrosine dipeptides behaved inconsistently in all four force fields. Using the FFscore as the metric, two force fields with *TIP4Pew* in phosphoserine, *ff03/TIP4Pew* in phosphothreonine, and *ff03CMAP/TIP4PD* in phosphotyrosine had the best performance. None of the four force fields performed optimally in all three types of phosphorylation (shown in Figure 3A).

In addition to chemical shifts, Ramachandran plots in Figure 3B presented the backbone dihedral angle distribution of all simulated dipeptides. TALOS-N [71]-predicted distributions from experimental chemical shifts were used as references. In reference Ramachandran plots, phosphoserine and phosphotyrosine were enriched in the PPII region, while phosphothreonine was enriched in the sheet region. Only simulation results from *ff03CMAP/TIP4PD* had a PPII- and sheet-like distribution, while those from other force fields had high helical propensity in phosphoserine and phosphothreonine.

In general, all force fields have similar accuracy and inconsistent performance in all three types of phosphorylated dipeptides, and *ff03CMAP* has a small advantage over *ff03* in force field scores and dihedral distributions. This is consistent with the previous studies of Vymetal J et al. [30], showing that force fields showed inconsistent performance on phosphorylated dipeptides. Therefore, further evaluation was performed for modified disordered proteins and folded proteins.

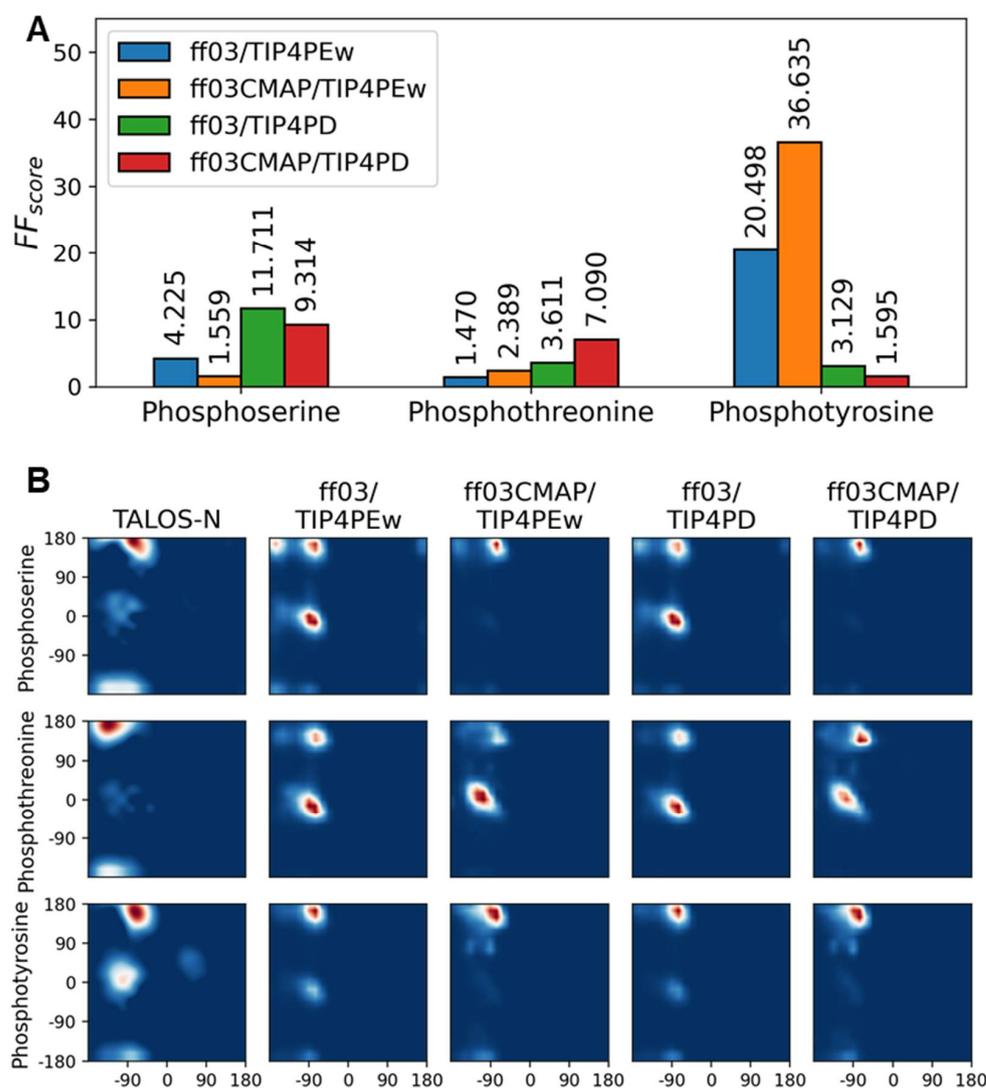


Figure 3. Simulation results of phosphorylated dipeptides. Comparison of force field scores (A) present the inconsistency of different force fields. Predicted and simulated Ramachandran plots (B) showed simulated distribution of backbone dihedral angle distribution difference between experiments and simulations.

3.4. Phosphorylated Disordered Proteins

As discovered in statistical analysis through PDB, most phosphorylation sites were located in disordered regions. Further simulation tests on phosphorylated disordered proteins can serve as paradigms (references) for further (future) research. Two modified disordered proteins, tissue factor cytoplasmic tail (TF, PDB:2CEF) [53] and vitellogenin (Vg, PDB:2LID) [54], were selected as simulation targets.

Figure 4 shows the RMSD of chemical shifts, overall force field score (FFscore) and modification site force field score (Modification Score). Overall force field scores show that *ff03/TIP4PD* and *ff03CMAP/TIP4PD* were the best, while modification site force field scores preferred *ff03CMAP/TIP4PEw*. This elucidated that if we focus on the modification site, the performance of *ff03CMAP* is better than that of the *ff03* force field. Radius of gyration (Rg) and end-to-end distance distribution showed that four force fields have a huge difference in simulated Rg distribution (shown in Figure S3). *ff03/TIP4PEw* has the lowest average Rg which represents a more condensed status, while *ff03CMAP/TIP4PD* has the largest average and the widest range of Rg, which represents the extended structure of disordered proteins. In simulated ensembles' secondary structure, *ff03CMAP* tends to have a more

disordered structure than *ff03*, while the difference between *TIP4PEw* and *TIP4PD* is not significant (shown in Figure S4).

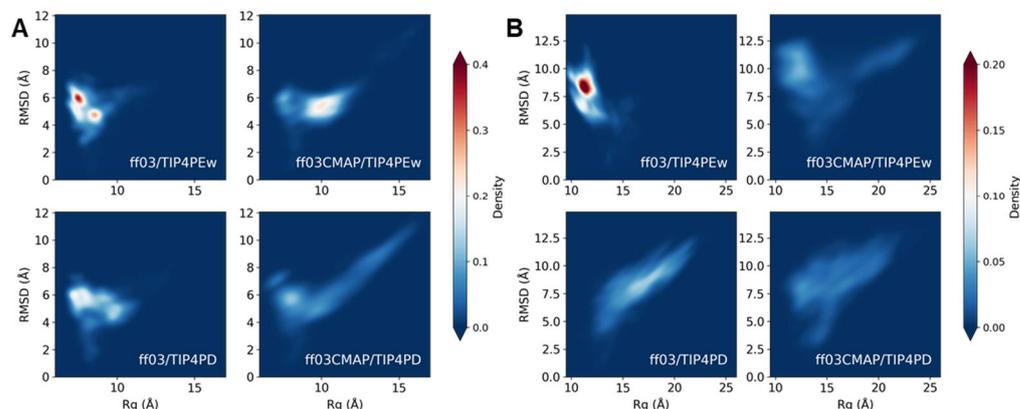


Figure 4. Energy landscape in MD simulation of phosphorylated disordered proteins. Comparison of different force field energy landscapes of phosphorylated disordered protein TF (A) and Vg (B) showed sampling efficiency advantage of *ff03CMAP* force field with both solvent models.

Sampling is another important factor for disordered protein simulation. The RMSD-Rg free energy landscape indicator can help to predict the sampling efficiency of different force fields. For disordered proteins, a wider energy landscape means the simulated structure will be more reasonable. As shown in Figure 3B, the energy landscape of *ff03/TIP4PEw* constrained in several conformations, while other force fields obtained a more adequate sampled conformation. With a better sampled landscape, *ff03CMAP* was able to obtain a more disordered conformation.

We compared the Ramachandran plot between simulation and the TALOS-N prediction from experimental data. The results show that *ff03CMAP/TIP4PD* has a propensity to be distributed in the PPII region, similar to the pattern found in phosphorylated dipeptides. This propensity has a superimposed effect, which means when we use the combination of *ff03CMAP* and *TIP4PD*, it will have a higher proportion in the PPII region. Analysis of hydrogen bonds also showed a difference between *ff03* and *ff03CMAP*, as *ff03CMAP* has less hydrogen bonds in all residues and modification sites (shown in Figure S5). After comparing chemical shifts, force field scores, Rg distributions, secondary structures, and energy landscapes, we can conclude that *ff03CMAP/TIP4PD* is the best system for phosphorylated disordered proteins.

3.5. Phosphorylated Folded Proteins and Complexes

Folded protein takes up the largest part of phosphorylated proteins. In nature, most protein function depends on specific structures, indicating that they are typically folded proteins with a defined, stable tertiary structure. In structural analysis, we found phosphorylation sites were usually located at the loop region (and also disordered regions) in a folded protein, or sometimes at a disordered protein which was bound to a folded protein. Therefore, we chose four folded proteins and two folded-disordered protein complexes as simulation targets for evaluation.

First, we compared the chemical shifts and force field scores. The overall force field scores of four force field systems had similar values and were close to 1. This means that from a holistic perspective, four force fields all have good performance, which is consistent with the previous works [35]. However, the modification site force field score was not unanimous among all four simulation systems. In folded modified proteins and folded-disordered protein complexes, *ff03CMAP* obviously had a better modification force field score than the *ff03* force field. In 2LKJ and 2LJE, *ff03CMAP/TIP4PEw* had the lowest modification site force field score, and in 2KMD and 5XK4, *ff03CMAP/TIP4PD* had the lowest modification site force field score.

In order to compare, two representative proteins were chosen for phosphorylated Ets1 (PDB: 2KMD) [57] and Ubiquitin (PDB: 5XK4) [58]. Ets1 is a reaction component of the Ras system. The two phosphorylated residues on Ets1 N terminal disordered region provide this motif with negative charge and repel its neighbor H0 motif from the negatively charged protein surface [57]. Therefore, the properties of the H0 helix should be key information for the simulation. In Figure 5A, we find that the secondary chemical shift for the H0 helix with *ff03CMAP/TIP4PD* has the best agreement with the experimental one among these force fields. RMSF (root mean square fluctuation) also shows that *ff03CMAP* has a relatively higher fluctuation than *ff03* simulated ensembles (shown in Figure S6). Additionally, the hydrogen bond analysis of Ets1 shows that *ff03CMAP/TIP4PD* had the least hydrogen bonds connected with two modification sites (shown in Figure S7). All of these results reflect that *ff03CMAP/TIP4PD* have a higher disorder propensity than other force fields, having secondary chemical shifts closer to zero, fewer hydrogen bonds and higher RMSF.

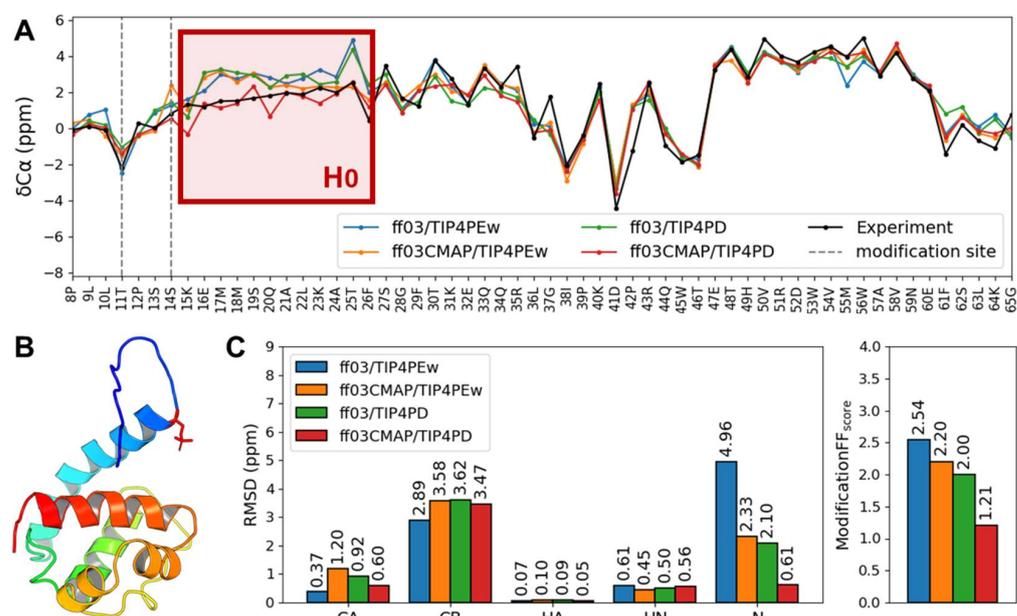


Figure 5. Simulation results of phosphorylated folded protein Ets1. Secondary chemical shifts of Ets1 around H0 region (A) showed only *ff03CMAP/TIP4PD* simulated the H0 helix best, whose conformation was greatly affected by close-site phosphorylations. Ets1 structure and its phosphorylation sites are shown in (B). Chemical shift RMSD of modified residues and modification force field scores (C) also support that *ff03CMAP/TIP4PD* performed best in phosphorylated residues.

For phosphorylated ubiquitin, multiple sites of phosphorylation can act as a pH sensor in the cell [58]. These multi-phosphorylated ubiquitin can be found around mitochondria. In phosphorylated ubiquitin, its conformational switch between protonated and unprotonated form plays a key role in its function, and we chose unprotonated single phosphorylated ubiquitin as the target-simulated protein. The result of chemical shifts and force field scores indicated that phosphorylated ubiquitin had low force field scores and was similar to the folded protein Ets1. However, at the modified serine on the disordered region, the modification site force field score of *ff03CMAP/TIP4PD* was obviously better than other force fields (Figure 5B). In addition, concerning the Ramachandran plot of modification sites, we also discovered that *ff03/TIP4PEw*, which were the benchmark methods in previous research, formed an unusual distribution of dihedral angles (shown in Figure S8).

We also tested the performance of CHARMM36m [50] and FB18 [51] force fields on the folded protein Ets1 and the disordered protein TF, the chemical shift RMSD and force field scores of which are shown in Figures S9 and S10. The results show that FB18 had better performance than CHARMM36m for both systems, and close with the performance of *ff03CMAP*. For the disordered protein TF, *ff03CMAP/TIP4PD* had the lowest general force

field score, while FB18 had the lowest modification force field score. For folded protein Ets-1, all force fields except for CHARMM36m had a relative optimal general force field score (less than 1.1), while ff03CMAP/TIP4PD had the lowest modification force field score.

4. Discussion

In our research, the main goal is to find an ideal force field for molecular dynamics simulations of post-translationally modified proteins. First, we studied the structural features of post-translationally modified proteins by analyzing the PDB protein structure database. We found that phosphorylation modifications, which occur most predominantly in post-translational modifications, have a distinct tendency to occur in the disordered regions of the protein. The conventional force fields for folded proteins sometimes might have inaccurate simulations of disordered proteins or disordered regions. Furthermore, previous studies have shown that existing force fields for post-translational modification are not accurate and consistent enough to simulate post-translationally modified peptides or proteins [29]. Therefore, we introduced the balanced *ff03CMAP* force field, which can perform well in molecular dynamics simulations of both folded and disordered proteins, to compare with the conventional *ff03* force field.

By simulating a series of phosphorylated dipeptides, disordered proteins, folded proteins, and disordered-folded protein complexes, we summarized the results of the simulations of several force fields and focused on the discussion of the *ff03CMAP* force field and *ff03* force field with *TIP4PEw* and *TIP4PD* solvent models (shown in Figure 6).

Force field		ff03 TIP4PEw	ff03CMAP TIP4PEw	ff03 TIP4PD	ff03CMAP TIP4PD	
General FF _{SCORE}	disorder	2CEF	1.29	1.21	1.24	1.11
		2LID	1.45	1.18	1.09	1.22
	folded	2LKJ	1.04	1.17	1.16	1.25
		2LJE	1.10	1.32	1.03	1.52
		2KMD	1.06	1.00	1.10	1.06
		5XK4	1.03	1.05	1.06	1.04
	complex	2RSY	1.00	1.03	1.03	1.03
		2MFQ	1.11	1.00	1.06	1.02
	average		1.14	1.12	1.10	1.16
	Modification FF _{SCORE}	dipeptide	GSG	4.23	1.56	11.71
GTG			1.47	2.39	3.61	7.09
GYG			20.49	36.63	3.13	1.6
disorder		2CEF	2.11	1.2	2.84	1.76
		2LID	6.33	1.54	4.13	3.78
folded		2LKJ	2.70	1.91	4.44	4.91
		2LJE	7.59	2.53	5.75	8.81
		2KMD	37.22	18.39	16.36	1.58
		5XK4	6.42	3.29	2.05	1.58
complex		2RSY	3.00	1.35	2.08	1.16
		2MFQ	3.42	6.75	4.18	9.52
average		8.63	7.05	5.48	4.65	

Figure 6. Overall force field score and modification site force field score of all simulated systems. Green represents a General FFscore lower than 1.2 and Modification score lower than 2. Light green represents a Modification FFscore between 2 and 5. Yellow represents a General FFscore from 1.2 to 1.3 and Modification FFscore from 5 to 10. Red represents a General FFscore higher than 1.3 and Modification Ffscore higher than 10.

The comparison with force fields, including *CHARMM36m* and *FB18* on folded proteins, indicates that the *ForceField_PTM* compatible *ff03* force field family is ideal for the simulation of phosphorylated proteins in a broader view of force fields. For the *ff03* force field family, the *ff03CMAP* force field showed improvement in simulation results of phosphorylated disordered proteins with less difference from experimental values compared with the *ff03* force field, and *TIP4PD* was better between the two solvent models. In folded proteins, the two force fields performed similarly, but when the post-translational modification site was located in the disordered region, the modification site simulation results were closer to the experimental values using the *ff03CMAP* force field. In the disordered-folded protein complexes, the test results were similar to those of the folded protein, and the post-translational modification site on the occurring disordered protein was closer to the experimental value in the simulation results of *ff03CMAP*. Overall, the simulations using the *ff03CMAP* force field outperformed the *ff03* force field when simulating biomolecules with post-translational modification sites located on disordered proteins or disordered regions.

Based on previous studies on potential charge in post-translational modification sites, it has been widely accepted that charge is an important component in the process of modeling post-translational modification sites [24,70]. In this research, our analysis reveals that the vast majority of phosphorylation occurs in the disordered region of the protein, which leads us to consider introducing CMAP energy for dihedral angles to improve the simulation of phosphorylated proteins. CMAP energy has been widely used to simulate disordered proteins [38–44] and can also be used to improve the balanced force field for disordered and structural proteins [35]. We have tried using the balanced force field to simulate phosphorylated proteins with both structured and disordered regions, and indeed the simulation of the modification sites in the disordered region was more accurate than other force field. This further illustrates that the methods used to develop disordered protein force fields can be used to improve the simulation of phosphorylated proteins.

However, the *ff03CMAP* force field also has limitations. Considering the simulation results of dipeptides, we found that neither the *ff03* force field nor the *ff03CMAP* force field was the consistent result for all three post-translational modifications. Similarly, in the protein simulations, we found that the dihedral angle distribution at the post-translational modification sites still notably differed from the predicted dihedral angle distribution. *ff03CMAP* still needs further improvement in the dihedral angle distribution at the post-translational modification sites.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms231911285/s1>.

Author Contributions: Conceptualization, B.Z. and H.-F.C.; methodology, B.Z. and G.S.; software, B.Z. and G.S.; validation, B.Z. and G.S.; formal analysis, B.Z. and G.S.; investigation, B.Z. and G.S.; data curation, B.Z. and G.S.; writing—original draft preparation, B.Z. and G.S.; writing—review and editing, H.-F.C.; visualization, H.-F.C.; supervision, H.-F.C.; project administration, H.-F.C.; funding acquisition, Y.H.-F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported by the Center for HPC at Shanghai Jiao Tong University, the National Key Research and Development Program of China (2020YFA0907700) and the National Natural Science Foundation of China (21977068 and 32171242).

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

References

1. Khoury, G.A.; Baliban, R.C.; Floudas, C.A. Proteome-Wide Post-Translational Modification Statistics: Frequency Analysis and Curation of the Swiss-Prot Database. *Sci. Rep.* **2011**, *1*, 90. [CrossRef] [PubMed]
2. Ardito, F.; Giuliani, M.; Perrone, D.; Troiano, G.; Muzio, L.L. The Crucial Role of Protein Phosphorylation in Cell Signaling and Its Use as Targeted Therapy (Review). *Int. J. Mol. Med.* **2017**, *40*, 271–280. [CrossRef] [PubMed]

3. Terman, J.R.; Kashina, A. Post-Translational Modification and Regulation of Actin. *Curr. Opin. Cell Biol.* **2013**, *25*, 30–38. [[CrossRef](#)] [[PubMed](#)]
4. Jenuwein, T.; Allis, C.D. Translating the Histone Code. *Science* **2001**, *293*, 1074–1080. [[CrossRef](#)]
5. Zecha, J.; Gabriel, W.; Spallek, R.; Chang, Y.-C.; Mergner, J.; Wilhelm, M.; Bassermann, F.; Kuster, B. Linking Post-Translational Modifications and Protein Turnover by Site-Resolved Protein Turnover Profiling. *Nat. Commun.* **2022**, *13*, 165. [[CrossRef](#)]
6. Appella, E.; Anderson, C.W. Post-Translational Modifications and Activation of P53 by Genotoxic Stresses. *Eur. J. Biochem.* **2001**, *268*, 2764–2772. [[CrossRef](#)]
7. Govin, J.; Caron, C.; Lestrat, C.; Rousseaux, S.; Khochbin, S. The Role of Histones in Chromatin Remodelling during Mammalian Spermiogenesis. *Eur. J. Biochem.* **2004**, *271*, 3459–3469. [[CrossRef](#)]
8. Almeida, K.H.; Sobol, R.W. A Unified View of Base Excision Repair: Lesion-Dependent Protein Complexes Regulated by Post-Translational Modification. *DNA Repair* **2007**, *6*, 695–711. [[CrossRef](#)]
9. Csizmok, V.; Forman-Kay, J.D. Complex Regulatory Mechanisms Mediated by the Interplay of Multiple Post-Translational Modifications. *Curr. Opin. Struct. Biol.* **2018**, *48*, 58–67. [[CrossRef](#)]
10. Elbaum, M.B.; Zondlo, N.J. OGIcNAcylation and Phosphorylation Have Similar Structural Effects in α -Helices: Post-Translational Modifications as Inducible Start and Stop Signals in α -Helices, with Greater Structural Effects on Threonine Modification. *Biochemistry* **2014**, *53*, 2242–2260. [[CrossRef](#)]
11. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W.; et al. Intrinsically Disordered Protein. *J. Mol. Graph. Model.* **2001**, *19*, 26–59. [[CrossRef](#)]
12. Bah, A.; Forman-Kay, J.D. Modulation of Intrinsically Disordered Protein Function by Post-Translational Modifications. *J. Biol. Chem.* **2016**, *291*, 6696–6705. [[CrossRef](#)] [[PubMed](#)]
13. Owen, I.; Shewmaker, F. The Role of Post-Translational Modifications in the Phase Transitions of Intrinsically Disordered Proteins. *Int. J. Mol. Sci.* **2019**, *20*, 5501. [[CrossRef](#)]
14. Bah, A.; Vernon, R.M.; Siddiqui, Z.; Krzeminski, M.; Muhandiram, R.; Zhao, C.; Sonenberg, N.; Kay, L.E.; Forman-Kay, J.D. Folding of an Intrinsically Disordered Protein by Phosphorylation as a Regulatory Switch. *Nature* **2015**, *519*, 106–109. [[CrossRef](#)]
15. Wu, J.; Li, D.; Liu, X.; Li, Q.; He, X.; Wei, J.; Li, M.; Rehman, A.U.; Xia, Y.; Wu, C.; et al. IDDB: A Comprehensive Resource Featuring Genes, Variants and Characteristics Associated with Infertility. *Nucleic Acids Res.* **2021**, *49*, D1218–D1224. [[CrossRef](#)]
16. Slupsky, C.M.; Gentile, L.N.; Donaldson, L.W.; Mackereth, C.D.; Seidel, J.J.; Graves, B.J.; McIntosh, L.P. Structure of the Ets-1 Pointed Domain and Mitogen-Activated Protein Kinase Phosphorylation Site. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12129–12134. [[CrossRef](#)] [[PubMed](#)]
17. Zeng, J.; Jiang, F.; Wu, Y. Mechanism of Phosphorylation-Induced Folding of 4E-BP2 Revealed by Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 320–328. Available online: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b00848?casa_token=C_nyHiGjT5kAAAAA:hbjhgAWkh2t36_rprkez0_2tUqifD0M5tOpTMqg0fMBfLYdIEt0iq2U5dvI5lji5LSyLcHtifeKX1Bi- (accessed on 16 April 2022). [[CrossRef](#)]
18. Humphrey, S.J.; James, D.E.; Mann, M. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Trends Endocrinol. Metab.* **2015**, *26*, 676–687. [[CrossRef](#)]
19. Millar, A.H.; Heazlewood, J.L.; Giglione, C.; Holdsworth, M.J.; Bachmair, A.; Schulze, W.X. The Scope, Functions, and Dynamics of Posttranslational Protein Modifications. *Annu. Rev. Plant Biol.* **2019**, *70*, 119–151. [[CrossRef](#)]
20. Smith, L.M.; Kelleher, N.L. Proteoform: A Single Term Describing Protein Complexity. *Nat. Methods* **2013**, *10*, 186–187. [[CrossRef](#)]
21. Willems, P.; Horne, A.; Van Parys, T.; Goormachtig, S.; De Smet, I.; Botzki, A.; Van Breusegem, F.; Gevaert, K. The Plant PTM Viewer, a Central Resource for Exploring Plant Protein Modifications. *Plant J.* **2019**, *99*, 752–762. [[CrossRef](#)] [[PubMed](#)]
22. Zulawski, M.; Schulze, W.X. The Plant Kinome. In *Plant Phosphoproteomics: Methods and Protocols*; Methods in Molecular Biology; Schulze, W.X., Ed.; Springer: New York, NY, USA, 2015; pp. 1–23, ISBN 978-1-4939-2648-0.
23. Homeyer, N.; Horn, A.H.; Lanig, H.; Sticht, H. AMBER Force-Field Parameters for Phosphorylated Amino Acids in Different Protonation States: Phosphoserine, Phosphothreonine, Phosphotyrosine, and Phosphohistidine. *J. Mol. Modeling* **2006**, *12*, 281–289. [[CrossRef](#)] [[PubMed](#)]
24. Khoury, G.A.; Thompson, J.P.; Smadbeck, J.; Kieslich, C.A.; Floudas, C.A. Forcefield_PTM: Ab Initio Charge and AMBER Forcefield Parameters for Frequently Occurring Post-Translational Modifications. *J. Chem. Theory Comput.* **2013**, *9*, 5653–5674. [[CrossRef](#)]
25. Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187–217. [[CrossRef](#)]
26. Brooks, B.R.; Brooks, C.L., III; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [[CrossRef](#)] [[PubMed](#)]
27. Margreitter, C.; Petrov, D.; Zagrovic, B. Vienna-PTM Web Server: A Toolkit for MD Simulations of Protein Post-Translational Modifications. *Nucleic Acids Res.* **2013**, *41*, W422–W426. [[CrossRef](#)] [[PubMed](#)]
28. Petrov, D.; Margreitter, C.; Grandits, M.; Oostenbrink, C.; Zagrovic, B. A Systematic Framework for Molecular Dynamics Simulations of Protein Post-Translational Modifications. *PLoS Comput. Biol.* **2013**, *9*, e1003154. [[CrossRef](#)] [[PubMed](#)]
29. Margreitter, C.; Reif, M.M.; Oostenbrink, C. Update on Phosphate and Charged Post-Translationally Modified Amino Acid Parameters in the GROMOS Force Field. *J. Comput. Chem.* **2017**, *38*, 714–720. [[CrossRef](#)]

30. Vymetal, J.; Jurásková, V.; Vondrášek, J. AMBER and CHARMM Force Fields Inconsistently Portray the Microscopic Details of Phosphorylation. *J. Chem. Theory Comput.* **2018**, *15*, 665–679. [[CrossRef](#)]
31. Rieloff, E.; Skepö, M. Phosphorylation of a Disordered Peptide—Structural Effects and Force Field Inconsistencies. *J. Chem. Theory Comput.* **2020**, *16*, 1924–1935. [[CrossRef](#)]
32. 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 Struct. Funct. Bioinform.* **2010**, *78*, 1950–1958. [[CrossRef](#)] [[PubMed](#)]
33. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M.C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012. [[CrossRef](#)]
34. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [[CrossRef](#)] [[PubMed](#)]
35. 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](#)] [[PubMed](#)]
36. Tian, C.; Kasavajhala, K.; Belfon, K.A.A.; Raguette, L.; Huang, H.; Miguez, A.N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; et al. Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16*, 528–552. [[CrossRef](#)]
37. Mackerell, A.D., Jr. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004**, *25*, 1584–1604. [[CrossRef](#)]
38. Song, D.; Luo, R.; Chen, H.-F. The IDP-Specific Force Field Ff14IDPSFF Improves the Conformer Sampling of Intrinsically Disordered Proteins. *J. Chem. Inf. Model.* **2017**, *57*, 1166–1178. [[CrossRef](#)]
39. Chen, J.; Liu, H.; Cui, X.; Li, Z.; Chen, H.-F. RNA-Specific Force Field Optimization with CMAP and Reweighting. *J. Chem. Inf. Model.* **2022**, *62*, 372–385. [[CrossRef](#)]
40. Song, D.; Liu, H.; Luo, R.; Chen, H.-F. Environment-Specific Force Field for Intrinsically Disordered and Ordered Proteins. *J. Chem. Inf. Model.* **2020**, *60*, 2257–2267. [[CrossRef](#)]
41. Yang, S.; Liu, H.; Zhang, Y.; Lu, H.; Chen, H. Residue-Specific Force Field Improving the Sample of Intrinsically Disordered Proteins and Folded Proteins. *J. Chem. Inf. Model.* **2019**, *59*, 4793–4805. [[CrossRef](#)]
42. Liu, H.; Song, D.; Lu, H.; Luo, R.; Chen, H.-F. Intrinsically Disordered Protein-Specific Force Field CHARMM36IDPSFF. *Chem. Biol. Drug Des.* **2018**, *92*, 1722–1735. [[CrossRef](#)]
43. Ye, W.; Ji, D.; Wang, W.; Luo, R.; Chen, H.-F. Test and Evaluation of Ff99IDPs Force Field for Intrinsically Disordered Proteins. *J. Chem. Inf. Model.* **2015**, *55*, 1021–1029. [[CrossRef](#)]
44. Wang, W.; Ye, W.; Jiang, C.; Luo, R.; Chen, H.-F. New Force Field on Modeling Intrinsically Disordered Proteins. *Chem. Biol. Drug Des.* **2014**, *84*, 253–269. [[CrossRef](#)]
45. Horn, H.W.; Swope, W.C.; Pitara, J.W.; Madura, J.D.; Dick, T.J.; Hura, G.L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678. [[CrossRef](#)]
46. 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](#)]
47. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)]
48. Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolym. Orig. Res. Biomol.* **1983**, *22*, 2577–2637. [[CrossRef](#)]
49. Touw, W.G.; Baakman, C.; Black, J.; Te Beek, T.A.; Krieger, E.; Joosten, R.P.; Vriend, G. A Series of PDB-Related Databanks for Everyday Needs. *Nucleic Acids Res.* **2015**, *43*, D364–D368. [[CrossRef](#)]
50. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B.L.; Grubmüller, H.; MacKerell, A.D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71–73. [[CrossRef](#)]
51. Stoppelman, J.P.; Ng, T.T.; Nerenberg, P.S.; Wang, L.-P. Development and Validation of AMBER-FB15-Compatible Force Field Parameters for Phosphorylated Amino Acids. *J. Phys. Chem. B* **2021**, *125*, 11927–11942. [[CrossRef](#)]
52. Bienkiewicz, E.A.; Lumb, K.J. Random-Coil Chemical Shifts of Phosphorylated Amino Acids. *J. Biomol. NMR* **1999**, *15*, 203–206. [[CrossRef](#)] [[PubMed](#)]
53. Sen, M.; Herzik, M., Jr.; Craft, J.W., Jr.; Creath, A.L.; Agrawal, S.; Ruf, W.; Legge, G.B. Spectroscopic Characterization of Successive Phosphorylation of the Tissue Factor Cytoplasmic Region. *Open Spectrosc. J.* **2009**, *3*, 58. [[CrossRef](#)] [[PubMed](#)]
54. Havukainen, H.; Underhaug, J.; Wolschin, F.; Amdam, G.; Halskau, Ø. A Vitellogenin Polyserine Cleavage Site: Highly Disordered Conformation Protected from Proteolysis by Phosphorylation. *J. Exp. Biol.* **2012**, *215*, 1837–1846. [[CrossRef](#)] [[PubMed](#)]
55. Deshmukh, L.; Meller, N.; Alder, N.; Byzova, T.; Vinogradova, O. Tyrosine Phosphorylation as a Conformational Switch a CASE Study of Integrin B3 Cytoplasmic Tail. *J. Biol. Chem.* **2011**, *286*, 40943–40953. [[CrossRef](#)]
56. Chua, G.-L.; Tang, X.-Y.; Amalraj, M.; Tan, S.-M.; Bhattacharjya, S. Structures and Interaction Analyses of Integrin AMβ2 Cytoplasmic Tails. *J. Biol. Chem.* **2011**, *286*, 43842–43854. [[CrossRef](#)] [[PubMed](#)]
57. Nelson, M.L.; Kang, H.-S.; Lee, G.M.; Blaszczyk, A.G.; Lau, D.K.; McIntosh, L.P.; Graves, B.J. Ras Signaling Requires Dynamic Properties of Ets1 for Phosphorylation-Enhanced Binding to Coactivator CBP. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10026–10031. [[CrossRef](#)] [[PubMed](#)]

58. Dong, X.; Gong, Z.; Lu, Y.-B.; Liu, K.; Qin, L.-Y.; Ran, M.-L.; Zhang, C.-L.; Liu, Z.; Zhang, W.-P.; Tang, C. Ubiquitin S65 Phosphorylation Engenders a PH-Sensitive Conformational Switch. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6770–6775. [[CrossRef](#)]
59. Tanaka, H.; Akagi, K.; Oneyama, C.; Tanaka, M.; Sasaki, Y.; Kanou, T.; Lee, Y.-H.; Yokogawa, D.; Dobenecker, M.-W.; Nakagawa, A.; et al. Identification of a New Interaction Mode between the Src Homology 2 Domain of C-Terminal Src Kinase (Csk) and Csk-Binding Protein/Phosphoprotein Associated with Glycosphingolipid Microdomains. *J. Biol. Chem.* **2013**, *288*, 15240–15254. [[CrossRef](#)]
60. Zeng, L.; Kuti, M.; Mujtaba, S.; Zhou, M.-M. Structural Insights into FRS2 α PTB Domain Recognition by Neurotrophin Receptor TrkB. *Proteins Struct. Funct. Bioinform.* **2014**, *82*, 1534–1541. [[CrossRef](#)]
61. Lee, T.-S.; Cerutti, D.S.; Mermelstein, D.; Lin, C.; LeGrand, S.; Giese, T.J.; Roitberg, A.; Case, D.A.; Walker, R.C.; York, D.M. GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J. Chem. Inf. Model.* **2018**, *58*, 2043–2050. [[CrossRef](#)]
62. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H.J.C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [[CrossRef](#)]
63. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092. [[CrossRef](#)]
64. Salomon-Ferrer, R.; Götz, A.W.; Poole, D.; Grand, S.L.; Walker, R.C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888. Available online: <https://pubs.acs.org/doi/abs/10.1021/ct400314y> (accessed on 16 April 2022). [[CrossRef](#)] [[PubMed](#)]
65. Rieloff, E.; Skepö, M. Molecular Dynamics Simulations of Phosphorylated Intrinsically Disordered Proteins: A Force Field Comparison. *Int. J. Mol. Sci.* **2021**, *22*, 10174. [[CrossRef](#)]
66. Rieloff, E.; Skepö, M. The Effect of Multisite Phosphorylation on the Conformational Properties of Intrinsically Disordered Proteins. *Int. J. Mol. Sci.* **2021**, *22*, 11058. [[CrossRef](#)] [[PubMed](#)]
67. Roe, D.R.; Cheatham, T.E., III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [[CrossRef](#)]
68. Shen, Y.; Bax, A. SPARTA+: A Modest Improvement in Empirical NMR Chemical Shift Prediction by Means of an Artificial Neural Network. *J. Biomol. NMR* **2010**, *48*, 13–22. [[CrossRef](#)] [[PubMed](#)]
69. Schrödinger, L. *The PyMOL Molecular Graphics System, Version 1.8*; Schrodinger, LLC: New York, NY, USA, 2015.
70. 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](#)]
71. Shen, Y.; Bax, A. Protein Backbone and Sidechain Torsion Angles Predicted from NMR Chemical Shifts Using Artificial Neural Networks. *J. Biomol. NMR* **2013**, *56*, 227–241. [[CrossRef](#)]