Biophysical Reviews (2021) 13:1173–1178
https://doi.org/10.1007/s12551-021-00850-5

REVIEW

Secondary structure dependence on simulation techniques and force field parameters: from disordered to ordered proteins

Orkid Coskuner-Weber[1] · **Sule Irem Caglayan**[1]

Received: 15 August 2021 / Accepted: 1 October 2021 / Published online: 13 October 2021

Abstract

Computer simulations are used for identifying the secondary structure properties of ordered and disordered proteins. However, our recent studies showed that the chosen computer simulation protocol, simulation technique, and force field parameter set for a disordered protein impact its predicted secondary structure properties. Here, we compare the outcome from computer simulations utilizing molecular dynamics simulations without parallel tempering techniques using various force field parameter sets and temperature-replica exchange molecular dynamics simulations both for a model ordered and two model disordered proteins. Specifically, the model ordered protein is the third IgG-binding domain of Protein G (GB3) and the two model disordered proteins are amyloid-β(1–40) and α-synuclein in water. Our findings clearly indicate that temperature-replica exchange molecular dynamics simulations and molecular dynamics simulations without special sampling techniques yield similar results for the ordered GB3 protein whereas such agreement between simulation techniques using various force field parameter sets could not be obtained for disordered proteins. These findings clearly indicate that a consensus has to be reached via further development in computer simulation technique and force field parameter sets for disordered proteins.

Keywords Disordered proteins · Ordered proteins · Force field parameters · Simulation techniques · Simulation protocols

Introduction

Molecular dynamics (MD) simulations represent a computer simulation technique that is used in the investigations of ordered and disordered proteins for analyzing the physical movements of protein atoms. However, traditional MD simulations face challenges in sampling the whole conformational space of proteins within reachable simulation time-scales as it can be trapped in local minimum energy states (Coskuner and Wise-Scira 2013). By linking MD simulations to a Monte Carlo algorithm, the temperature-replica exchange molecular dynamics simulation (REMD) technique is capable of overcoming high energy barriers and enhances conformational space sampling of proteins (Qi et al. 2018). Furthermore, a force field parameter set is used for estimating the forces acting between atoms within a protein and also between proteins (Allison 2011). Specifically, a force field parameter set refers to the functional form utilized to compute the potential energy of a system of interest that consists of atoms (Carballo-Pacheco and Strodel 2017). These potential functions may be derived directly from experiments or from quantum mechanical calculations (Strodel and Coskuner-Weber 2019). Force field parameters define the energy landscape from which the acting forces on each particle are calculated as a gradient of the potential energy with respect to atom coordinates (Allison 2011). Commonly utilized force field parameter sets for proteins are the AMBER, CHARMM, OPLS-AA, and GROMOS parameters.

Our recent investigations showed that MD simulations using different force fields, as well as T-REMD simulations, are not capable of yielding results for the 140 amino acid long intrinsically disordered α-synuclein protein that present full agreement with available experiments (Mandaci et al. 2020). Interestingly, a recent additional study by our group revealed that best results from computer simulations that are in accord with experiments for the 40 amino acid long intrinsically disordered amyloid-β(1–40) peptide are obtained utilizing the charmm36m, a99sb-disp and a99sb*-ildn parameter sets in MD simulations as well as REMD

✉ Orkid Coskuner-Weber
weber@tau.edu.tr

[1] Molecular Biotechnology, Turkish-German University, Sahinkaya Caddesi, No. 106, Beykoz, Istanbul 34820, Turkey

simulations (Caliskan et al. 2021). However, a comparison of results obtained using different simulation techniques, simulation protocols, and force field parameter sets both for model ordered and disordered proteins' secondary structure properties is currently missing in the literature. Therefore, we conduct extensive REMD simulations on GB3 in water and obtained results are compared to those gained from MD simulations using different force fields. These findings are then compared to the results obtained for disordered proteins; α-synuclein and amyloid-β in an aqueous medium. For these purposes, REMD simulations on GB3 in water were conducted between 280 and 320 K temperatures using 32 replicas that were distributed exponentially between these temperatures. We used the charmm36m (c36m) parameters (Huang et al. 2017) for GB3 (PDBID: 1P7E) (Ulmer et al. 2003) and the TIP3P water model in REMD simulations (Jorgensen et al. 1983). After solvating GB3 in water by using a 20 Å water layer with 13,368 water molecules, the initial structure was equilibrated for 2 ns (per replica) via using first the canonical ensemble and then for 2 ns utilizing the isothermal-isobaric ensemble. Replica exchanges were attempted with a time step of 2 fs for every 5 ps. Configurations were saved every 500 steps. For maintaining the temperature of a replica with a collision frequency of 2 ps^{-1}, Langevin dynamics was used (Allison 2011). Particle mesh Ewald (PME) method was used to determine the long-range interactions (Allison 2011). All bonds to hydrogen atoms were constrained utilizing the SHAKE algorithm (Allison 2011). The MD simulation trajectories obtained from Shaw and co-workers using various force fields; the AMBER parameters (a99SB*-ILDN along with the TIP3P water model, the a03ws parameters, the a99SB parameters, and a99SB-ILDN parameters with the TIP4P-D model for water), a99SB-disp and the CHARMM force field parameter sets (c36m and c22*) were analyzed (Robustelli et al. 2018). These MD simulations were conducted for tens of microseconds without applying parallel tempering techniques. More details can be found in Ref. 11. For investigating the secondary structure properties, we utilized the DSSP program and an in-house script is utilized for determining the secondary structure abundances per amino acid residue (Mandaci et al. 2020). In all these calculations, the convergence is tested following our previous studies (Caliskan et al. 2021).

Secondary structure dependence on force field parameters and simulation techniques

The aim of this short review is to understand the secondary structure elements and their abundances for a model ordered and two model disordered proteins by REMD

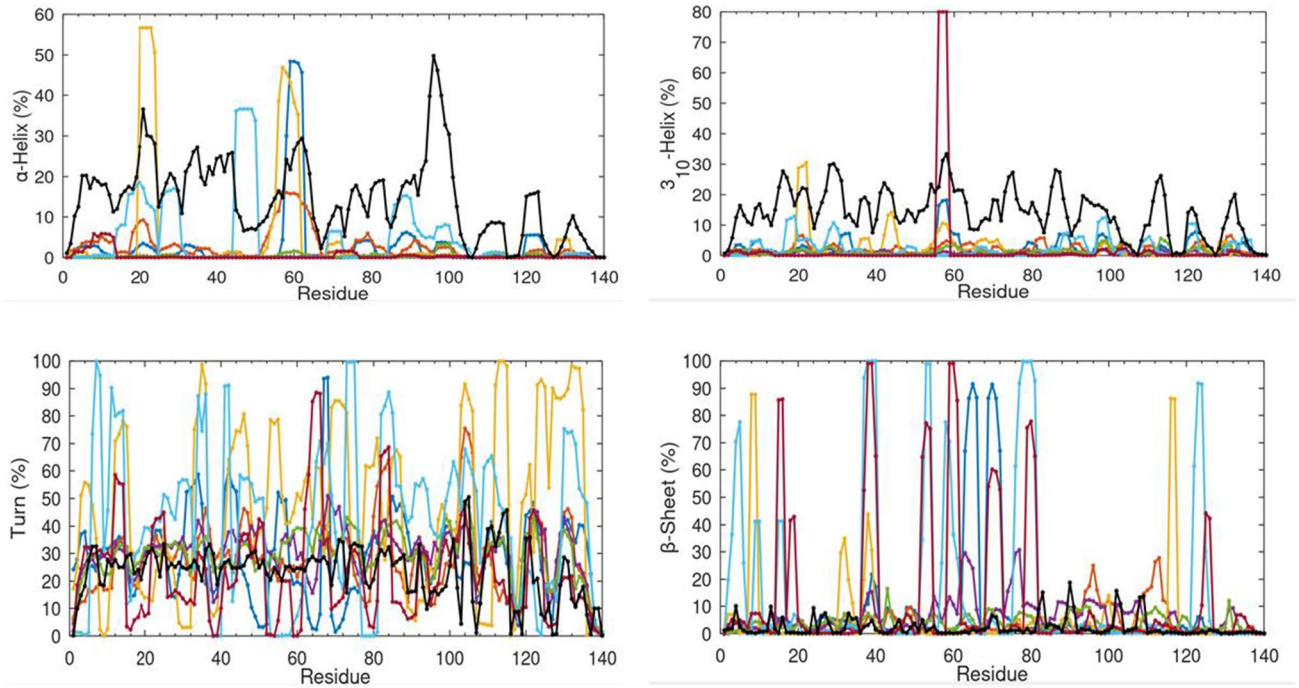

Fig. 1 A-helix, 3_{10}-helix, turn, and β-sheet propensities of disordered α-synuclein in an aqueous medium. A03ws (dark blue), a99sb-disp (orange), c36m (cherry), c22* (light blue), a99sb-UCB (green), and a99sb*-ildn (yellow) parameters with TIP3P water model, a99sb-ildn parameters with the TIP4P-D water model (violet) by MD simulations and AMBER ff99sb parameters utilizing the OBC water model by REMD simulations (black). Reproduced with permission from Ref. 6. Copyright, 2021, Chem. Biol. & Drug Des

and MD simulations using different force field parameter sets. Figure 1 presents the secondary structure abundances for α-helix, 3_{10}-helix, turn, and β-sheet structures for α-synuclein in an aqueous medium.

Based on these secondary structure component probabilities, MD simulations utilizing different force fields yield different secondary structure trends for α-synuclein in an aqueous solution medium. Significant variations exist between results obtained using different simulation protocols, simulation techniques, and force field parameters when we consider the large size disordered α-synuclein protein's epitope region identification. More details can be found in Ref. 6. Overall, experiments demonstrated more probable β-sheet formation between Pro120-Ala140 in the C-terminal region of α-synuclein (Marsh et al. 2006). This characteristic could not be captured in MD simulations using different force field parameter sets or REMD simulations. Also, based on experiments, larger abundant α-helical structure formation occurs in the N-terminal, NAC, and C-terminal regions of α-synuclein (specifically, between Lys60–Leu100) (Marsh et al. 2006). REMD simulations catch this trend for α-helix structure formation but this trend might also indicate that our REMD simulations are stabilizing α-structure formation in α-synuclein. Experiments reported abundant α-helical structure formation in the N-terminal region whereas NAC region demonstrates α-helical structure formation with a smaller probability in comparison to the N-terminal region. Additionally, C-terminal region possesses α-helical structure formation with a smaller probability than the NAC region of α-synuclein (Marsh et al. 2006). Such a trend could not be obtained from MD simulations (except with the usage of a99sb*-ildn parameters) or from our parallel tempering REMD simulations. Bussel and Elizier proposed that the NAC region`s ability for forming fibrils requires residues Val71-Val82, indicating that this region might be forming abundant β-sheet structure and might be therefore involved in early inter-molecular interactions (Bussell and Eliezer 2001). Interestingly, MD simulations utilizing the c22* and c36m parameter sets yield such a trend. A comparison of REMD simulations to experiments yields that the α-helix abundance is larger in the N-terminal and NAC regions rather than in the C-terminal region of α-synuclein. This finding is in accord with NMR experiments (Marsh et al. 2006). In fact, α-helix formation in the NAC and N-terminal regions of α-synuclein was reported to be a key feature for membrane and vesicle binding.

Figure 2 presents the secondary structure component probabilities for amyloid-β(1–40) in an aqueous medium.

Obtained results from MD simulations—utilizing different force field parameters—present varying epitope regions for smaller size disordered amyloid-β(1–40) in an aqueous medium regarding secondary structure propensity

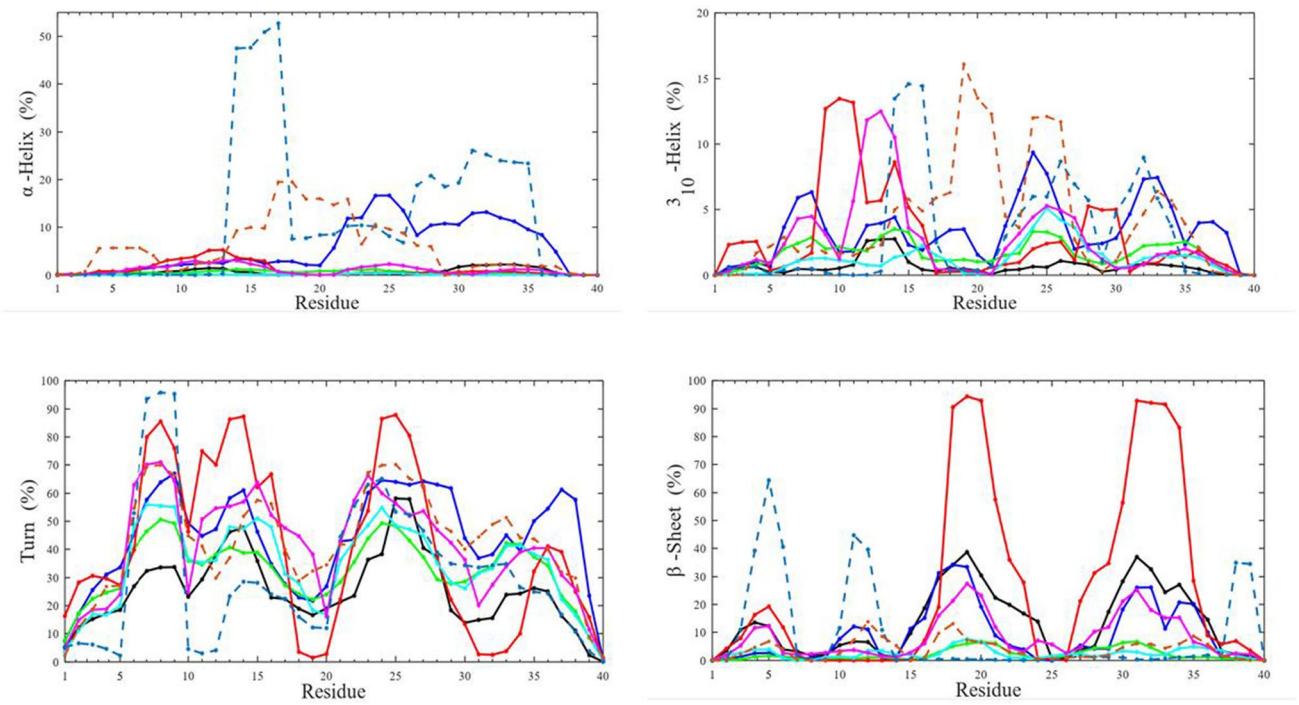

Fig. 2 A-helix, 3_{10}-helix, turn, and β-sheet propensities of disordered amyloid-β(1–40) in an aqueous medium using a03ws (dark blue), a99sb-disp (pink), c36m (black), a99sb-UCB (green), c22* (purple), and a99sb*-ildn (red) parameters along with the TIP3P water model, a99sb-ildn parameters with the TIP4P-D water model (light blue) by MD simulations and AMBER ff99sb parameters using the OBC water model by REMD simulations (brown). Reproduced with permission from Ref. 7. Copyright, 2021, Chem. Biol. & Drug. Des

calculations (Fig. 2) (Caliskan et al. 2021). Significant differences exist in determined epitope regions utilizing different simulation protocols and simulation techniques as can be seen in Fig. 2 and more details including comparison to experiments about these findings can be found in Ref. 7. All in all, simulation results show better agreement with experiments for the smaller disordered amyloid-β(1–40) than for the larger size disordered α-synuclein protein in an aqueous solution environment. Specifically, NMR measurements demonstrated slightly prominent α-helical structure formation at Phe4–Ser8 in the N-terminal region of amyloid-β(1–40) (Hou et al. 2004). MD and REMD simulations (utilizing different force fields) show less than 5% of α-helical structure formation in the N-terminus region of amyloid-β(1–40) (Fig. 2). For the Lys17–Lys28 region, experiments showed a larger prominent α-helical structure formation in comparison to the C-terminal region (Ile31-Ile40) of amyloid-β(1–40) (Tomaselli et al. 2006). REMD simulations reproduce these experimental findings for amyloid-β(1–40). Also, MD simulations using the c22* parameter set yield results in agreement with experiments. Furthermore, NMR measurements illustrated that the Ala21-Ala30 region of amyloid-β(1–40) does not demonstrate α-helical structure formation (Fawzi et al. 2008). This finding could be reproduced by MD simulations using the c36m, a99sb, a99sb-disp, and a99sb*-ildn parameter sets. Moreover, experiments demonstrated 3_{10}-helical structure formation at Asp1-Phe4 (Sgourakis et al. 2007). This characteristic is reproduced by REMD as well as MD simulations using different force fields (Fig. 2). In addition, NMR measurements showed 3_{10}-helical structure formation in the N-terminal region of amyloid-β(1–40) (Turner et al. 2003). These findings could be replicated by REMD and MD simulations (using different force fields except with the usage of the a03ws parameters). Also, NMR studies reported 3_{10}-helical structure formation at His13–Asp23; (Vivekanandan et al. 2011) these findings could also be obtained from REMD simulations. In addition, circular dichroism experiments reported 3_{10}-helical structure formation at Val12–Lys28 and this characteristic is also obtained from our REMD simulations (Jarvet et al. 2003). Asp7–Glu11 and Tyr10–Val12 regions adopt turn structure based on NMR measurements (Rosenman et al. 2013) and this trend could be replicated by all simulations of interest in this study. Furthermore, turn structure formation was reported for the Val24–Lys28 region (Rosenman et al. 2013) and this trend is obtained from MD and REMD simulations (Fig. 2). Turn structure formation was also reported for the Ala21–Ala30 region of the disordered amyloid-β(1–40) peptide (Williams et al. 2004), which could be replicated by MD and REMD simulations. In addition, NMR experiments demonstrated β-sheet formation at Leu17–Phe20 and no β-sheet was detected in the C-terminal region of the disordered amyloid-β(1–40) peptide (Danielsson et al. 2005). These trends could be obtained from all simulations except MD simulations utilizing the a03ws parameter set. Moreover, CD experiments reported abundant β-sheet formation at Gly25–Met35 (Danielsson et al. 2005) and MD as well as REMD simulations yield the same findings except MD simulations using the a03ws parameter set (Fig. 2). NMR measurements demonstrated an additional β-sheet structure formation at Ala21–Ala30 (Fawzi et al. 2008) and our results demonstrate a stark increase in β-sheet structure formation abundance at Ala21–Ser26 and a smaller probability for this secondary structure element in the rest of the mid-domain region.

Figure 3 demonstrates our secondary structure calculation results for the model ordered GB3 protein in an aqueous solution environment by MD simulations—utilizing different force fields—and REMD simulations.

MD and REMD simulation results—using various force fields—yield for the same region (Ala23–Asn37) prominent α-helix structure formation in an aqueous solution environment for the ordered GB3 protein. However, the probability of α-helix formation is smaller and differs by up to 15% from the results obtained from REMD simulations especially for the residues located between Glu24 and Phe30. Abundant 3_{10}-helix formation is detected at Asp47–Thr49 and the highest probability is obtained from MD simulations using the a99sb-UCB parameter set (47.3%). We notice that the c22* parameter set yields the lowest abundance for 3_{10}-helix formation in this region of GB3 in an aqueous solution environment. Interestingly, additional 3_{10}-helix formation with smaller probability (about 5%) is detected at Ala26–Lys28 using the a99sb-UCB parameter set in MD simulations (Fig. 3). Turn structure is formed in the N-terminal, mid-domain, and C-terminal regions of GB3 (Fig. 3). Specifically, all simulation results yield that Gly9–Lys13 in the N-terminal region adopts turn structure with probabilities varying between 3 and 100%. Additionally, prominent turn structure formation is noted at Ala20–Asp22. The C-terminal residues Asp47–Lys50 adopt prominent turn structure (up to 95%). A more enhanced turn structure is noted in the mid-domain region of GB3 utilizing the a99sb-UCB parameters in MD simulations. B-sheet structure formation is detected with highest probabilities in the C- and N-terminal regions of GB3 (Fig. 3). Specifically, MD and REMD simulations—utilizing different force field parameter sets—yield that Gln2–Gly9, Leu12–Lys19, Val42–Asp46, and Thr51–Thr55 adopt prominent β-sheet with probabilities up to 100%. These results are in excellent agreement with X-ray and NMR measurements (Ulmer et al. 2003).

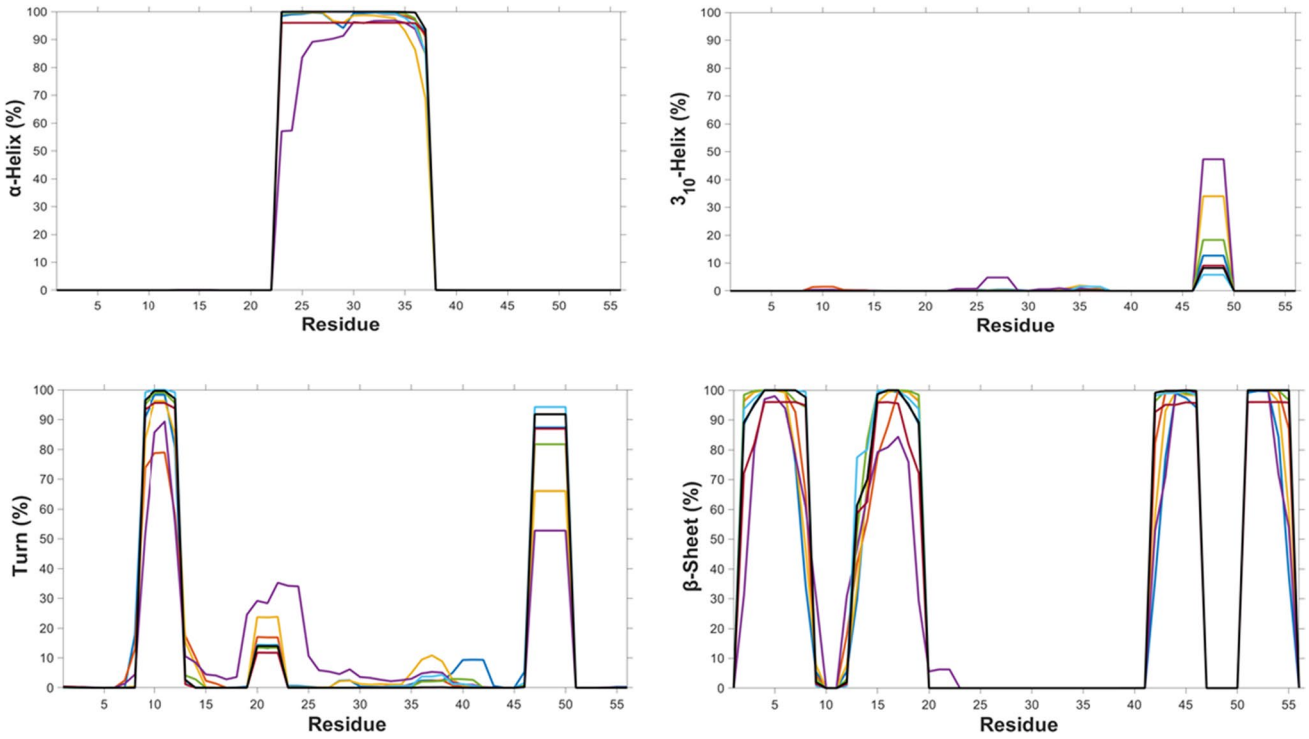

Fig. 3 A-helix, 3_{10}-helix, turn, and β-sheet propensities for the ordered GB3 protein in aqueous solution using a03ws (dark blue), c36m (dark red), a99sb-disp (orange), c22* (light blue), a99sb-UCB (violet), and a99sb*-ildn (green) parameters along with TIP3P water model, a99sb-ildn parameters with the TIP4P-D water model (dark yellow) by MD simulations and AMBER ff99sb parameters using the TIP3P model for water by REMD simulations (black)

Conclusion

Here, we present a short review about the secondary structure properties for model disordered proteins (α-synuclein and amyloid-β) and for a model ordered protein (GB3) by MD simulations using different force field parameter sets as well as REMD simulations. MD simulations using different force fields and REMD simulations present varying secondary structure abundances and regions for α-synuclein and amyloid-β(1–40) in an aqueous solution environment. Less agreement with available experiments is detected for the larger size disordered α-synuclein in comparison to the smaller size disordered amyloid-β(1–40) peptide in an aqueous solution environment. There are significant differences between obtained secondary structure properties of disordered proteins based on the chosen simulation protocol, simulation technique, and force field parameter set. For the smaller size disordered amyloid-β(1–40) peptide, we report that all force field parameter sets except a03ws parameters present findings in partial or full agreement with available experiments. REMD simulations for amyloid-β(1–40) show agreement with existing experiments as well. These validation studies showed that the MD simulations using the a99sb*-ildn, c36m, and a99sb-disp parameters for the disordered smaller size amyloid-β(1–40) peptide or REMD simulations yield the best results in excellent agreement with various experiments. Regarding the ordered GB3 protein, we find that almost all MD simulations using different force field parameter sets and REMD simulations yield the same regions for secondary structure element location but the probabilities depend on the chosen force field parameter set. Results obtained from MD and REMD simulations are in excellent agreement with X-ray and NMR measurements for the ordered GB3 protein in water. We should also mention that the literature captures many more force field parameters that are based on modifications of existing generally used force field parameter sets, such as ff19IDPs and ff14IDPSFF for intrinsically disordered proteins but these are not validated—based on secondary structure properties—by us in this short review.

All in all, our short review demonstrates that MD simulations and REMD simulations yield excellent results for ordered proteins in water in comparison to disordered proteins. Partial or more agreement with experiments is detected for smaller size disordered proteins but less agreement with experimental data is obtained for large size disordered proteins in an aqueous medium. These results indicate that the primary structure and the size of the disordered protein impact the accuracy of computer simulations. A consensus and accuracy that is obtained for the ordered protein

is also required for small and large size disordered proteins regarding the chosen MD simulation technique, simulation protocol, and force field parameter set. Furthermore, the chosen water model may impact the simulation outcome and accuracy as well. Current literature includes also recent development of various water models for computational simulation studies of intrinsically disordered proteins. However, without having a consensus reached for protein force field parameters, it is difficult to validate the usefulness of such water models.

Acknowledgements We thank D. E. Shaw and co-workers for sharing their MD simulation trajectories for this review.

References

Allison TC (2011) Metallic Systems: A Quantum Chemist's Perspective, 0 ed. Allison TC, Coskuner O, Gonzalez CA eds. CRC Press. https://doi.org/10.1201/b10835

Bussell R, Eliezer D (2001) Residual structure and dynamics in Parkinson's disease-associated mutants of alpha-synuclein. J Biol Chem 276(49):45996–46003. https://doi.org/10.1074/jbc.M106777200

Caliskan M, Mandaci SY, Uversky VN, Coskuner-Weber O (2021) Secondary structure dependence of amyloid-β(1–40) on simulation techniques and force field parameters. Chem Biol Drug Des 97(5):1100–1108. https://doi.org/10.1111/cbdd.13830

Carballo-Pacheco M, Strodel B (2017) Comparison of force fields for Alzheimer's Aβ42: a case study for intrinsically disordered proteins: comparison of force fields for Alzheimer's A B42. Protein Sci 26(2):174–185. https://doi.org/10.1002/pro.3064

Coskuner O, Wise-Scira O (2013) Arginine and disordered amyloid-β peptide structures: molecular level insights into the toxicity in Alzheimer's disease. ACS Chem Neurosci 4(12):1549–1558. https://doi.org/10.1021/cn4001389

Danielsson J, Jarvet J, Damberg P, Gräslund A (2005) The Alzheimer beta-peptide shows temperature-dependent transitions between left-handed 3-helix, beta-strand and random coil secondary structures. FEBS J 272(15):3938–3949. https://doi.org/10.1111/j.1742-4658.2005.04812.x

Fawzi LN, Phillips AH, Ruscio JZ, Doucleff M, Wemmer DE, Head-Gordon T (2008) Structure and dynamics of the Aβ $_{21–30}$ peptide from the interplay of NMR experiments and molecular simulations. J Am Chem Soc 130(19):6145–6158. https://doi.org/10.1021/ja710366c

Hou L, Shao H, Zhang Y, Li H, Menon NK, Neuhaus EB, Brewer JM, Byeon I-JL, Ray DG, Vitek MP, Iwashita T, Makula RA, Przybyla AB, Zagorski MG (2004) Solution NMR studies of the Aβ(1–40) and Aβ(1–42) peptides establish that the Met35 oxidation state affects the mechanism of amyloid formation. J Am Chem Soc 126(7):1992–2005. https://doi.org/10.1021/ja036813f

Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmüller H, MacKerell AD (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods 14(1):71–73. https://doi.org/10.1038/nmeth.4067

Jarvet J, Damberg P, Danielsson J, Johansson I, Eriksson LEG, Gräslund A (2003) A left-handed 3(1) helical conformation in the Alzheimer abeta(12–28) peptide. FEBS Lett 555(2):371–374. https://doi.org/10.1016/s0014-5793(03)01293-6

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79(2):926–935. https://doi.org/10.1063/1.445869

Mandaci SY, Caliskan M, Sariaslan MF, Uversky VN, Coskuner-Weber O (2020) Epitope region identification challenges of intrinsically disordered proteins in neurodegenerative diseases: secondary structure dependence of A-synuclein on simulation techniques and force field parameters. Chem Biol Drug Des cbdd.13662. https://doi.org/10.1111/cbdd.13662

Marsh JA, Singh VK, Jia Z, Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: implications for fibrillation. Protein Sci 15(12):2795–2804. https://doi.org/10.1110/ps.062465306

Qi R, Wei G, Ma B, Nussinov R (2018) Replica exchange molecular dynamics: a practical application protocol with solutions to common problems and a peptide aggregation and self-assembly example. Methods Mol Biol Clifton NJ 1777:101–119. https://doi.org/10.1007/978-1-4939-7811-3_5

Robustelli P, Piana S, Shaw DE (2018) Developing a molecular dynamics force field for both folded and disordered protein states. Proc Natl Acad Sci 115(21):E4758–E4766. https://doi.org/10.1073/pnas.1800690115

Rosenman DJ, Connors CR, Chen W, Wang C, García AE (2013) Aβ monomers transiently sample oligomer and fibril-like configurations: ensemble characterization using a combined MD/NMR approach. J Mol Biol 425(18):3338–3359. https://doi.org/10.1016/j.jmb.2013.06.021

Sgourakis NG, Yan Y, McCallum SA, Wang C, Garcia AE (2007) The Alzheimer's peptides Aβ40 and 42 adopt distinct conformations in water: a combined MD / NMR study. J Mol Biol 368(5):1448–1457. https://doi.org/10.1016/j.jmb.2007.02.093

Strodel B, Coskuner-Weber O (2019) Transition metal ion interactions with disordered amyloid-β peptides in the pathogenesis of Alzheimer's disease: insights from computational chemistry studies. J Chem Inf Model 59(5):1782–1805. https://doi.org/10.1021/acs.jcim.8b00983

Tomaselli S, Esposito V, Vangone P, van Nuland NAJ, Bonvin AMJJ, Guerrini R, Tancredi T, Temussi PA, Picone D (2006) The α-to-β conformational transition of Alzheimer's Aβ-(1–42) peptide in aqueous media is reversible: a step by step conformational analysis suggests the location of β conformation seeding. ChemBioChem 7(2):257–267. https://doi.org/10.1002/cbic.200500223

Turner PR, O'Connor K, Tate WP, Abraham WC (2003) Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. Prog Neurobiol 70(1):1–32. https://doi.org/10.1016/s0301-0082(03)00089-3

Ulmer TS, Ramirez BE, Delaglio F, Bax A (2003) Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. J Am Chem Soc 125(30):9179–9191. https://doi.org/10.1021/ja0350684

Vivekanandan S, Brender JR, Lee SY, Ramamoorthy A (2011) A partially folded structure of amyloid-beta(1–40) in an aqueous environment. Biochem Biophys Res Commun 411(2):312–316. https://doi.org/10.1016/j.bbrc.2011.06.133

Williams AD, Portelius E, Kheterpal I, Guo J, Cook KD, Xu Y, Wetzel R (2004) Mapping abeta amyloid fibril secondary structure using scanning proline mutagenesis. J Mol Biol 335(3):833–842. https://doi.org/10.1016/j.jmb.2003.11.008