Received: 30 June 2023 | Revised: 1 September 2023 | Accepted: 7 September 2023

DOI: 10.1002/prot.26590

RESEARCH ARTICLE

PROTEINS WILEY

Atomistic molecular simulations of Aβ-Zn conformational ensembles

Julen Aduriz-Arrizabalaga | Xabier Lopez | David De Sancho

Polimero eta Material Aurreratuak: Fisika, Kimika eta Teknologia, Kimika Fakultatea, UPV/EHU & Donostia International Physics Center (DIPC), Donostia-San Sebastian, Euskadi, Spain

Correspondence
David De Sancho, Polimero eta Material Aurreratuak: Fisika, Kimika eta Teknologia, Kimika Fakultatea, PV/EHU & Donostia International Physics Center (DIPC), PK 1072, 20018, Donostia-San Sebastian, Euskadi, Spain.
Email: david.desancho@ehu.eus

Funding information
Eusko Jaurlaritza; Ministerio de Ciencia e Innovación; the Office of Science Research, Grant/Award Number: PID2021-127907NB-I00; Ramón y Cajal contract, Grant/Award Number: RYC-2016-19590

Abstract

The amyloid-forming Aβ peptide is able to interact with metal cations to form very stable complexes that influence fibril formation and contribute to the onset of Alzheimer's disease. Multiple structures of peptides derived from Aβ in complex with different metals have been resolved experimentally to provide an atomic-level description of the metal-protein interactions. However, Aβ is intrinsically disordered, and hence more amenable to an ensemble description. Molecular dynamics simulations can now reach the timescales needed to generate ensembles for these type of complexes. However, this requires accurate force fields both for the protein and the protein-metal interactions. Here we use state-of-the-art methods to generate force field parameters for the Zn(II) cations in a set of complexes formed by different Aβ variants and combine them with the Amber99SB*-ILDN optimized force field. Upon comparison of NMR experiments with the simulation results, further optimized with a Bayesian/Maximum entropy approach, we provide an accurate description of the molecular ensembles for most Aβ-metal complexes. We find that the resulting conformational ensembles are more heterogeneous than the NMR models deposited in the Protein Data Bank.

KEYWORDS
amyloid, force field, intrinsically disordered proteins, molecular dynamics

1 | INTRODUCTION

Alzheimer's disease is the leading cause of senile dementia and has over 55 million cases reported by the World Health Organization as of 2021.[1] Although the exact cause of Alzheimer's disease is unknown, the aggregation and deposition of amyloid beta (Aβ) in neural tissue is widely accepted as a contributing factor to the onset of the disease. Transition metal ions and oxidative metabolism have been proposed to play fundamental roles in the process of aggregation and deposition of Aβ.[2] In particular, the binding of divalent metals, such as copper, iron, and zinc influences the aggregation process of the protein, contributing directly to the severity of the disease.[3–6] It has been reported that both the monomeric and oligomeric forms of Aβ are neurotoxic[7] and that these cations directly influence toxicity.[8] Interestingly, the Zn(II) ion concentration in the brain[9] of around 150 μM is an order of magnitude higher than the ion concentration in blood. Furthermore, even though the Zn(II) levels remain relatively constant through adult life, a significantly elevated concentration has been found in the brains of patients affected by Alzheimer's disease.[10] Therefore, the role Zn(II) plays in Alzheimer's disease has become of great interest.[3–5]

Despite the relevance of protein-metal interactions to a thorough understanding of Aβ-Zn(II) complexes, their details have remained elusive. This is partly due to Aβ being an intrinsically disordered protein

(IDP), which makes an ensemble description preferable to the static structures derived from experimental methods such as X-ray crystallography or Nuclear Magnetic Resonance (NMR).[11] Additionally, Zn(II) characterization is highly limited by its physico-chemical properties. Out of its three stable isotopes, only ^{67}Zn is NMR active, but due to its low natural abundance and low receptivity, only solid-state, low-temperature NMR studies of zinc compounds are practically achievable. In addition, zinc complexes have no absorbance in the UV–Vis and microwave spectral regions, and its completely filled d^{10} orbitals render it diamagnetic, being, therefore, invisible in EPR spectroscopy.[12] Although Aβ's metal binding region has been located between residues 1–16,[13,14] even in this short fragment there are multiple possible chelating amino acids (Asp^1, Glu^3, His^6, Asp^7, Glu^{11}, His^{13} and His^{14}).[15] All these factors combined make experimental studies of Aβ-Zn(II) systems highly challenging. Computer simulations are ideally suited to complement experiments in our understanding of these systems.

Modeling complexes of Aβ with transition metals like Zn(II) is however not without its own challenges. First, the chemical flexibility of Zn(II) allows it to adopt different coordination modes with Aβ (see Figure 1). In the experimentally resolved structures of Aβ with Zn(II) complexes, we find the metal tetrahedrally coordinated by glutamic/aspartic acid and/or histidine residues,[16] making different combinations possible. Hence, modelers have multiple possibilities to simulate the protein-transition metal interactions.[17] The simplest option is through nonbonded models, where the metal ion interacts via electrostatic and Van der Waals interactions. Nonbonded models have successfully been used in the past.[18–21] However, in the case of Zn(II) nonbonded models often result in the wrong coordination, and the metal may even leave the coordination site.[17] This problem may be particularly acute for complexes with IDPs, where the metal may be more solvent-exposed than in globular proteins. In order to overcome these limitations, one possibility is to use "dummy models," which contain virtual atoms with point charges that promote the target coordination.[22,23] Unfortunately, they have been reported to undergo dissociation in long-time scale simulations.[17] Alternatively, one can use constraints to force the coordination of metal binding.[16,24,25] Nonetheless, it has been reported that improbable fluctuations appeared using this approach.[17] Alternatively, one can use bonded models, where chemical bonds are defined between the protein and the metal, which keep the coordination stable throughout the whole simulation time. Bonded models have been used extensively for zinc-containing systems,[2,26–28] including Aβ-metal complexes.[6,29–32]

The modeling of Aβ-Zn(II) complexes is further complicated by the disordered nature of Aβ.[11] IDPs lack a well-defined native structure and are thus best described as an ensemble of conformations. Simulating IDPs hence requires classical force fields carefully parametrized to avoid too strong structural propensities favoring folded states.[33,34] In the last few years, extensive work has dramatically improved the ability of force fields to reproduce experimental data on peptides, unfolded states of proteins, and IDPs.[35–41] The remaining inaccuracies in the simulated ensembles can be alleviated using integrative approaches, where experimental data is used as input in the modeling.[42] One possibility is to restrain simulations to match the experimental data available. In the case of heterogeneous systems, this approach may bias the sampling toward conformations that are not representative of any of the relevant states.[43] Alternatively, one can use maximum-entropy approaches to generate conformational ensembles that are also compatible with the prior knowledge of the system given by the force field.[43,44] Specifically, Bayesian/Maximum entropy reweighting methods optimize statistical weights of the snapshots sampled in molecular simulations in order to maximize the agreement with experimental data. These approaches have recently been used in various systems, such as RNA tetranucleotides[45] or IDPs.[46,47]

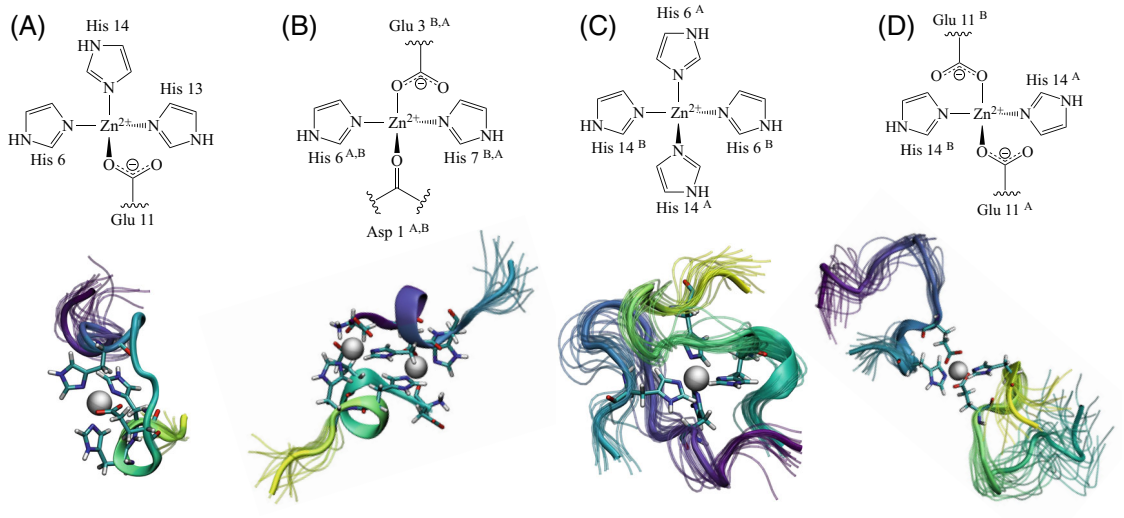

FIGURE 1 Perspective drawing of the metal centers (top) and cartoon representation of Aβ-Zn(II) NMR ensembles of experimentally characterized systems (bottom). PDB ID: 1ZE9 (A), 5LFY (B), 2LI9 (C), and 2MGT (D).

TABLE 1 Molecular systems included in this study. In the sequences, zinc chelating residues are highlighted in red.

PDB ID	BMRB ID	MRBLOCK ID	Description	Sequence	Oligomerization	#Zn(II) cations
1ZE9[15]	–	16 076	Human Aβ WT	Ac-DAEFRHDSGYEVHHQK-NH2	Monomer	1
5LFY[54]	34 019	620 016	Human Aβ D7H	DAEFRHHSGY-NH2	Dimer	2
2LI9[53]	19 602	530 791	Rat Aβ	Ac-DAEFGHDSGFEVRHQK-NH2	Dimer	1
2MGT[57]	17 884	582 988	Human Aβ H6R	Ac-DAEFRRDSGYEVHHQK-NH2	Dimer	1

In this work, we focus on the N-terminal region of Aβ, which contains its zinc-binding region.[13] Specifically, we study four different complexes formed by slightly different peptides from Aβ with varying sequences, coordination and stoichiometry, which provide a challenging test for the simulation models. Using atomistic molecular dynamics (MD) simulations and an optimized force field, we generate conformational ensembles of the Aβ-Zn(II) complexes with different coordinations and in different oligomeric states that we have parametrized using the metal center parameter builder package (MCPB.py) developed by Li and Merz.[48] This approach has recently been used to study complexes of Aβ with Zn(II),[49] Cu(II) and Al(III).[32] Here we combine the resulting parameters with the optimized Amber99SB*-ILDN force field,[50] which includes corrections for backbone[36] and sidechain torsions[37] to the ff99SB force field[51] that better capture experimental structural propensities in short peptides. In a recent study comparing multiple force fields to study $A\beta_{42}$, this modified force field produced results that compared extremely well against experiments.[52] We have used these parameter sets to run molecular simulations and validated them against experimental data including NMR chemical shifts and NOEs. Although we find a good agreement between simulation and experiment, we further improve the ensembles using a recently developed Bayesian/Maximum entropy reweighting method.[43] Our results indicate that Aβ-Zn(II) conformational ensembles may be more heterogeneous than suggested by the NMR models deposited in the Protein Data Bank.

2 | MATERIALS AND METHODS

2.1 | Aβ-Zn(II) models

In order to find minimal model systems for the study of protein-metal interactions, we queried the Protein Data Bank for entries from the full-length Aβ peptide in complex with Zn(II). This query returned four structural models of peptides of different length and sequence from the $A\beta_{1-16}$ N-terminal fragment, widely regarded as the metal-binding domain.[13] These models include variants of human and rat Aβ whose structures have been resolved using solution NMR in the presence of Zn^{2+} (see Table 1). The peptide-metal complexes manifest the structural polymorphism of Aβ.[4] While the human WT sequence appears in the monomeric form (1ZE9[15]), the D7H and H6R mutants (5LFY[53] and 2MGT,[54] respectively) and rat Aβ[53] are all dimers, with either one or two cations per complex. In the structures, chelation takes place from combinations of most of the amino acid residues that have been proposed to interact with the Zn(II) cation,[3,15] specifically Asp[1], Glu[3], His[6], Glu[11], His[13], and His[14] (see Figure 2). In all four cases, NOE restraints are available at the NMR restraints grid,[55] and for the dimers chemical shifts are also available at the Biological Magnetic Resonance Database (BMRB)[56] (see Table 1).

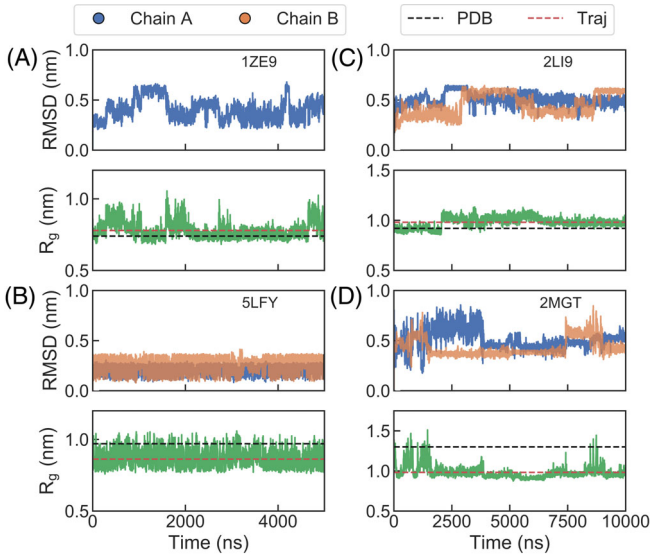

FIGURE 2 *RMSD* and R_g values calculated for 1ZE9 (A), 5LFY (B), 2LI9 (C) and 2MGT (D) from the simulations using the parametrized bonded models. Mean R_g values calculated from PDB models are shown by black lines and mean R_g values calculated from our simulations are shown by red lines.

2.2 | Parametrization of protein-metal interactions

All bonded models used have been parametrized using the MCPB.py tool by Merz and co-workers,[48] which is distributed with Ambertools.* We performed all DFT quantum mechanical calculations using the B3LYP/6-31G*[58,59] method in the Gaussian16 package.[60] For each of the model systems, generating new parameters requires three different steps: geometry optimization, estimation of force constants using the Seminario method[61] and calculation of point charges using the restrained electrostatic potential (RESP) model.[62]

Three out of four of the systems we are studying are protein dimers, where two exact copies of the same protein chain are

*https://ambermd.org/AmberTools.php

coordinated to either one or two Zn(II) cations. Due to minor discrepancies between dimers present in the experimental structures, we have obtained slightly different values of equilibrium geometry parameters, force constants, and charges for the atoms in each of the chains. To make the two chains in the dimers indistinguishable in our simulations, the protein-metal force field parameters obtained should be exactly the same. We have followed the approach taken by Peters et al. when parametrizing the ZAFF force field[27] and averaged the charges, equilibrium geometry parameters, and force constants of bonds, angles, and dihedrals of all duplicate values. Hence, we have obtained the same parameter sets for both protein chains present in dimer systems.

All the files used to parametrize the bonded models and the resulting parameters are available at https://osf.io/y4zk5/. All parameters are within the same range of values and all models follow the same qualitative tendencies, and atomic charges and spring constants are consistent with those of existing zinc force fields such as ZAFF[27] and EZAFF.[28]

2.3 | Molecular dynamics simulations

We have inserted all systems within cubic boxes and solvated them with TIP3P water molecules.[63] We added a 0.1 M concentration of Cl^- and Na^+ ions to all boxes and also neutralized the total charge of the system when necessary. The systems were energy-minimized using the steepest descent algorithm and then equilibrated in two stages. First, we run a 100 ps simulation in the NVT ensemble and then another 100 ps in the NPT ensemble, both including position restraints in the protein-heavy atoms. Molecular Dynamics (MD) simulations were performed in the NVT ensemble, using an integration time step of 2 fs. Constant temperature and pressure were set by coupling the system to a Parrinello-Rahman barostat at 1.0 bar and a velocity-rescaling thermostat at 278 K,[64–66] corresponding to the temperature of the NMR experiments. Equilibrium simulations were run for 5–10 μs, depending on the size of the system and convergence of the simulations, with the Amber99SB*-ILDN force field[36,37] using the Gromacs software package (version 2020).[67]

2.4 | Analysis, validation, and reweighting

To analyze the results of the simulations we have used a combination of Gromacs tools and in-house scripts using the MDtraj Python library.[68] We have validated the results of our simulations against experiments back-calculating a series of NMR observables from the simulations. We used the SPARTA+ program to back-calculate chemical shifts[69] and the TALOS+ server to derive secondary structures from the experimental data.[70] Additionally, we estimate distances corresponding to the observed NOEs using r^{-6} averaging.[71–73]

We additionally perform a Bayesian/Maximum entropy reweighting using the BME program[43] to refine simulations using the experimental data available, in this case NOE distance restraints. The main goal of this reweighting is to obtain an optimized set of statistical weights for the frames in the simulation ensemble that are in better agreement with experiments given a forward model for the experimental observable.[42] This results in a reduction in the χ^2 value, which quantifies the agreement against experiments and is expressed as[43]

$$\chi^2 = \left\langle \frac{\left(F_i^{sim} - F_i^{exp}\right)^2}{\sigma_i^2} \right\rangle. \quad (1)$$

where $F_i^{sim/exp}$ is the value of an observable i obtained from simulation/experiment, σ_i is the uncertainty in the observable and the brackets denote an average over observables. More information about BME is provided in Data S1. We note that the BME approach is different (and can be combined with) other methods that may be used to bias the data generation, like metadynamics or accelerated MD.

3 | RESULTS AND DISCUSSION

3.1 | Molecular simulations with new force fields parameters

In a first attempt to model the complexes of Aβ-Zn systems, we used non-bonded models using the recently optimized Amber99SB-disp force field.[40] Unfortunately, these efforts resulted in the loss of the tetrahedral coordination of the metal within a few hundreds of nanoseconds (see Figures S1–S5 for additional details). This result adds to the mounting evidence that non-bonded models cannot be used to appropriately model Aβ-metal complexes.[17] We hence resorted to generate parameters for the metal using a bonded model. Specifically, we used the MCBP.py package to generate parameters for all the Aβ-Zn(II) complexes in Table 1 (see Section 2). The resulting atomic charges and spring constants are consistent with those of existing zinc force fields.[27,28]

Using the new parameters, we have run 5 μs atomistic MD simulations for all four systems. This amount of simulation time seems sufficient to converge the distribution of R_g for the monomer (1ZE9) and D7H mutant dimer (5LFY), which remains stable for over half the simulation runs (see Figure S6). In the case of the rat and H6R mutant, we extended the simulation runs up to 10 μs in order to sample conformational space more exhaustively. In Figure 2, we show the values of the $C^\alpha - RMSD$ and the radius of gyration (R_g) for each of them. The *RMSD* values typically remain under 5 Å for most of the simulation time. However, with the exception of the D7H mutant dimer (5LFY), we observe relatively large fluctuations that suggest that the simulations explore a variety of conformations that are dissimilar from those reported in the PDB. On the other hand, we find that the R_g values remain stable throughout the simulations, although two of the systems are notably more compact than the PDB models. This effect is likely due to the influence of the TIP3P water model used in our simulations, which is known to produce overcompaction.[74]

In order to identify the regions of the protein with the largest fluctuations, we also report values for the *RMSF* for all systems (see Figure 3A). As expected, the highest *RMSF* values correspond to

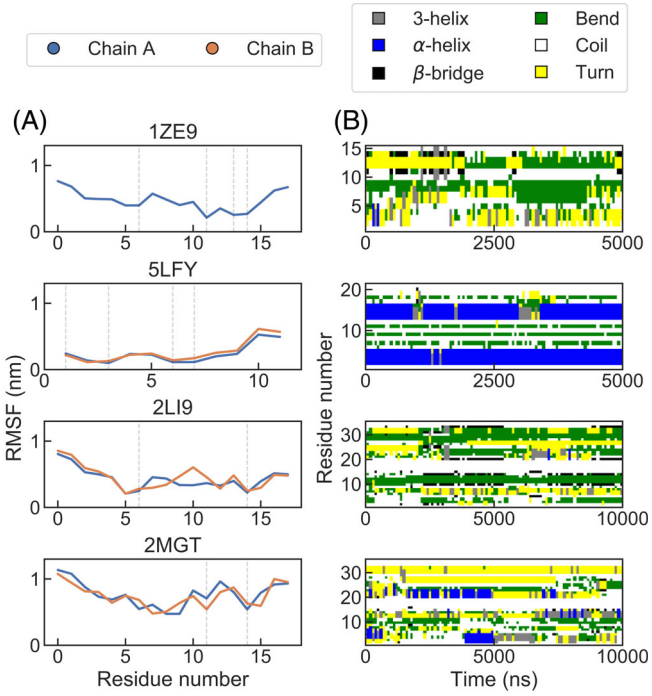

FIGURE 3 (A) Per-residue *RMSF* values calculated for each of the Aβ-Zn systems. For dimers, we represent each chain with a different color. Vertical dashed lines mark the Zn-chelating residues. (B) Secondary structure assignment calculated with DSSP.

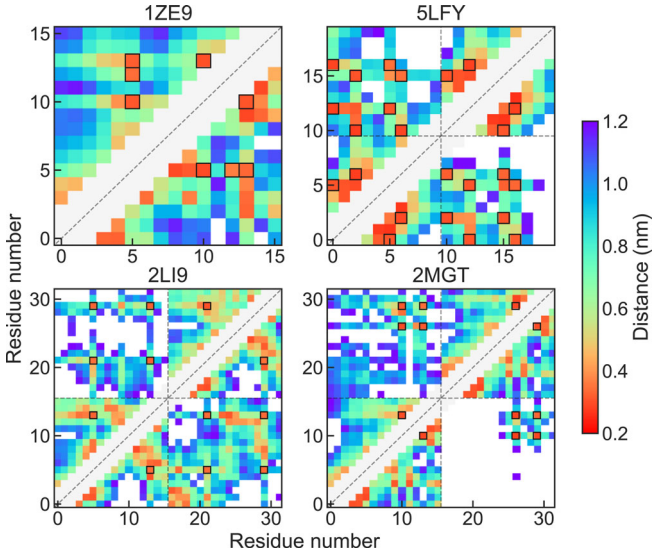

FIGURE 4 Contact maps calculated from PDB structures (lower triangle) and MD trajectories (upper triangle). Residues belonging to the binding site are highlighted in black squares.

regions at the N and C-termini of each system, except in the case of the D7H mutant dimer (5LFY). In this case, the low *RMSF* values in the N-termini are related to the coordination of the Zn^{2+} cation by the oxygen from the carbonyl of the first amino acid in the peptide sequence. On the contrary, in the case of the monomer (1ZE9), H6R mutant dimer (2MGT) and rat Aβ dimer (2LI9), fluctuations are on average larger and their values are greatest at the N and C-termini.

The trends observed in the *RMSF* are consistent with the very different structural propensities in the metal-bound peptides. In Figure 3B, we show the assignments to different types of secondary structure obtained using the DSSP algorithm.[75] The regions that more stably keep the secondary structure coincide with those with the lowest *RMSF* values, which also contain the chelated residues. This behavior is clearest for the D7H mutant dimer (5LFY), where the α-helices remain fully formed in both peptide chains through almost the complete duration of the simulation trajectory.

In order to visualize the differences between the ensemble of conformations sampled during the simulations and the NMR structures, we have calculated the contact maps, which we compare with the averages from the 20 structural models deposited in the PDB for each of our systems (see Figure 4). We find that despite the intermediate values of the *RMSD* reported above, the contact maps from the simulations are largely consistent with those of the experimental models. Specifically, the binding regions of all models are recapitulated accurately. This is expected because the conformational dynamics in our complexes are restricted by the interactions with the metal cations. More notable differences are observed for the intermolecular interactions in the case of dimers. In one case, the H6R mutant dimer, we find that there are substantially more intermolecular contacts in the simulated ensemble than in the NMR models.

3.2 | Validation against NMR data

In order to assess the validity of our simulations more carefully, we compare the results obtained using the new metal parameters and the Amber99SB*-ILDN force field directly against available experimental NMR data. First, we compare experimental chemical shifts with values back-calculated from the simulations using SPARTA+[69] (see Section 2). In Table 2 we report the *RMSD* and χ^2 values of the calculated chemical shifts, and compare them with those for the NMR models. Overall, both the *RMSD* and χ^2 values obtained from the 20 structures deposited in the PDB and our 5–10 μs simulations are in the same range, indicating that our simulations have reached a comparable level of accuracy as the experimental models for the Aβ-Zn (II) systems.

There are, however, some noteworthy inconsistencies between the calculated chemical shifts and the experimental measurements (see Figure 5). For example, the simulations do not completely track the overall tendency of the chemical shift of Hα for residues 1 and 3 of the D7H mutant (5LFY model). In the case of C^{α} and C^{β} chemical shifts, the agreement is worst for the rat peptide (2LI9), although the errors are also largest for the models deposited in the PDB and the overall trends are captured correctly. Instead, for both the D7H and H6R mutants (5LFY and 2MGT, respectively) the agreement with experiments is excellent for the carbon chemical shifts. Lastly, for the HN protons, the simulations for the rat peptide (2LI9) result in the best agreement obtained shown both from *RMSD* and χ^2 values. Chemical shifts are well-known to report on secondary structures of

TABLE 2 Calculated *RMSD* and χ^2 values of chemical shifts back-calculated from PDB structures and simulations.

	5LFY		2LI9		2MGT	
***RMSD* (ppm)**	**PDB**	**Simulation**	**PDB**	**Simulation**	**PDB**	**Simulation**
H^{α}	0.21	0.35	0.24	0.34	0.27	0.25
C^{α}	0.88	0.65	1.44	1.76	1.23	1.08
C^{β}	1.02	1.03	1.53	2.22	1.05	1.14
H^{N}	0.53	0.48	0.55	0.43	0.46	0.49
χ^2	PDB	Simulation	PDB	Simulation	PDB	Simulation
H^{β}	0.90	2.43	1.45	2.04	1.39	1.12
C^{α}	0.95	0.86	3.01	3.67	2.04	1.52
C^{β}	0.88	1.05	2.00	4.26	1.03	1.13
H^{N}	1.32	1.28	1.62	0.83	1.01	1.12

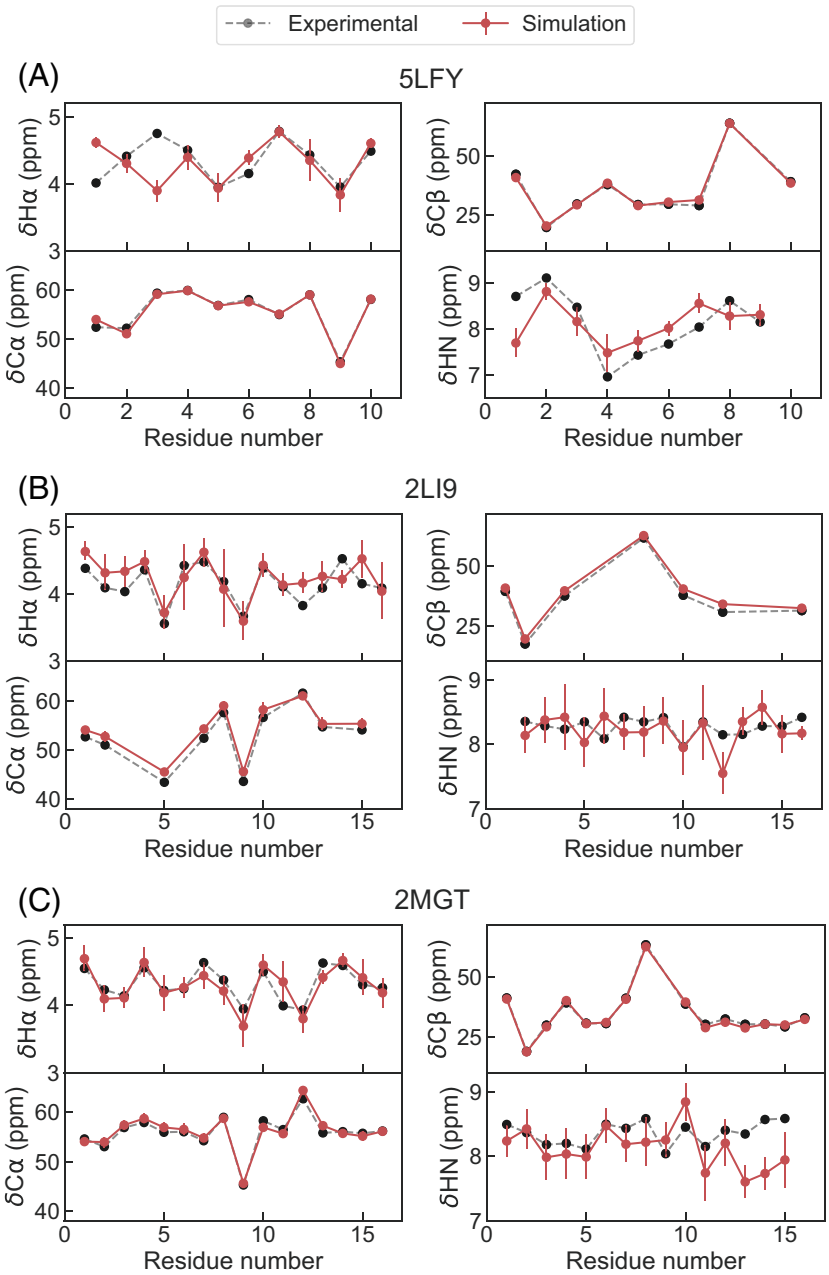

FIGURE 5 Chemical shifts calculated for the 5LFY (A), 2LI9 (B) and 2MGT (C) systems with experimental data available. Experimental values are shown in black and back calculated chemical shifts are shown in red.

proteins. For the three dimers, we used the software TALOS+ to assign secondary structures from the chemical shift data.[70] In Figure S7, we show that the resulting assignments are consistent with the helical propensities calculated from the simulations. This is particularly important for the D7H mutant and rat Aβ, where the *RMSD* and χ^2 of chemical shifts are the largest.

Additionally, we have calculated average distances from the simulations for all the atom pairs involved in measurable NOEs (see Figure 6). We note that for all the dimers, the intra-chain NOE distance restraint values are duplicated for chains A and B in the BMRB entries (see Figure S8). To estimate the average distances from the simulations, we assume that the chains are indistinguishable, as they are in the experiment, and keep only one set of NOE values. Accordingly, we report a single set of distances averaging distances from both chains. When no errors were reported, we arbitrarily assign a 20% error to the measured value.

Overall, the simulations with the new bonded parameters and the Amber99SB*-ILDN force field successfully capture the vast majority of the experimental NOE distances. Interestingly the χ^2 value obtained for each system is 0.078 for the monomer (1ZE9), 0.076 for the D7H mutant (5LFY), 0.081 for the rat dimer (2LI9), and 0.083 for the H6R mutant (2MGT). Nonetheless, some values of NOE distances are overestimated from the simulations, for example, distances 65 and 66 of the monomer (1ZE9), distances 0–10 in the rat dimer (2LI9) or distances from 15 to 20 of the H6R mutant (2MGT) (see Figure 6).

3.3 | BME reweighting improves agreement with NMR data

The overestimation of some NOE distances from the unbiased simulations suggests there is room for improvement in the conformational ensembles. Therefore, we have reweighted our simulations to further improve the quality of the ensembles obtained for the systems using the BME program developed by Bottaro et al.[43] which relies on Bayesian/Maximum entropy approach (see Section 2) using the NOE data from Figure 6. Ensemble reweighting helps improve the agreement with experiments for all systems, significantly in some cases, as indicated by a decrease in χ^2. After optimization the χ^2 values are of 0.050 for the monomer (1ZE9), 0.045 for the D7H mutant (5LFY), 0.046 for the rat dimer (2LI9), and 0.032 for the H6R (2MGT), a 36%, 55%, 43%, and 61% improvement, respectively.

In order to illustrate the effect of the reweighting, we show the distribution of R_g calculated for all systems (see Figure 7). In the case of the monomer system (1ZE9), a bimodal distribution of the R_g is obtained. Before reweighting, the distribution shows that contorted conformations are much more populated while reweighting reduces that bias. In Figure 7B, we show selected snapshots corresponding to different ensembles of the monomer system (1ZE9). Similarly, in the case of the rat dimer (2LI9), the force field prior distribution indicates a greater propensity for extended conformations, which is modulated by the reweighting. As for the rat dimer, we show snapshots

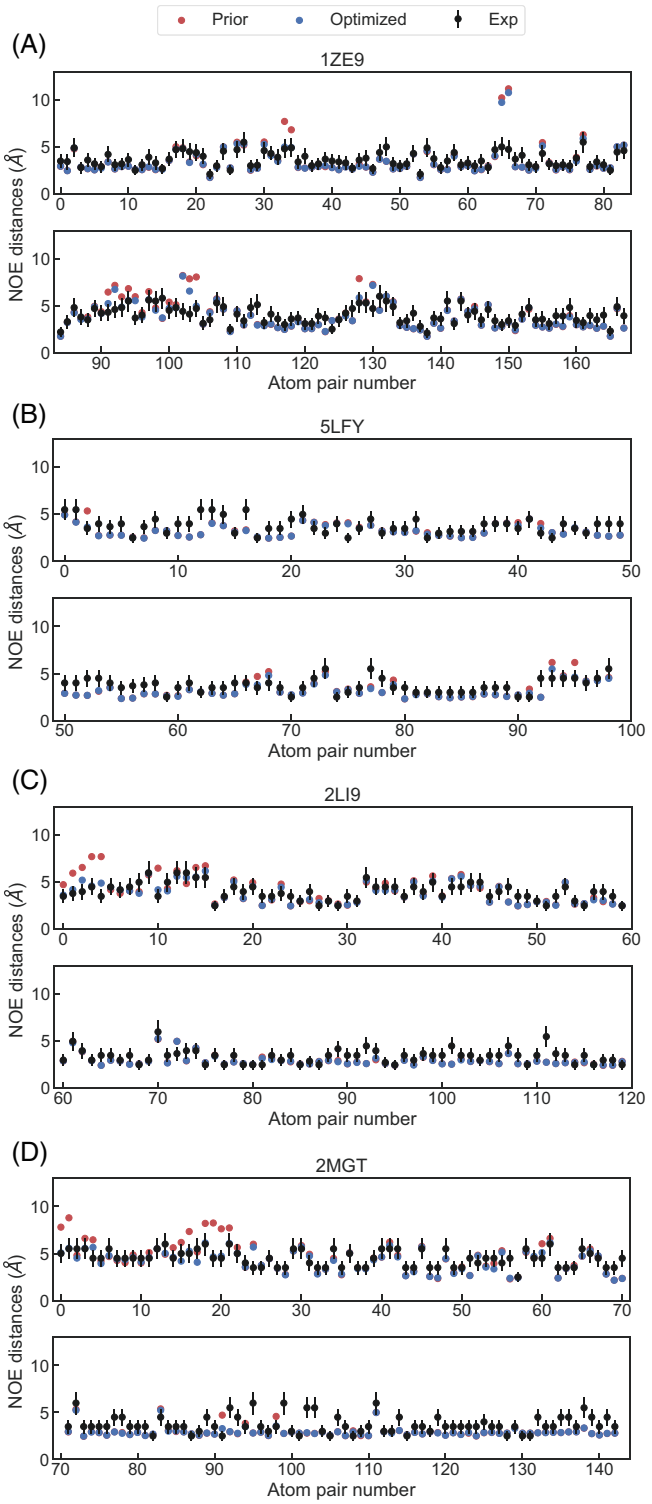

FIGURE 6 NOEs calculated for the 1ZE9 (A), 5LFY (B), 2LI9 (C), and 2MGT (D) systems with experimental data available. Experimental values are shown in black, back-calculated NOEs are shown in red and NOEs calculated after optimization are shown in blue.

corresponding to both populations in Figure 7C. The effect of the reweighting in the case of the H6R mutant (2MGT) is more modest but again shows that ensembles including extended conformations

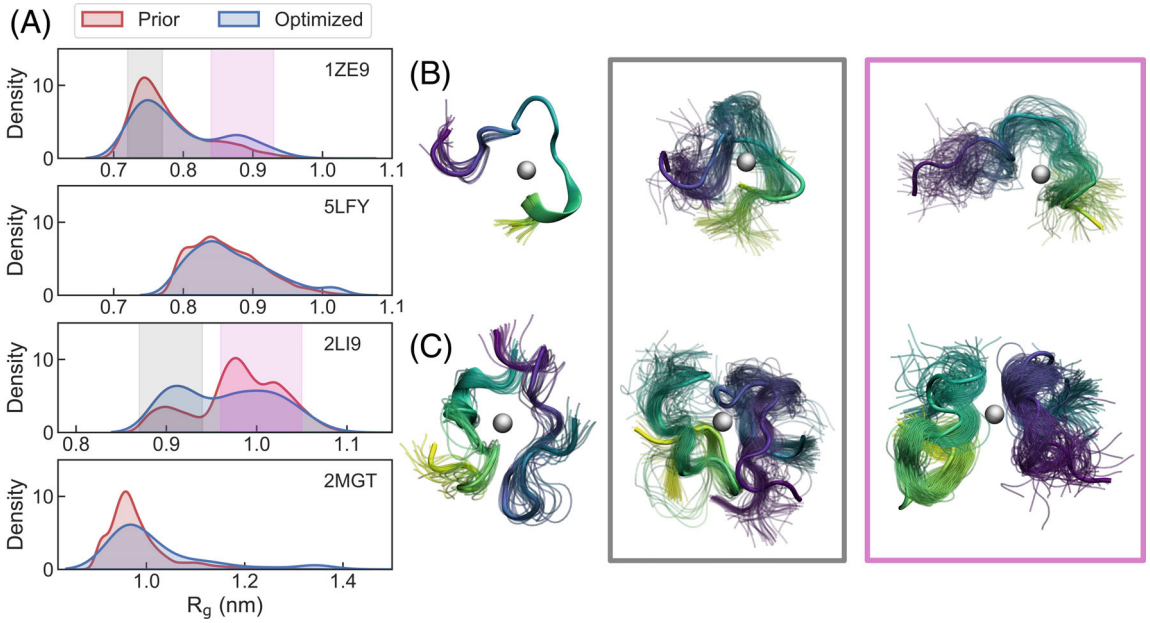

FIGURE 7 (A) Distribution of R_g calculated from the prior distribution and optimized distribution. Compact and extended regions are highlighted in black and magenta, respectively. (B) Snapshots from the NMR ensemble (left), and compact (center) and extended (right) subpopulations of the simulated ensembles for 1ZE9. (C) Same for 2LI9. Ensembles from compact and extended subpopulations are marked by black and magenta rectangles, respectively.

are in better agreement with experiments. In this case, we find that the reweighted ensemble has a more sparse contact map than the raw MD ensemble (see Figure S18), resulting in a better agreement with the models in the NMR ensemble. Furthermore, the improvement in the distances where the agreement with NOEs was worst is highlighted in Figure S19. Clearly, the reweighting greatly improves agreement with experiments (see Figure S19B), a trend common to all the systems under consideration. To quantify the heterogeneity of the NMR and simulated ensembles, we have calculated the within-ensemble RMSDs for the sets illustrated in Figure 7.[76] For the NMR ensemble of the monomer (1ZE9), the RMSD value is 0.25 nm, while the two subpopulations of the simulated ensemble have RMSDs of 0.53 and 0.54 nm (for the compact and extended ensembles, respectively). Similarly, in the case of the rat dimer (2LI9) the within-ensemble RMSD is 0.29 nm for the NMR models, while it is 0.41 and 0.53 nm for the compact and extended ensembles, respectively. These results clearly show the effect the BME reweighting has over simulations and prove that, generally, the most accurate ensemble description of Aβ-Zn(II) systems are built with more heterogeneous ensembles than would be expected from the NMR models alone.

4 | CONCLUSIONS

In this work, we have produced bonded models of four Aβ-Zn(II) systems for the Amber99SB*-ILDN force field. Using these parameters, we have run 5–10 μs simulations of four different systems and validated the results against experiments. The simulations have sampled well beyond the conformational space of the models reported in the PDB as shown by the RMSD of C^{α} and contact maps. This extended sampling nevertheless compares well with experimentally measured chemical shifts and NOEs. A better agreement between simulation and experiment could be derived from a detailed characterization of force field effects in the conformational ensembles. Specifically, recent work using different explicit and implicit solvent models shows strong differences in the dimensions of simulated ensembles,[77] which for some of our systems seem overly compact relative to the NMR models.

In order to improve the agreement between simulations and experiments, we have reweighted our simulations with the BME program[43] using experimental NOE data. Interestingly, the improvement of the simulations shows that the best agreement is obtained when both collapsed and extended conformations of the systems are present in the conformational ensembles. This may have implications for our understanding of the role of Zn(II) in the aggregation of Aβ, which has been explored in the past using structured NMR models as initial structures for MD simulations.[16] The more heterogeneous ensembles that we have obtained may contribute to the polymorphism that has been proposed for Aβ-Zn(II) complexes in the past.[4] Heterogeneous conformational states for the different coordinations considered in our models may serve to seed simulations in studies of primary and secondary nucleation of fibrils,[78] and compare whether these processes relevant to aggregation are facilitated by the presence of the cation in one or several of the intervening molecules. Additionally, the ensembles that we have generated may be useful to understand the differences between aggregation in the absence and presence of metals.[5] Specifically, the heterogeneous Aβ-Zn(II) ensembles may facilitate the predominantly non-fibrillar aggregation observed in experiments, which is later followed by maturation of aggregates to fibrils. The timescales for this type of processes may however be

difficult to study using atomistic molecular simulations and will require coarse-grained models that in some way incorporate the effects of the metal.[79]

AUTHOR CONTRIBUTIONS

Julen Aduriz-Arrizabalaga: Writing – original draft; investigation. **Xabier Lopez:** Conceptualization; funding acquisition; writing – original draft; supervision. **David De Sancho:** Conceptualization; funding acquisition; writing – original draft; supervision.

ACKNOWLEDGMENTS

Financial support comes from Eusko Jaurlaritza (Basque Government) through the project IT1584-22 and from the Spanish Ministry of Science and Universities through the Office of Science Research (MINECO/FEDER) through grant PID2021-127907NB-I00. DDS acknowledges the Spanish Ministry of Science and Universities for a Ramón y Cajal contract (Grant RYC-2016-19590). The authors thank Sandro Bottaro for help using the BME program and Kresten Lindorff Larsen for useful discussions. The authors also acknowledge the staff at the DIPC Supercomputing Center for technical support.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/prot.26590.

DATA AVAILABILITY STATEMENT

All the input files required to reproduce our results, together with structure and parameter files have been deposited in the Open Science Framework repository https://osf.io/y4zk5/ with DOI:10.17605/OSF.IO/Y4ZK5.

ORCID

David De Sancho https://orcid.org/0000-0002-8985-2685

REFERENCES

1. World Health Organization. Dementia. https://www.who.int/news-room/fact-sheets/detail/dementia, Accessed: 2023-02-14
2. Barnham KJ, Bush AI. Metals in Alzheimer's and Parkinson's diseases. *Curr Opin Chem Biol*. 2008;12:222-228.
3. Faller P, Hureau C. Bioinorganic chemistry of copper and zinc ions coordinated to amyloid-β peptide. *Dalton Trans*. 2009;1080-1094.
4. Miller Y, Ma B, Nussinov R. Polymorphism in Alzheimer Aβ amyloid organization reflects conformational selection in a rugged energy landscape. *Chem Rev*. 2010;110:4820-4838.
5. Tõugu V, Tiiman A, Palumaa P. Interactions of Zn (II) and Cu (II) ions with Alzheimer's amyloid-beta peptide. Metal ion binding, contribution to fibrillization and toxicity. *Metallomics*. 2011;3:250-261.
6. Wise-Scira O, Xu L, Perry G, Coskuner O. Structures and free energy landscapes of aqueous zinc (II)-bound amyloid-β (1–40) and zinc (II)-bound amyloid-β (1–42) with dynamics. *J Biol Inorg Chem*. 2012;17:927-938.
7. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*. 2002;297:353-356.
8. Chen T, Wang X, He Y, et al. Effects of Cyclen and cyclam on zinc (II)- and copper (II)-induced amyloid β-peptide aggregation and neurotoxicity. *Inorg Chem*. 2009;48:5801-5809.
9. Sensi SL, Paoletti P, Bush AI, Sekler I. Zinc in the physiology and pathology of the CNS. *Nat Rev Neurosci*. 2009;10:780-791.
10. Deibel M, Ehmann W, Markesbery W. Copper, iron, and zinc imbalances in severely degenerated brain regions in Alzheimer's disease: possible relation to oxidative stress. *J Neurol Sci*. 1996;143:137-142.
11. Das P, Matysiak S, Mittal J. Looking at the disordered proteins through the computational microscope. *ACS Cent Sci*. 2018;4:534-542.
12. Laitaoja M, Valjakka J, Janis J. Zinc coordination spheres in protein structures. *Inorg Chem*. 2013;52:10983-10991.
13. Mekmouche Y, Coppel Y, Hochgräfe K, et al. Characterization of the ZnII binding to the peptide amyloid-β_{1-16} linked to Alzheimer's disease. *Chembiochem*. 2005;6:1663-1671.
14. Tsvetkov PO, Kulikova AA, Golovin AV, et al. Minimal Zn2+ binding site of amyloid-β. *Biophys J*. 2010;99:L84-L86.
15. Zirah S, Kozin SA, Mazur AK, et al. Structural changes of region 1-16 of the Alzheimer disease amyloid β-peptide upon zinc binding and in vitro aging. *J Biol Chem*. 2006;281:2151-2161.
16. Miller Y, Ma B, Nussinov R. Zinc ions promote Alzheimer Aβ aggregation via population shift of polymorphic states. *Proc Natl Acad Sci U S A*. 2010;107:9490-9495.
17. Strodel B, Coskuner-Weber O. 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*. 2019;59:1782-1805.
18. Hoops SC, Anderson KW, Merz KM Jr. Force field design for metalloproteins. *J Am Chem Soc*. 1991;113:8262-8270.
19. Stote RH, Karplus M. Zinc binding in proteins and solution: a simple but accurate nonbonded representation. *Proteins*. 1995;23:12-31.
20. Bredenberg J, Nilsson L. Modeling zinc sulfhydryl bonds in zinc fingers. *Int J Quantum Chem*. 2001;83:230-244.
21. Yang B, Zhu Y, Wang Y, Chen G. Interaction identification of Zif268 and TATAZF proteins with GC−/AT-rich DNA sequence: a theoretical study. *J Comput Chem*. 2011;32:416-428.
22. Duarte F, Bauer P, Barrozo A, et al. Force field independent metal parameters using a nonbonded dummy model. *J Phys Chem B*. 2014;118:4351-4362.
23. Liao Q, Pabis A, Strodel B, Kamerlin SCL. Extending the nonbonded cationic dummy model to account for ion-induced dipole interactions. *J Phys Chem Lett*. 2017;8:5408-5414.
24. Man VH, Nguyen PH, Derreumaux P. High-resolution structures of the amyloid-β_{1-42} dimers from the comparison of four atomistic force fields. *J Phys Chem B*. 2017;121:5977-5987.
25. Li W, Zhang J, Wang J, Wang W. Metal-coupled folding of Cys2His2 zinc-finger. *J Am Chem Soc*. 2008;130:892-900.
26. Lin F, Wang R. Systematic derivation of AMBER force field parameters applicable to zinc-containing systems. *J Chem Theory Comput*. 2010;6:1852-1870.
27. Peters MB, Yang Y, Wang B, Fusti-Molnar L, Weaver MN, Merz KM Jr. Structural survey of zinc-containing proteins and development of the zinc AMBER force field (ZAFF). *J Chem Theory Comput*. 2010;6:2935-2947.
28. Yu Z, Li P, Merz KM Jr. Extended zinc AMBER force field (EZAFF). *J Chem Theory Comput*. 2018;14:242-254.
29. Pan L, Patterson JC. Molecular Dynamics Study of Zn(Aβ) and Zn(Aβ) 2. *PloS One*. 2013;8:e70681.
30. Xu L, Wang X, Wang X. Effects of Zn2+ binding on the structural and dynamic properties of amyloid B peptide associated with Alzheimer's disease: Asp1 or Glu11? *ACS Chem Nerosci*. 2013;4:1458-1468.
31. Boopathi S, Dinh Quoc Huy P, Gonzalez W, Theodorakis PE, Li MS. Zinc binding promotes greater hydrophobicity in Alzheimer's $A\beta_{42}$

peptide than copper binding: molecular dynamics and solvation thermodynamics studies. *Proteins*. 2020;88:1285-1302.
32. Roldán-Martín L, Peccati F, Sciortino G, Sodupe M, Maréchal J-D. Impact of Cu (ii) and Al (iii) on the conformational landscape of amyloid$\beta_{1\text{-}42}$. *Phys Chem Chem Phys*. 2021;23:13023-13032.
33. Huang J, MacKerell AD Jr. Force field development and simulations of intrinsically disordered proteins. *Curr Opin Struct Biol*. 2018;48:40-48.
34. Shea J-E, Best RB, Mittal J. Physics-based computational and theoretical approaches to intrinsically disordered proteins. *Curr Opin Struct Biol*. 2021;67:219-225.
35. Best RB, Buchete N-V, Hummer G. Are current molecular dynamics force fields too helical? *Biophys J*. 2008;95:L07-L09.
36. Best RB, Hummer G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J Phys Chem B*. 2009;113:9004-9015.
37. Lindorff-Larsen K, Piana S, Palmo K, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*. 2010;78:1950-1958.
38. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. Systematic validation of protein force fields against experimental data. *PloS One*. 2012;7:e32131.
39. Huang J, Rauscher S, Nawrocki G, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods*. 2017;14:71-73.
40. Robustelli P, Piana S, Shaw DE. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A*. 2018;115:E4758-E4766.
41. Zerze GH, Zheng W, Best RB, Mittal J. Evolution of all-atom protein force fields to improve local and global properties. *J Phys Chem Lett*. 2019;10:2227-2234.
42. Bottaro S, Lindorff-Larsen K. Biophysical experiments and biomolecular simulations: a perfect match? *Science*. 2018;361:355-360.
43. Bottaro S, Bengtsen T, Lindorff-Larsen K. Integrating molecular simulation and experimental data: a Bayesian/maximum entropy reweighting approach. In: Gáspári Z, (ed) Structural Bioinformatics. *Methods in Molecular Biology*. Humana, New York, NY. 2020;2112:219-240. https://doi.org/10.1007/978-1-0716-0270-6_15
44. Hummer G, Köfinger J. Bayesian ensemble refinement by replica simulations and reweighting. *J Chem Phys*. 2015;143:12B634_1.
45. Bottaro S, Bussi G, Kennedy SD, Turner DH, Lindorff-Larsen K. Conformational ensembles of RNA oligonucleotides from integrating NMR and molecular simulations. *Sci Adv*. 2018;4:eaar8521.
46. Crehuet R, Buigues PJ, Salvatella X, Lindorff-Larsen K. Bayesian-maximum-entropy reweighting of IDP ensembles based on NMR chemical shifts. *Entropy*. 2019;21:898.
47. Mantsyzov AB, Maltsev AS, Ying J, Shen Y, Hummer G, Bax A. A maximum entropy approach to the study of residue-specific backbone angle distributions in α-synuclein, an intrinsically disordered protein. *Protein Sci*. 2014;23:1275-1290.
48. Li P, Merz KM Jr. MCPB.py: a python based metal center parameter builder. *J Chem Inf Model*. 2016;56:599-604.
49. Al-Shammari N, Savva L, Kennedy-Britten O, Platts JA. Forcefield evaluation and accelerated molecular dynamics simulation of Zn (II) binding to N-terminus of amyloid-β. *Comput Biol Chem*. 2021;93:107540.
50. Piana S, Lindorff-Larsen K, Shaw DE. How robust are protein folding simulations with respect to force field parameterization? *Biophys J*. 2011;100:L47-L49.
51. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins*. 2006;65:712-725.
52. Carballo-Pacheco M, Strodel B. Comparison of force fields for Alzheimer's a: a case study for intrinsically disordered proteins. *Protein Sci*. 2017;26:174-185.
53. Istrate AN, Tsvetkov PO, Mantsyzov AB, et al. NMR solution structure of rat Aβ (1–16): toward understanding the mechanism of Rats' resistance to Alzheimer's disease. *Biophys J*. 2012;102:136-143.
54. Polshakov VI, Mantsyzov AB, Kozin SA, et al. A binuclear zinc interaction fold discovered in the homodimer of Alzheimer's amyloid-β fragment with Taiwanese mutation D7H. *Angew Chem Int ed*. 2017;129:11896-11901.
55. Doreleijers JF, Mading S, Maziuk D, et al. BioMagResBank database with sets of experimental NMR constraints corresponding to the structures of over 1400 biomolecules deposited in the Protein Data Bank. *J Biomol NMR*. 2003;26:139-146.
56. Ulrich EL, Akutsu H, Doreleijers JF, et al. BioMagResBank. *Nucl Acids Res*. 2007;36:D402-D408.
57. Istrate AN, Kozin SA, Zhokhov SS, et al. Interplay of histidine residues of the Alzheimer's disease Aβ peptide governs its Zn-induced oligomerization. *Sci Rep*. 2016;6:1-14.
58. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J Phys Chem A*. 1994;98:11623-11627.
59. Hehre WJ, Ditchfield R, Pople JA. Self—consistent molecular orbital methods. XII. Further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules. *Chem Phys*. 1972;56:2257-2261.
60. Frisch MJ, Trucks GW, Schlegel HB, et al. *GaussianËœ16 Revision C.01*. Gaussian Inc.; 2016.
61. Seminario JM. Calculation of intramolecular force fields from second-derivative tensors. *Int J Quantum Chem*. 1996;60:1271-1277.
62. Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem*. 1993;97:10269-10280.
63. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *Chem Phys*. 1983;79:926-935.
64. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys*. 2007;126:014101.
65. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys*. 1981;52:7182-7190.
66. Nosé S, Klein M. Constant pressure molecular dynamics for molecular systems. *Mol Phys*. 1983;50:1055-1076.
67. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1-2:19-25.
68. McGibbon RT, Beauchamp KA, Harrigan MP, et al. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys J*. 2015;109:1528-1532.
69. 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.
70. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. 2009;44:213-223.
71. Tropp J. Dipolar relaxation and nuclear overhauser effects in nonrigid molecules: the effect of fluctuating internuclear distances. *J Chem Phys*. 1980;72:6035-6043.
72. Rieping W, Habeck M, Nilges M. Inferential structure determination. *Science*. 2005;309:303-306.
73. Daura X, Antes I, van Gunsteren WF, Thiel W, Mark AE. The effect of motional averaging on the calculation of NMR-derived structural properties. *Proteins*. 1999;36:542-555.
74. Shabane PS, Izadi S, Onufriev AV. General purpose water model can improve atomistic simulations of intrinsically disordered proteins. *J Chem Theory Comput*. 2019;15:2620-2634.

75. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983;22:2577-2637.
76. Kuzmanic A, Zagrovic B. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. *Biophys J*. 2010;98:861-871.
77. Rauscher S, Gapsys V, Gajda MJ, Zweckstetter M, De Groot BL, Grubmuller H. Structural ensembles of intrinsically disordered proteins depend strongly on force field: a comparison to experiment. *J Chem Theory Comput*. 2015;11:5513-5524.
78. Bellaiche MM, Best RB. Molecular determinants of $A\beta_{42}$ adsorption to amyloid fibril surfaces. *J Phys Chem Lett*. 2018;9:6437-6443.
79. Wu C, Shea J-E. Coarse-grained models for protein aggregation. *Curr Opin Struc Biol*. 2011;21:209-220.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Aduriz-Arrizabalaga J, Lopez X, De Sancho D. Atomistic molecular simulations of Aβ-Zn conformational ensembles. *Proteins*. 2024;92(1):134-144. doi:10.1002/prot.26590