ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

A Structure Independent Molecular Fragment Interfuse Model for Mesoscale Dissipative Particle Dynamics Simulation of Peptides

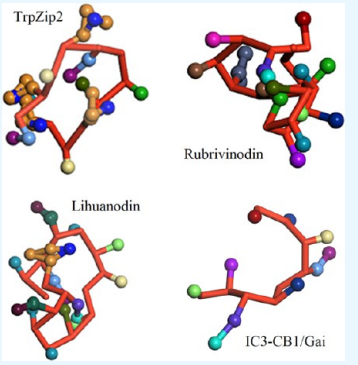
Ricky Anshuman Dash and Esmaiel Jabbari*

Cite This: *ACS Omega* 2024, 9, 18001−18022

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: There is a need to develop robust computational models for mesoscale simulation of the structure of peptides over large length scales toward the discovery of novel peptides for medical applications to address the issues of peptide aggregation, enzymatic degradation, and short half-life. The primary objective was to predict the structure and conformation of peptides whose native structures are not known. This work presents a new model for computation of interaction parameters between the beads in coarse-grained dissipative particle dynamics (DPD) simulation that is properly calibrated for amino acids, supports compressibility requirement of water molecules, and accounts for subtle differences in the structure of amino acids and the charge in the side chain of charged amino acids. This new model is referred to as Structure Independent Molecular Fragment Interfuse Model, abbreviated as SIMFIM, because it accounts for specific interactions between different beads, which represent molecular fragments of the amino acids, in calculating nonbonded interaction parameters in the absence of knowing the actual peptide structure. The electrostatic interactions are incorporated in this model by using a normal distribution of charges around the center of the beads to prevent the collapse of oppositely charged soft beads. The uniquely parameterized DPD force field in the SIMFIM model is optimized for a given peptide with respect to the degree of coarse-grained graining for simulating the peptide over long times and length scales. The SIMFIM model was tested in this work using four peptides, namely, TrpZip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai peptides, whose structures were sourced from the Protein Data Bank. The SIMFIM model predicted radius of gyration (R_g) values for the peptides closer to the actual structures as compared to the conventional model, and there was less deviation between the predicted and actual structures of the peptides.



INTRODUCTION

Peptides have a wide range of applications in medicine, drug delivery, and tissue engineering.[1−3] These include, among others, peptides as antimicrobial, anticancer, and cell penetrating agents, as vaccines and biosensors, as self-assembly agents to form hierarchical multiscale structures, and as interference peptides in precision medicine.[4−10] Recently, peptides based on the recognition sites of morphogenetic proteins have received attention as an alternative to protein therapy in tissue engineering to guide differentiation and maturation of transplanted cells to the desired lineage to reduce side effects like immunogenic response and tumorigenicity.[11−15] However, peptides derived from morphogenetic and other proteins have orders of magnitude lower biological activity.[16,17] This lower activity is attributed to misfolding in the absence of other protein sequences (conformational differences between the peptide as a part of the protein versus a free molecule), which further causes peptide aggregation and micelle formation.[17] Furthermore, natural peptides are susceptible to degradation by proteolytic enzymes and their function is limited by short half-life in physiological medium.[18] Therefore, there is a need to redesign peptides that are based on active receptor-binding domains of proteins to address the issues of misfolding, aggregation, enzymatic degradation, and short half-life.[19]

Morphogenetic peptide sequences are, in general, 10−30 amino acids in length. As there are 20 different amino acids, there is an enormous number of 10−30 sequences to be studied for identifying possible sequences with morphogenetic activity. Currently, structure−property data on morphogenetic peptides exist only for a limited number of peptides,[20,21] which is insufficient to make a database for predicting the bioactivity of unknown peptides using machine learning algorithms. Experimental methods like electron microscopy and light scattering are limited by the number of peptide sequences that can be studied in a reasonable time and the complexities of measuring structural properties at the molecular scale.[22] Classical molecular dynamics (MD) or all-atom (AA) simulations can predict the physicochemical properties of peptides with unknown structures, but they are limited by length scales and require long computational times.[23] Therefore, there is a need to

Received: November 29, 2023
Revised: March 7, 2024
Accepted: April 2, 2024
Published: April 12, 2024

develop robust computational models to simulate the conformation of unknown peptides of different sequences toward the creation of a database for the physicochemical properties of morphogenetic peptides.

Peptides, unlike synthetic polymers, exhibit diverse chemical groups for energetic interaction including nonpolar, polar, neutral, hydrophobic, hydrophilic, amphipathic, charged, hydrophobic−charged, and hydrophilic−charged groups. Coarse-grained and mesoscopic simulations have been used to study the structure and conformation of peptides in aqueous solution over large times and length scales within computationally practical time scales.[17,24−31] As short-range interactions dominate between the beads in coarse-grained models, they have the benefit of smoothing the potential energy function of the system.[32,33] However, coarse-grained simulations require robust models to account for secondary interactions between the beads, as the interplay among these interactions significantly influences the structure and conformation of the simulated peptides. Among mesoscale methods, the stochastic dissipative particle dynamics (DPD) simulation, commonly used for modeling the self-assembly of macromolecular systems,[34] uses soft dynamics of spherical beads that interact to form secondary bonds via pairwise additive force fields. The DPD method has been used to simulate dynamics and structure of macromolecular systems after Groot and Warren related the conservative force terms to the Flory−Huggins interaction parameter for pairwise interaction between the beads, hereafter referred to as χ_{ij} for beads i and j.[35]

Peptides have been studied extensively by computational methods using different force fields to account for bead interactions.[17,22,24,28−31,36] Marrink et al. introduced the MARTINI force field as a computational tool for conducting coarse-grained molecular dynamics simulations of large molecules. This approach was developed by computing the partition free energies between distinct polar and apolar molecular fragments.[37] Subsequently, it was expanded to encompass amino acids, peptides, and proteins as well.[30,31] In this framework, the amino acid backbone is represented by a single polar bead in the simulation, whereas the side chain of the amino acid is represented by one or more beads contingent on its size and chemical properties. For instance, the amino acid lysine is represented by three beads: a polar backbone bead, an apolar bead for the hydrocarbon portion of the side chain, and a charged bead for the NH_3^+ terminal group of the side chain. This strategy entails predefined values for bonded interactions, such as bond lengths, bond angles, and dihedral angles, that mirror the actual conformations of the peptide. These constants are determined via molecular dynamics simulation from the lowest-energy conformation of the peptide in the equilibrium state. Nonbonded interactions in the peptide are defined by a Lennard-Jones potential functional form. Bond stretching and bond angles follow a harmonic potential model, while dihedral angle potentials adopt a shifted dihedral functional form. However, some amino acids are not distinguishable within the MARTINI approach. For example, leucine and isoleucine, as well as serine and threonine, share the same coarse-grained model and bead types in the MARTINI force field. As the types of bead−bead interactions in the MARTINI force field are limited to charged, polar, apolar, and nonpolar interactions, this model fails to accurately capture subtle differences between certain amino acids with partial similarities. Furthermore, the use of hard spheres with hard-core potential functions makes the MARTINI model computationally time-consuming. It has been noted that the MARTINI model hinders proper protein folding, attributing this effect to the limited types of bead−bead interactions as well as other aspects of the model.[38,39]

In an earlier investigation, Vishnyakov et al. developed a simplified framework for DPD simulation of polypeptides in solution.[25] This framework replicated the hydrogen-bonding interaction, which is pivotal for stabilizing α-helical and β-structural elements by employing detachable Morse bond potentials.[25] However, this approach led to nonrealistic conformations of the peptides as a consequence of the spherical symmetry stemming from the involvement of multiple Morse bonds.[25] Peter et al. introduced an electrostatic polarizable coarse-grained protein for DPD.[29] The polarizability in the amino acids was described by massless drude particles representing two opposite charges, which were transparent to other coarse-grained beads. Being a polarizable force field, the computational process in the Peter et al. model was time-intensive and the explicit water self-interaction parameter in the model did not support the property of compressibility of liquid water.[29] Vaiwala et al. recently used three peptides with known conformations to determine the parameters of his generic force field and used these parameters to simulate the structure of other peptides and proteins with unknown structures.[39] However, his force field required a priori knowledge of the secondary structure of the peptide under simulation. In another study, Choudhury et al. simulated peptides having an α-helical secondary structure using DPD by imposing dihedral angle constraints on the peptide chain.[40] Periole et al. used an elastic network as a structural scaffold to describe and maintain the overall shape of a protein and a physics-based coarse-grained model, like the MARTINI model, to describe the interaction between molecular fragments in the system.[41] As elastic networks need a priori knowledge of the peptide structure to establish the constraints, it is not applicable to the simulation of peptides with unknown structures. Jumper et al. developed a neural network model for predicting the structure of unknown peptides and proteins, but the model could not predict the dynamics of the peptide structure for studying properties like critical micelle concentration (CMC) or peptide self-assembly.[42] There is a need for new models to determine the pairwise interaction parameter between the beads in coarse-grained DPD simulations that do not require a priori knowledge of the secondary structure of the peptide, account for subtle differences in the structure of amino acids, support compressibility of water molecules, account for the charge in the side chain of charged amino acids, and are properly calibrated for amino acids. Such a model should have the capability to simulate many peptide sequences with unknown structures to determine their physicochemical and biological properties for the creation of a database for machine learning.

This work presents a new model named the Structure Independent Molecular Fragment Interfuse Model, hereafter abbreviated as SIMFIM, for DPD simulation of peptides that addresses the above limitations. The novelty of our model is that the matrix of χ_{ij} values for nonbonded pairwise bead interactions of the peptide is determined using the energy of mixing of the molecular fragments, as opposed to using bead−bead solubility parameters in the conventional model. Our model predicts both positive and negative (repulsive and attractive) values for χ_{ij} as opposed to the conventional method, which predicts only positive (repulsive) values, and the model is independent of the knowledge of the actual peptide structure. Moreover, electrostatic interactions are incorporated into our model by using a

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

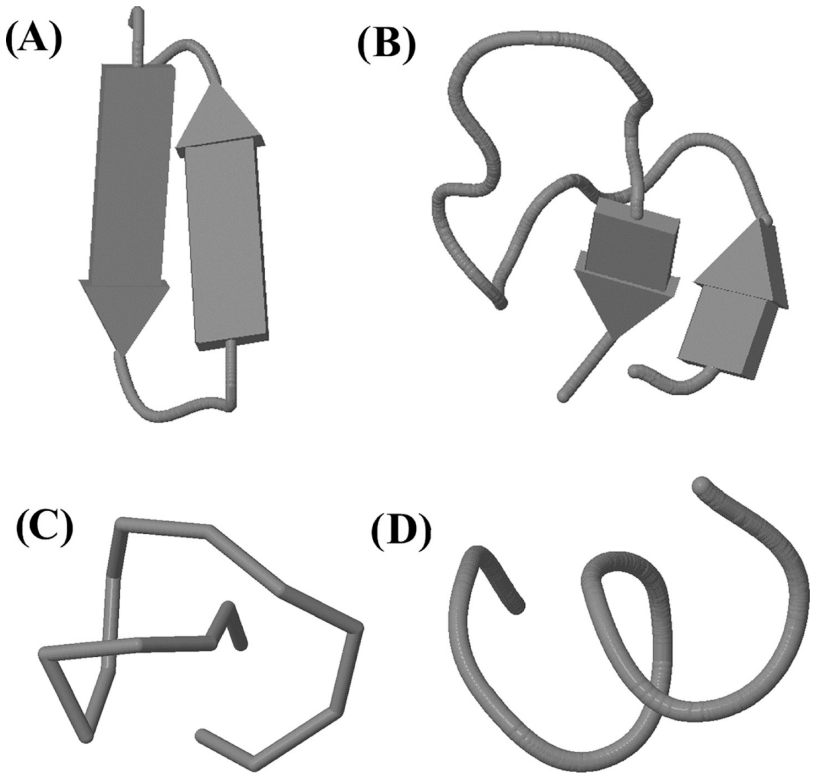


Figure 1. Images of TrpZip2 (A), Rubrivinodin (B), Lihuanodin (C), and IC3-CB1/Gai (D) peptides generated from the PDB coordinates.

normal distribution of charges around the center of the beads to prevent the collapse of oppositely charged soft beads. Furthermore, our uniquely parameterized DPD force field model is optimized for a given peptide with respect to the degree of coarse graining to simulate a given peptide over long times and length scales. The model is tested using four peptides with known structures, namely, TrpZip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai, whose structures were sourced from the Protein Data Bank. TrpZip2 sequence has a β-hairpin;[43] Rubrivinodin, which is a lasso peptide, has four residues that participate in a β sheet;[44] Lihuanodin is a random coil lasso peptide;[45] and IC3-CB1/Gai is an α helical peptide.[46] These four peptides contain amino acids with negative or positive charges and amino acids with hydrogen bonding, polar, or hydrophobic groups; thus, these peptides have diverse physicochemical properties for validating our model. The primary objective of this computational model is not to perfectly simulate the secondary structure of peptides but to create a database of peptides' physicochemical properties, whose native structures are not known, for use in machine learning algorithms. It should be noted that the experimental structures of the four peptides were used to compare the simulated backbone coordinates with the experimental coordinates, but a priori knowledge of the structure of peptides was not required in the simulations and no information from the experimental structures was used in running the simulations.

METHODS AND THEORY

Model Peptides. Four peptides with known structures and different numbers of amino acids were used to evaluate the proposed simulation methodology. These peptides were 12-mer TrpZip2 with amino acid sequence (SWTWENGKWTWK), 18-mer Rubrivinodin with sequence (GAPSLINSEDNPAFPQRD), 15-mer Lihuanodin with sequence (GSKYSDTADESSYRW), and 9-mer third cytoplasmic loop (IC3) of the cannabinoid receptor-1 (CB1) bound to G(alphai1), hereafter abbreviated by IC3-CB1/Gai, with sequence (DIRLAKTLV). The Protein Data Bank identification numbers (PDB ID) for TrpZip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai were 6H7Q, 5OQZ, 7LCW, and 1LVQ, respectively. Information on their structures, acquired by nuclear magnetic resonance (NMR) spectroscopy, is available at the Protein Data Bank (https://www.rcsb.org/). The data for the coordinates of all atoms in the structure of the peptide sequences and the experimental conditions for acquisition of their NMR spectra were extracted from the PDF file using MATLAB. The database had 20 structures each for TrpZip2 and Lihuanodin and one structure each for Rubrivinodin and IC3-CB1/Gai. The structures of these peptides in their lowest-energy states are shown in Figure 1. The NMR-acquired structures of the peptides in their lowest-energy state are hereafter referred to as the "actual" structures for comparison with the simulated structures. The lowest-energy state of the peptides was determined from the chemical shifts in the NMR spectra by fragment density functional theory (DFT) as previously described.[47] Briefly, DFT was used to determine the distribution of electrons in different molecular orbitals of the system by solving the Schrodinger equation. Accurate electronic energy levels were obtained by using an exchange−correlation functional in DFT to account for the correlation and exchange interactions between the electrons. The total energy of the system was obtained as the summation of the kinetic energy of electrons, the potential energy of electron−nucleus interactions, and the potential energy of electron−electron interactions.[47]

Coarse Graining. To simulate larger simulation boxes, groups of atoms or molecular fragments in the simulation box were coarse-grained or mapped into different beads. These atoms were part of one molecule or multiple molecules in the

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

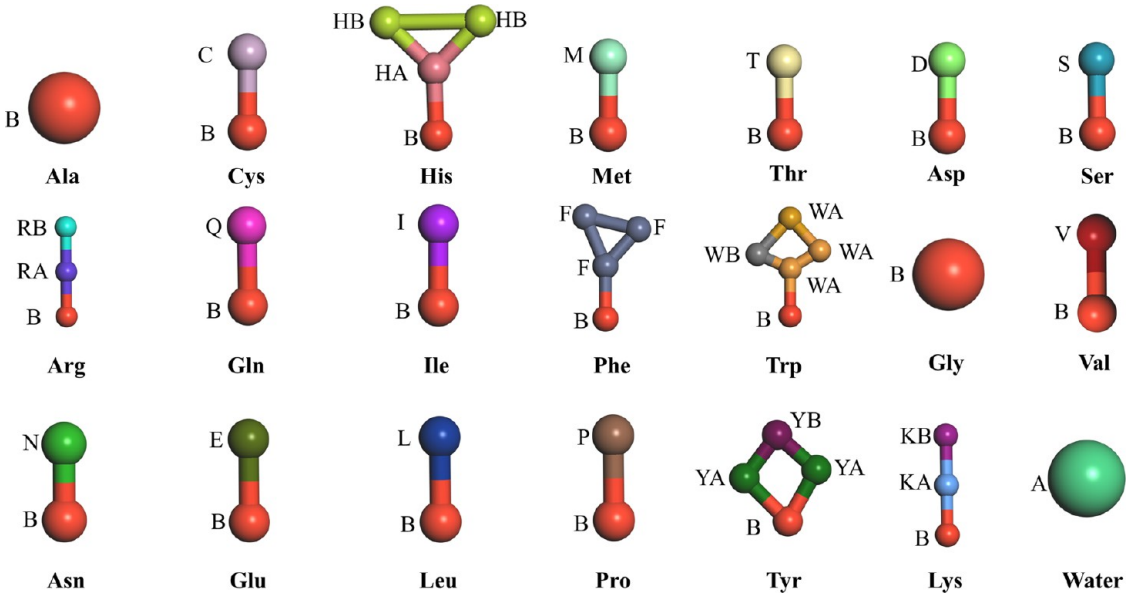


Figure 2. Color-based representation of the coarse-grained beads for the main and side chains of all simulated amino acids and the water bead. See Table S1 in the Supporting Information for the molecular fragments in the amino acids that are represented by the beads.

simulation box. For example, multiple water molecules as the solvents in the box were mapped into one bead. For the peptides in the box, multiple neighboring atoms in the amino acid sequence were mapped into the beads. Coarse graining of amino acids of the peptide was done as described previously in the MARTINI model by Monticelli et al. and de Jong et al.[30,31] Briefly, each amino acid was coarse-grained into one backbone bead and one or multiple side-chain beads depending on the molecular size of the side chain. Figure 2 shows mapping of the amino acids into the corresponding coarse-grained beads. The mass and size of the beads, time, and length scales of the simulation depended on the mapping. Table S1 shows the molecular fragments in the amino acids that are represented by the beads.

DPD Equations of Motion. The positions and velocities of the soft beads are governed by Newton's equations of motion of the beads in the simulation box as follows

$$\frac{dr_i}{dt} = v_i,\ m_i\frac{dv_i}{dt} = f_i \tag{1}$$

where r_i, m_i, v_i, and f_i represent the position vector, mass, velocity vector, and total force acting on the ith bead, respectively. The total force f_i and its components in a cutoff radius R_C around the ith bead are given by the following set of equations

$$f_i = \sum_{j\neq i}^{n} (f_{ij}^{C} + f_{ij}^{D} + f_{ij}^{R}) \tag{2}$$

$$f_{ij}^{D}(r_{ij}) = -\gamma[\omega^{D}r_{ij}](e_{ij}\cdot v_{ij})e_{ij} \tag{3}$$

$$f_{ij}^{R} = \sigma[\omega^{R}(r_{ij})]\theta_{ij}e_{ij} \tag{4}$$

$$f_{ij}^{c} = \alpha_{ij}\left(1 - \frac{r_{ij}}{R_c}\right)e_{ij} \tag{5}$$

where f_{ij}^{D}, f_{ij}^{R}, and f_{ij}^{C} represent the dissipative, random, and conservative components of the force vector joining ith and jth beads, respectively; parameters r_{ij} and e_{ij} represent the magnitude and unit vector joining ith and jth beads, respectively; and the parameter α_{ij} is the interaction parameter that describes the maximum repulsion between the interacting ith and jth beads. The parameter α_{ij}, which is assumed to be independent of the properties of the mixture, quantifies the extent of the interaction between different beads in the DPD simulation. It is assumed that the conservative force vanishes when the distance between ith and jth beads is more than the cutoff radius R_C.[48] R_C is defined as the side of a cube enclosing three beads for all mappings as described previously.[49] The parameter α_{ii}, which is dependent on the chosen mapping, is the self-repulsion parameter, and it is adjusted to mimic the compressibility of the system. The soft molecular fragments or beads in DPD occupy a certain volume, which is accounted for by adjusting the corresponding α_{ii} values at each mapping. The parameter α_{ii} is adjusted to achieve the desired compressibility in response to changes in the thermodynamic properties of the system. Increasing α_{ii} values makes it more difficult to compress the beads, thus decreasing the compressibility and vice versa. The previously reported values of α_{ii} were used in the simulations.[48] The parameters f_{ij}^{D} and f_{ij}^{R} are weight functions that are dependent on the radius. The parameter f_{ij}^{D} models the viscous drag on a bead due to interaction with the surrounding beads, while f_{ij}^{R} represents the force exerted on a bead due to thermal noise in the system.[50] The constants γ and σ are the magnitudes of dissipative and random forces, respectively. The parameter θ_{ij} is a random variable with a Gaussian probability distribution to account for the inherent random Brownian motion of the beads in the system. The parameter e_{ij} is the unit vector joining the centers of beads i and j. The parameters ω^{R} and ω^{D} are the weight functions of the random and dissipative forces to provide the range of interactions, respectively.[51] The above parameters and functions satisfy the fluctuation dissipation theorem,[35,51] which ensures that the random and dissipative forces act as a thermostat for the simulation to maintain a constant temperature. The cutoff radius R_C, which is dependent on the mapping, was calculated as described previously.[48,49]

Force Field Parameterization. The interaction parameter for nonbonded interactions, α_{ij}, is calculated by relating it to χ_{ij}

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

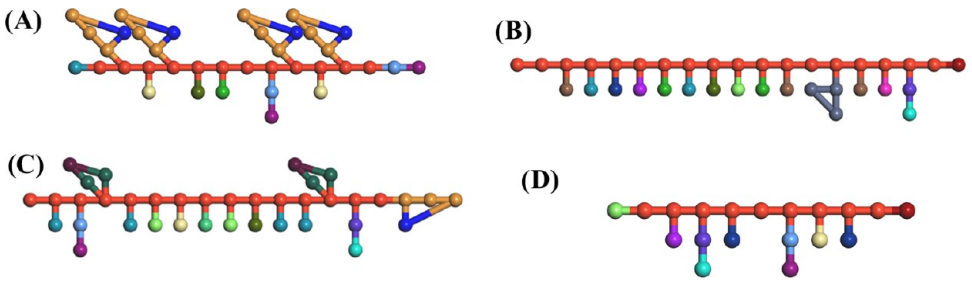


Figure 3. Bead representations of the coarse-grained molecular structures of TrpZip2 (A), Rubrivinodin (B), Lihuanodin (C), and IC3-CB1/Gai (D) peptide mesomolecules.

by the following relation derived from the equation of state of the system by Groot and Warren[35]

$$\alpha_{ij} = \alpha_{ii} + 3.27\chi_{ij} \tag{6}$$

In the conventional model, χ_{ij} is calculated from the solubility parameters of the molecular fragments representing beads i and j by the following equation[24,27]

$$\chi_{ij} = \frac{v}{RT}(\delta_i - \delta_j)^2 \tag{7}$$

where δ_i and δ_j represent the solubility parameters of beads i and j, respectively, v is the molar volume of each bead, R is the universal gas constant, and T is temperature. The solubility parameters of the fragments are then determined using molecular dynamics simulation as previously described.[52] In the conventional model of calculating χ_{ij}, the interaction between the molecular fragments representing ith and jth beads is not realistically taken into account, as the solubility parameters are independent of the interacting molecular fragments of the beads. In this work, we propose a new approach to calculate χ_{ij} using a model based on interfused molecular fragments that is independent of the knowledge of the actual peptide structure. This new model for the calculation of χ_{ij} is hereafter referred to as the SIMFIM model. The χ_{ij} for mixing two molecular fragments, representing two beads in the simulation, was calculated by the blends module of the materials studio software (BIOVIA, Dassault Systèmes, Materials Studio, 23.1.0.3829, San Diego: Dassault Systèmes, 2023).[53,54] The blends module is used to predict the physical and chemical properties of multicomponent systems like, among others, thermal conductivity, diffusion coefficient, and energy of mixing using the COMPASS force field. The COMPASS force field accounts for all energetic interactions including nonpolar, polar, neutral, hydrophobic, hydrophilic, amphipathic, charged, hydrophobic−charged, and hydrophilic−charged groups. The molecular fragments represented by different types of beads are shown in Table S1 of the Supporting Information File. As the blends module accepts fragments where all atoms in the fragments satisfy their valency, the head and tail atoms in the fragments were identified for valency to be maintained and for the fragment to be treated by the blends module as a monomer on a longer molecule. Next, the molecular fragments were mixed and the change in the energy of mixing was determined.[55] The interactions between these molecular fragments were assumed to be governed by the COMPASS force field.[56] The χ_{ij} was determined by dividing the energy of mixing by RT.

The bond length D, bond angle θ, and dihedral angle ϕ of the bonded interactions were defined through different functional forms. Bond stretching between two beads i and j was defined as a soft harmonic potential E_{bs}. Bond angles between three beads i, j, and k had a cosine harmonic potential E_{ba}. The dihedral angle formed by four beads i, j, k, and l, which is the angle between the planes that contained the first three (i, j, k) and the second three (j, k, l) beads, was defined through a shifted dihedral potential E_{da}. The force constant k_0, the equilibrium bond length D_0, the equilibrium bond angle θ_0, and the equilibrium dihedral angle ϕ_0 were taken from the values mentioned in the MARTINI model. These values were determined through molecular dynamics simulation and used in the following equations[30,31]

$$E_{bs} = \frac{k_0}{2}(D - D_0)^2 \tag{8}$$

$$E_{ba} = \frac{k_0}{2}(\cos\theta - \cos\theta_0)^2 \tag{9}$$

$$E_{da} = \frac{k_0}{2}(1 + \cos(\phi - \phi_0)) \tag{10}$$

Assigning point charges to charged soft beads is a limitation as it leads to the collapse of two oppositely charged beads into the same point in the simulation box, which results in the unphysical stretching of the bonds and the release of excessive amounts of energy. This effect was avoided in the simulation by spreading the charge within a specified volume using a normal distribution around the center of the bead. The following functional form proposed by Warren et al., known as the ERFR form, was used to describe the electrostatic interaction in systems with distributed charges[57]

$$E_E = c\frac{q_i q_j \operatorname{erf}\left(\frac{R_{ij}}{\beta}\right)}{\epsilon R_{ij}} \tag{11}$$

where q_i and q_j are the charges on beads i and j, respectively; C is a conversion factor; ϵ is the relative dielectric constant; R_{ij} is the distance between beads i and j; and β is a parameter related to the length scale of the simulation, which depends on the mapping.

Mesomolecule Construction. For a given mapping, all beads were defined with the same mass and size. Considering that the conservative force vanishes when the distance between ith and jth beads is more than the cutoff radius R_C[48] and defining R_C as the side of a cube enclosing three beads for all mappings,[49] the values of mass and size of the beads were calculated as previously described using the following approach.[49] As an example of 3−1 mapping, three water molecules were mapped into a bead. With the volume occupied by one water molecule equal to 30 Å^3, the volume occupied by one bead is 90 Å^3, and the volume occupied by three such beads is 270 Å^3. If the three beads are enclosed within a cube with sides equal to R_C, then R_C^3 is 270 Å^3, and as a result, R_C = 6.46 Å. R_C was similarly calculated

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

for the other mappings. The radius of the beads was taken to be equal to half of the R_C, and the bead mass depended on the number of water molecules mapped into one bead (for 3−1 mapping, the bead mass was 3 × 18 amu = 54 amu). The backbone of the peptide was constructed first, followed by the addition of side-chain beads based on the amino acid sequence. The side chain of amino acids with a closed-ring structure like tryptophan and phenylalanine was constructed by a simple perl script.[58] A negative or positive charge was assigned to certain beads based on the actual charge on the amino acid at the pH of the experimental conditions of the NMR spectroscopy of the peptide. The constructed mesomolecules of the four simulated peptides are shown in Figure 3.

Mesostructure Construction. A mesostructure template or simulation box with dimensions of 200 Å × 200 Å × 200 Å was constructed. These dimensions were chosen to avoid the finite size effect, which arises when the number of beads in a simulation box is insufficient to provide a statistically appropriate depiction of the conditions of the actual physical system.[59] The box size was at least an order of magnitude larger than the actual radius of gyration of the simulated peptide.[60] The box was filled with water beads and the peptide mesomolecule, and a periodic boundary condition was applied to the simulation box. The concentration of the peptide in the aqueous solution was adjusted so that there was only one peptide in the box. To avoid any bias, initially, the peptide was randomly positioned inside the simulation box in a randomly selected three-dimensional (3D) conformation. As certain amino acids in the peptide were charged in aqueous solution at the simulation pH corresponding to the experimental conditions in NMR spectroscopy, oppositely charged water beads were placed in proximity to the charged beads to maintain electrical neutrality.

DPD Simulation. The energy of the mesostructure was initially too high, as it was created with a random initial conformation. Therefore, the mesostructure was geometrically optimized by removing any discrepancy between the force field and the random structure by adjusting the bond lengths, bond angles, and dihedral angles. The optimization was done using the geometry optimization task in the mesocite module of Materials Studio (BIOVIA, Dassault Systèmes) by an iterative procedure in which the geometry of the system was refined at each iteration to minimize the overall system's energy. The iterative process was terminated when the threshold value in energy between two iterations had been reached.

The refined mesostructure was simulated using the Dissipative Particle Dynamics simulation task in the mesocite module of Materials Studio (BIOVIA, Dassault Systèmes). Mesocite DPD simulations were performed in the NVT ensemble with the temperature set at 298 K, and the simulation algorithm acted as a thermostat such that the dissipative and random forces satisfied the fluctuation−dissipation theorem.[48] The mesostructure underwent a DPD run for 10,000 ps, which was sufficiently longer than the time for equilibration of the system. The simulation reached equilibrium when the mean square displacement of the water beads increased linearly monotonically with the simulation time.[61] A time step was chosen that was not too small to unnecessarily increase the computational time and not too large to sacrifice accuracy. This optimum time step, which depended on mapping, was approximately $0.005*\tau$, where τ was the simulation time scale. The position and velocity of each bead, determined using the velocity-Verlet algorithm,[62] are defined by the following equations

$$r_i(t + \Delta t) = r_i(t) + \Delta t v_i(t) + \frac{1}{2} f_i(t) \Delta t^2 \quad (12)$$

$$v_i(t + \Delta t) = v_i(t) + \lambda \Delta t f_i(t) \quad (13)$$

where r_i, v_i, and f_i are the position, velocity, and force vectors acting on bead i at time t, respectively. The tuning parameter λ was set at the optimum value of 0.65 as defined by Groot and Warren.[35,48]

Simulation Study, Design, and Analysis. For each peptide, simulations were performed for 5 different mappings of 2, 3, 4, 5, and 6 water molecules into one bead. The mass and radius of the beads, and the length scale of the simulation were adjusted based on the mapping as previously described.[48] Two models were used to calculate the interaction parameter, α_{ij}, from the energy of mixing as described in the Force Field Parameterization section, namely, the conventional and SIMFIM models. Two different functional forms of electrostatic interactions were used, namely, beads with their point charge at their center of mass (Coulombic) and beads with their charge distributed normally around their center of mass (ERFR). The reference group for each peptide in the simulations was beads with no charges (None). The coordinates of the beads at different time points were obtained using a perl script.[58] Data were collected at time points after equilibration of the mesostructure in the simulation box. At each time point, the coordinates of the beads after equilibration were used to determine the biophysical dynamic properties of the peptide using in-house codes in MATLAB as described below.

Radius of Gyration of Simulated Peptides. The coordinates of the backbone beads were used to calculate the radius of gyration (R_g) of the simulated peptide at different time points using the following equation[63]

$$R_g = \sqrt{\frac{\sum_{i=1}^{N} ((x_i - \bar{x})^2 + (y_i - \bar{y})^2 + (z_i - \bar{z})^2)}{N}} \quad (14)$$

where x_i, y_i, and z_i are the coordinates of the backbone beads of the simulated peptide at any time point; $\bar{x}$, $\bar{y}$, and $\bar{z}$ were the coordinates of the center of mass of the backbone beads in the simulated structure; and N was the total number of amino acids in the peptide sequence. It should be noted that the simulated R_g of the peptide at each time point is denoted by "R_g" and the average simulated R_g over all time points is denoted by "$\langle R_g \rangle$".

Root-Mean-Square Deviation of Simulated Peptides. The root-mean-square deviation (RMSD) is a measure of the deviation of the coordinates of the beads of the simulated peptide from their actual (experimentally acquired by NMR) coordinates as reported in the PDB database. The RMSD of the peptide was determined using the following steps as shown in Figure S1 of the Supporting Information File. First, the coordinates of the α-carbons in the backbone of the simulated (red) and actual (blue) peptide structures were determined. Next, the centers of the coordinate systems of the actual/simulated peptide structures were shifted so that one end of the backbone of both peptide structures had coordinates (0,0,0). Then, the first backbone−backbone bond of the simulated peptide structure was rotated to align with the line joining the two α-carbons of the backbone of the actual peptide structure. Based on the rotation for alignment, the Euler angles ψ, θ, and φ, which were the components of the rotation about the x, y, and z axes, were calculated. Finally, the coordinates of any point P on the peptide were multiplied by the rotation matrix $\mathrm{Rot}_{XYZ}(\psi, \theta, \varphi)$ to obtain the transformed coordinates P'.[64] The trans-

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

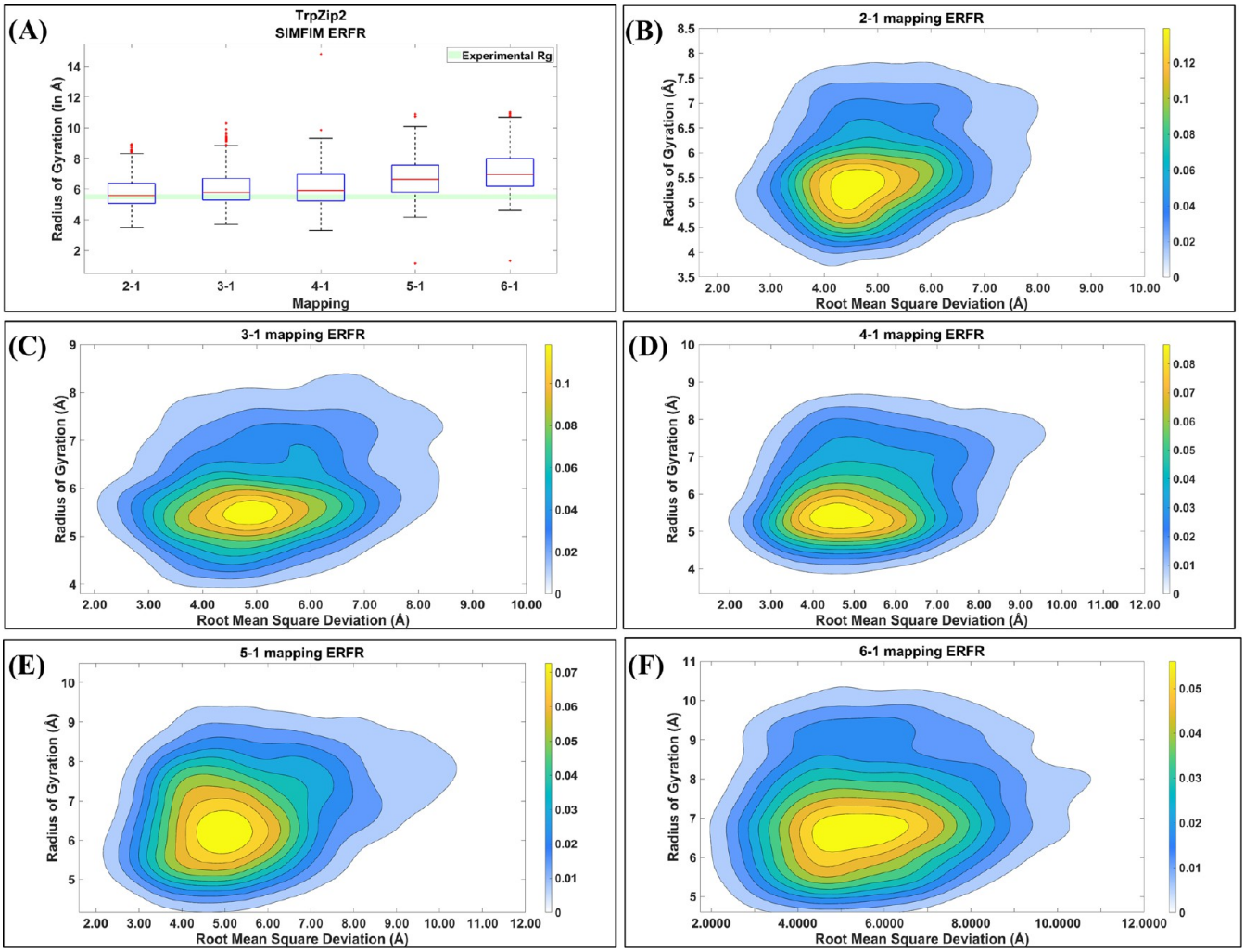


Figure 4. (A) Box plot of R_g ranges for the TrpZip2 peptide simulated at different mappings using SIMFIM model and ERFR functional form for electrostatic interaction; the blue box encloses points that are within the 25th and 75th percentiles of the data, with a red line showing the median value; the black whiskers show the maximum and minimum values in the data set that are not considered as outliers, and outliers appear as red "+" signs; the actual R_g values from the PDB database are shown as a light green horizontal line, with thickness indicating the range. The R_g versus RMSD contour plots simulated at (B) 2−1, (C) 3−1, (D) 4−1, (E) 5−1, and (F) 6−1 mappings.

formation of point P to P' is given by $P' = \mathrm{Rot}_{XYZ}(\psi, \theta, \varphi) \times P$,

where $\mathrm{Rot}_{XYZ}(\psi, \theta, \varphi)$ is given by the following matrix

$$\begin{bmatrix} \cos(\varphi)\cos(\theta) & \cos(\varphi)\sin(\theta)\sin(\psi) - \sin(\varphi)\cos(\psi) & \cos(\varphi)\sin(\theta)\cos(\psi) + \sin(\varphi)\sin(\psi) \\ \sin(\varphi)\cos(\theta) & \sin(\varphi)\sin(\theta)\sin(\psi) + \cos(\varphi)\cos(\psi) & \sin(\varphi)\sin(\theta)\cos(\psi) - \cos(\varphi)\sin(\psi) \\ -\sin(\theta) & \cos(\theta)\sin(\psi) & \cos(\varphi)\sin(\psi) \end{bmatrix} \quad (15)$$

The RMSD of the simulated peptide was calculated using the following equation

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{N} [(x_i - x_i')^2 + (y_i - y_i')^2 + (z_i - z_i')^2]}{N}} \quad (16)$$

where x_i, y_i, and z_i are the transformed coordinates of the backbone beads of the simulated peptide structure at any time point; x'_i, y'_i, and z'_i are the coordinates of the α-carbons of the backbone of the actual peptide structure; and N is the total number of amino acids in the peptide sequence.

Statistical Analysis. Data are expressed as means ± standard deviation (SD). Significant differences between R_g values obtained at different time points were analyzed for statistical significance with respect to mapping using a one-way ANOVA test followed by a Tukey post hoc test.

RESULTS AND DISCUSSION

Mesostructure Equilibration. Once the mesostructure reached equilibrium, the distribution of end-to-end distance vectors in each direction for the simulated peptide in aqueous solution collected at different simulation times should be normally distributed.[65] The normal distribution function in the

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

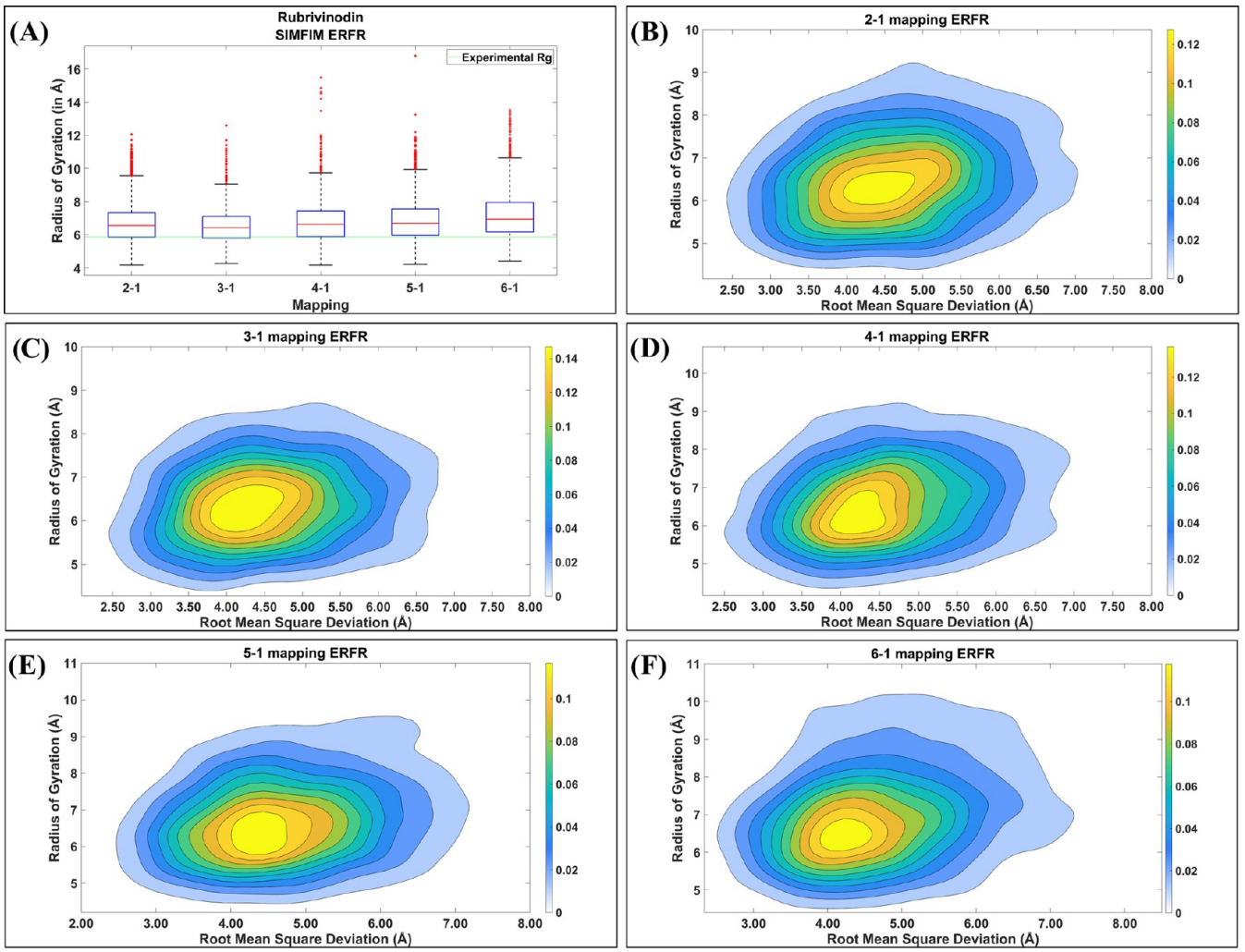


Figure 5. (A) Box plots of R_g ranges for Rubrivinodin peptide simulated at different mappings using SIMFIM model and ERFR functional form for electrostatic interaction; the blue box encloses points that are within the 25th and 75th percentiles of the data, with a red line showing the median value; the black whiskers show the maximum and minimum values in the data set that are not considered as outliers and outliers appear as red "+" signs; the actual R_g value from the PDB database is shown as a thin light green horizontal line. The R_g versus RMSD contour plots at (B) 2−1, (C) 3−1, (D) 4−1, (E) 5−1, and (F) 6−1 mappings.

x-direction is given by the following equation, where the mean is at zero

$$P_{1d}(N, x) = \frac{1}{\sigma\sqrt{2\Pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \quad (17)$$

The forms of the normal distribution function in the y and z directions should be the same as the above equation. The coordinates of the backbone beads at 100,000 different time points during the simulation were collected using a perl script. Figures S2−S4 in the Supporting Information File show the normalized distribution of end-to-end vectors for the TrpZip2 peptide in x, y, and z directions, respectively, based on the coordinates of the backbone beads at different simulation time points. The directional end-to-end distributions from the simulation (blue) showed a good fit to the normal distribution function (red) by MATLAB. Figures S2−S4 show that the simulated normalized distribution of end-to-end vectors for the TrpZip2 peptide was the same in the x, y, and z directions. The simulated values of the mean and SD in each direction were $\mu_x = 1.00 \times 10^{-2}$, $\sigma_x = 8.24$, $\mu_y = -9.95 \times 10^{-2}$, $\sigma_y = 8.18$, $\mu_z = 6.26 \times 10^{-2}$, and $\sigma_z = 8.27$, which showed that the mean values of the distributions were close to zero in all directions. The distribution of the data from the simulation was narrower than the normal distribution, which was attributed to the finite, small length of the peptide sequences with constraints in their geometries unlike infinitely long chains with a normal distribution of end-to-end vectors.

Effect of Mapping. The beads were coarse-grained at different levels of mapping from 2−1 to 6−1 to determine the optimum mapping for each peptide. Figures 4−7 show the simulated range of R_g and RMSD values of TrpZip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai peptides, respectively, using box plots and probability histograms as a function of mapping using SIMFIM model and ERFR functional form of the electrostatic interaction for charged amino acids. For TrpZip2 and Lihuanodin peptides, the 20 NMR-acquired structures of each peptide were used to determine their actual range of R_g values, and the range is shown in Figures 4A or 6A as a light green horizontal line with thickness showing the range. For Rubrivinodin and IC3-CB1/Gai peptides, the single NMR-acquired structure of each peptide was used to determine the actual R_g and the value is shown in Figures 5A or 7A as a thin light green horizontal line. The R_g values of the simulated

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

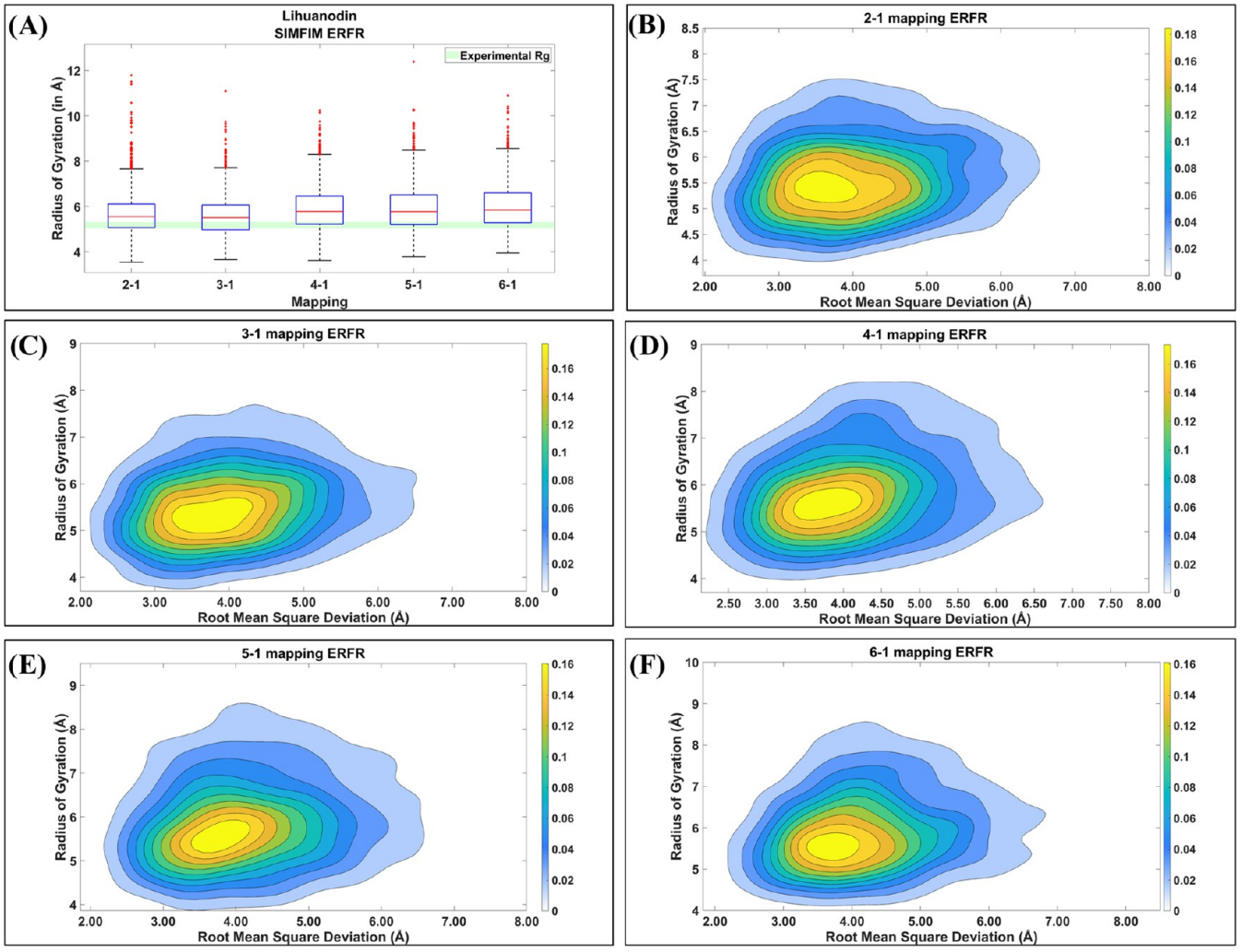


Figure 6. (A) Box plots of R_g ranges for Lihuanodin peptide simulated at different mappings using SIMFIM model and ERFR functional form for electrostatic interaction; the blue box encloses points that are within the 25th and 75th percentiles of the data, with a red line showing the median value; the black whiskers show the maximum and minimum values in the data set that are not considered as outliers and outliers appear as red "+" signs; the actual R_g values from the PDB database are shown as a light green horizontal line, with thickness indicating range. The R_g versus RMSD contour plots at (B) 2−1, (C) 3−1, (D) 4−1, (E) 5−1, and (F) 6−1 mappings.

peptides were determined using the coordinates of the backbone beads, whereas the R_g values of the actual peptide structures were calculated from the coordinates of the α-carbons of the peptide backbone from PDB data. Data from the simulations were collected for 2000 different time points after the equilibration of the mesostructures. The box plots in Figures 4−7 show that the 2−1 and 3−1 mappings predicted $\langle R_g \rangle$ values closer to the actual values for all four peptides. At 3−1 mapping, the predicted values of $\langle R_g \rangle$ for TrpZip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai peptides were 6.10, 6.51, 5.59, and 4.68 Å, respectively, compared to the actual values of 5.83, 5.87, 5.17, and 4.37 Å.

The contour plots for the joint probability distributions of R_g and RMSD for the four peptides are also shown in Figures 4−7. The color bar scales to the right show the probability depth of R_g and RMSD, and the numbers next to the color bars are dimensionless probability values. For 2−1 and 3−1 mappings, RMSD and R_g values corresponding to the maximum probability, hereafter abbreviated as mpRMSD and mpRg, were in the lower ranges of RMSD and R_g values. Furthermore, the distributions of RMSD and R_g were narrower for the lower mappings of 2−1 and 3−1 as compared to those for higher mappings. As the mapping increased, $\langle R_g \rangle$ values increased (Figures 4A,5A, 6A, and 7A) for all peptides. For the TrpZip2 peptide (Figure 4) with 2−1 mapping, mpRMSD and mpRg ranges were 4.12−5.14 and 4.85−5.67 Å, respectively (Figure 4B), and for 3−1 mapping, the ranges were 4.38−5.38 and 5.25−5.80 Å with the actual R_g of 5.83 Å close to the simulated upper bound for both mappings (Figure 4C); for higher mappings, mpRg and mpRMSD ranges were significantly above those of 2−1 or 3−1 mappings (Figure 4D−F). For example, at 6−1 mapping, mpRMSD and mpRg ranges were 4.33−6.56 and 6.14−7.11 Å, respectively (Figure 4F). The simulated maximum probabilities were larger for 2−1 and 3−1 mappings as compared to those for higher mappings.

For the Rubrivinodin peptide (Figure 5), the simulated $\langle R_g \rangle$ was at a minimum for 3−1 mapping. The simulated $\langle R_g \rangle$ was 6.71 Å at 2−1 mapping, decreased to 6.51 Å at 3−1, and then increased to 6.75, 6.89, and 7.24 Å for 4−1, 5−1, and 6−1 mappings, respectively. For 2−1 mapping (Figure 5B), mpRMSD and mpRg ranges were 3.95−4.90 and 5.91−6.68 Å, respectively, and for 3−1 mapping (Figure 5C), the ranges

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

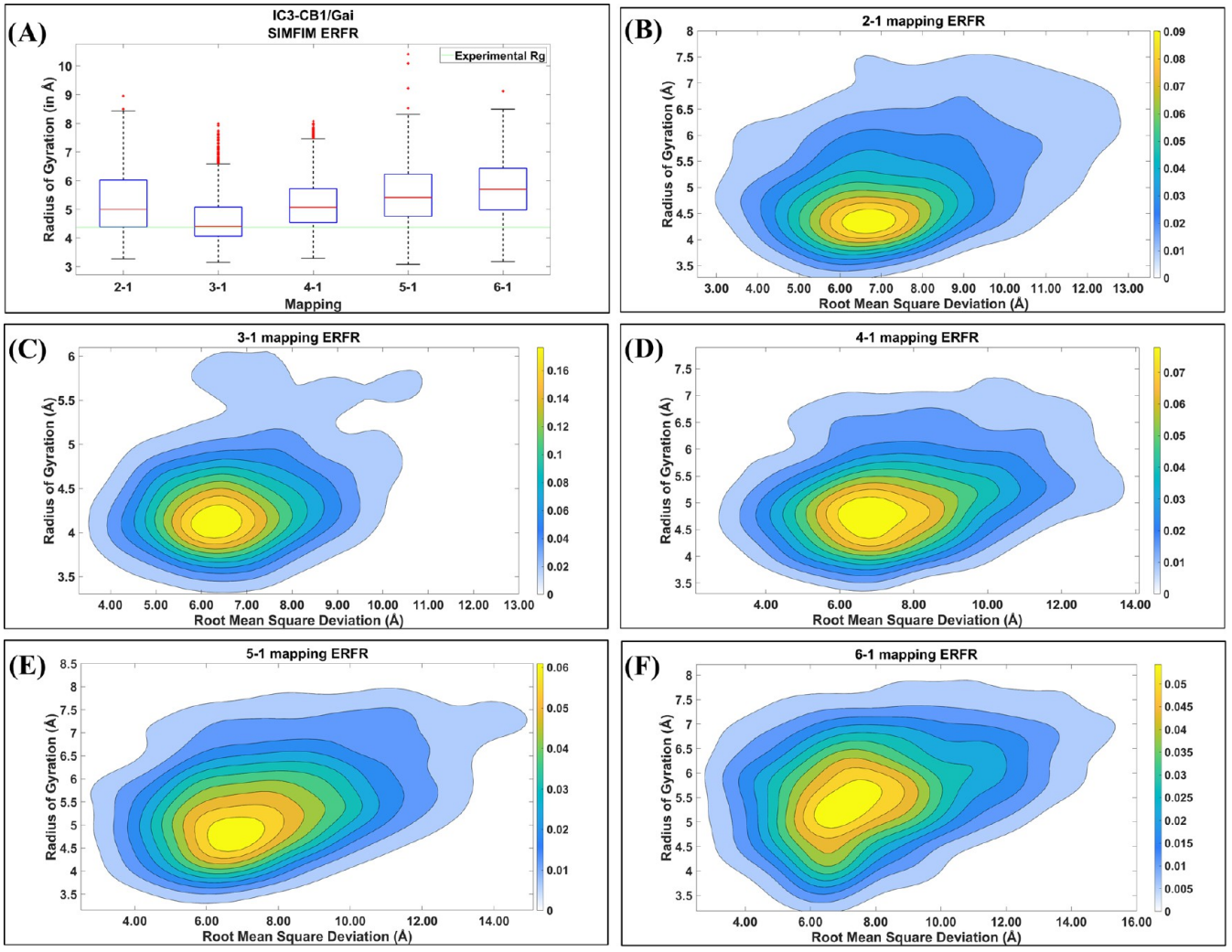


Figure 7. (A) Box plots of R_g ranges for the IC3-CB1/Gai peptide simulated at different mappings using SIMFIM model and ERFR functional form for electrostatic interaction; the blue box encloses points that are within the 25th and 75th percentiles of the data, with a red line showing the median value; the black whiskers show the maximum and minimum values in the data set that are not considered as outliers and outliers appear as red "+" signs; the actual R_g value from the PDB database is shown as a thin light green horizontal line. The R_g versus RMSD contour plots at (B) 2−1, (C) 3−1, (D) 4−1, (E) 5−1, and (F) 6−1 mappings.

Table 1. Comparison of χ_{ij} Values for Two Molecular Fragments Calculated by the Conventional and SIMFIM Models

molecular fragment i	molecular fragment j	χ_{ij} (conventional)	χ_{ij} (SIMFIM)
KA	A	24.6	8.2
KB	A	11.6	−0.4

were 3.78−4.70 and 5.75−6.82 Å, with the actual R_g of 5.87 Å close to the simulated lower bound for both mappings. At a higher mapping of 6−1, mpRMSD and mpRg ranges were 3.90−4.59 and 6.03−6.89 Å, respectively (Figure 5F). The actual R_g of Rubrivinodin was outside the simulated R_g range for higher 4−1, 5−1, and 6−1 mappings.

For the Lihuanodin peptide (Figure 6), the 2−1 and 3−1 mappings predicted $\langle R_g \rangle$ values closer to the actual value of 5.17 Å (Figure 6A). For 2−1 mapping (Figure 6B), mpRMSD and mpRg ranges were 3.29−4.06 and 5.12−5.72 Å, respectively, and for 3−1 mapping (Figure 6C), the ranges were 3.23−4.32 and 4.95−5.71 Å, with the actual R_g of 5.17 Å within the simulated range for both mappings; for higher mappings, mpRg ranges were outside the actual $\langle R_g \rangle$, and mpRMSD ranges were slightly higher than the lower mappings (Figure 6D−F). For example, at 6−1 mapping, the mpRg range was 5.18−5.89 Å, which did not contain the actual $\langle R_g \rangle$ of Lihuanodin, and the mpRMSD range was 3.34−4.11 Å (Figure 6F).

For the IC3-CB1/Gai peptide (Figure 7), only the 3−1 mapping predicted an $\langle R_g \rangle$ value close to the actual value of 4.37 Å (Figure 7A). For 2−1 mapping (Figure 7B), mpRMSD and mpRg ranges were 6.05−7.35 and 4.12−4.58 Å, respectively, with the actual value of 4.37 Å within the simulated R_g range, and for 3−1 mapping (Figure 7C), the ranges were 5.78−6.90 and 3.94−4.32 Å, with the actual R_g close to the simulated upper bound; for higher mappings, the mpRg range was outside the actual $\langle R_g \rangle$, and the mpRMSD range was slightly higher than the 3−1 mapping (Figure 7D−F). For example, at 6−1 mapping, mpRMSD and mpRg ranges were 3.12−8.20 and 4.88−5.86 Å, respectively (Figure 7F). The actual R_g for IC3-CB1/Gai was within the simulated range only for the 3−1 mapping.

These simulation results demonstrate that lower mappings of 2−1 and 3−1 provided higher resolution, which implied that the mass of beads at lower mappings was closer to the average mass

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

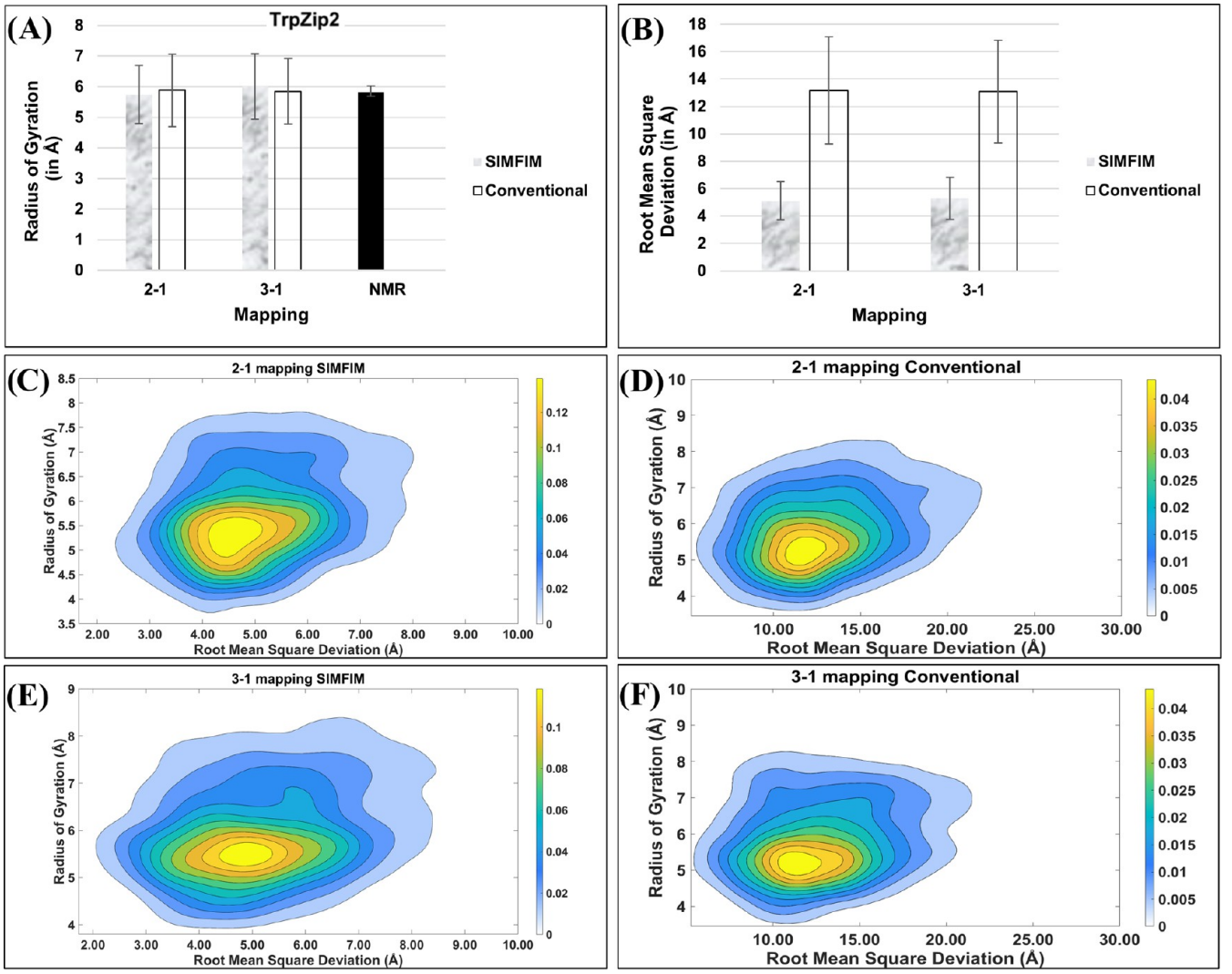


Figure 8. Comparison of (A) $\langle R_g \rangle$ and (B) ⟨RMSD⟩ values of the TrpZip2 peptide simulated with the conventional and SIMFIM models and ERFR functional form for electrostatic interaction for 2−1 and 3−1 mappings; the simulated $\langle R_g \rangle$ and ⟨RMSD⟩ values are compared with the actual value from the PDB database. The R_g versus RMSD contour plots simulated with (C) the conventional and (D) SIMFIM models for 2−1 mapping. The R_g versus RMSD contour plots simulated with (E) the conventional and (F) SIMFIM models for 3−1 mapping. The lowest (5.69 Å) and highest (6.03 Å) experimental R_g values among the 20 structures from the PDB database are shown as whiskers under the "NMR" column. The error bars show SD for R_g and RMSD.

of the molecular fragments represented by the beads. The simulation times were reduced at higher mappings but at the expense of less accurate predictions for $\langle R_g \rangle$ and RMSD of the peptides. Further, these results show that the predicted $\langle R_g \rangle$ values for smaller peptides were closer to the actual value. From one-way ANOVA analysis, a statistically significant difference was observed in the measured R_g of the four peptides between at least two mappings at different time points [$F(49{,}996, p < 0.05) = 424.08$ for TrpZip2, $F(49{,}996, p < 0.05) = 86.57$ for Rubrivinodin, $F(49{,}996, p < 0.05] = 74.06$ for Lihuanodin, $F(49{,}996, p < 0.05) = 298.59$ for IC3-CB1/Gai). From Tukey's post hoc test for multiple comparisons shown in Tables S2−S5 of the Supporting Information File, the $\langle R_g \rangle$ values were significantly different for all mappings in TrpZip2 simulations, whereas significant differences were observed for most of the mappings, but not all, in Rubrivinodin, Lihuanodin, and IC3-CB1/Gai simulations. Many of the p-values in Tables S2−S5 are zero because the size of our data set, which is the number of time points, was very high (10,000).

The data in Figures 4−7 show that the simulated R_g of the peptides increased with mapping from 2−1 to 6−1. The coarse-grained molecular structures of the peptides in Figure 3 indicate that the mass and radius of these beads depend on the mapping. For example, at 2−1 mapping, beads in the simulation box had a mass of 36 amu and a radius of 2.78 Å, whereas at 3−1 mapping, beads had a mass of 54 amu and a radius of 3.23 Å. As mapping was increased, the mass of beads increased, which increased the volume occupied by the beads. Therefore, the larger volume occupied by individual beads at higher mappings explains the increase in simulated R_g values of the peptides constructed from these coarse-grained beads.

Effect of Model for Interaction Parameter Calculations. Tables S6−S15 in the Supporting Information File show the matrix of α_{ij} values calculated by the conventional model and our model (SIMFIM) for all mappings. The difference in the calculated χ_{ij} values between the two models is demonstrated by considering the molecular fragments "KA" ($CH_2-CH_2-CH_2-CH_2$; see Table S1) and "KB" (H−N−H; see Table S1) of lysine

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

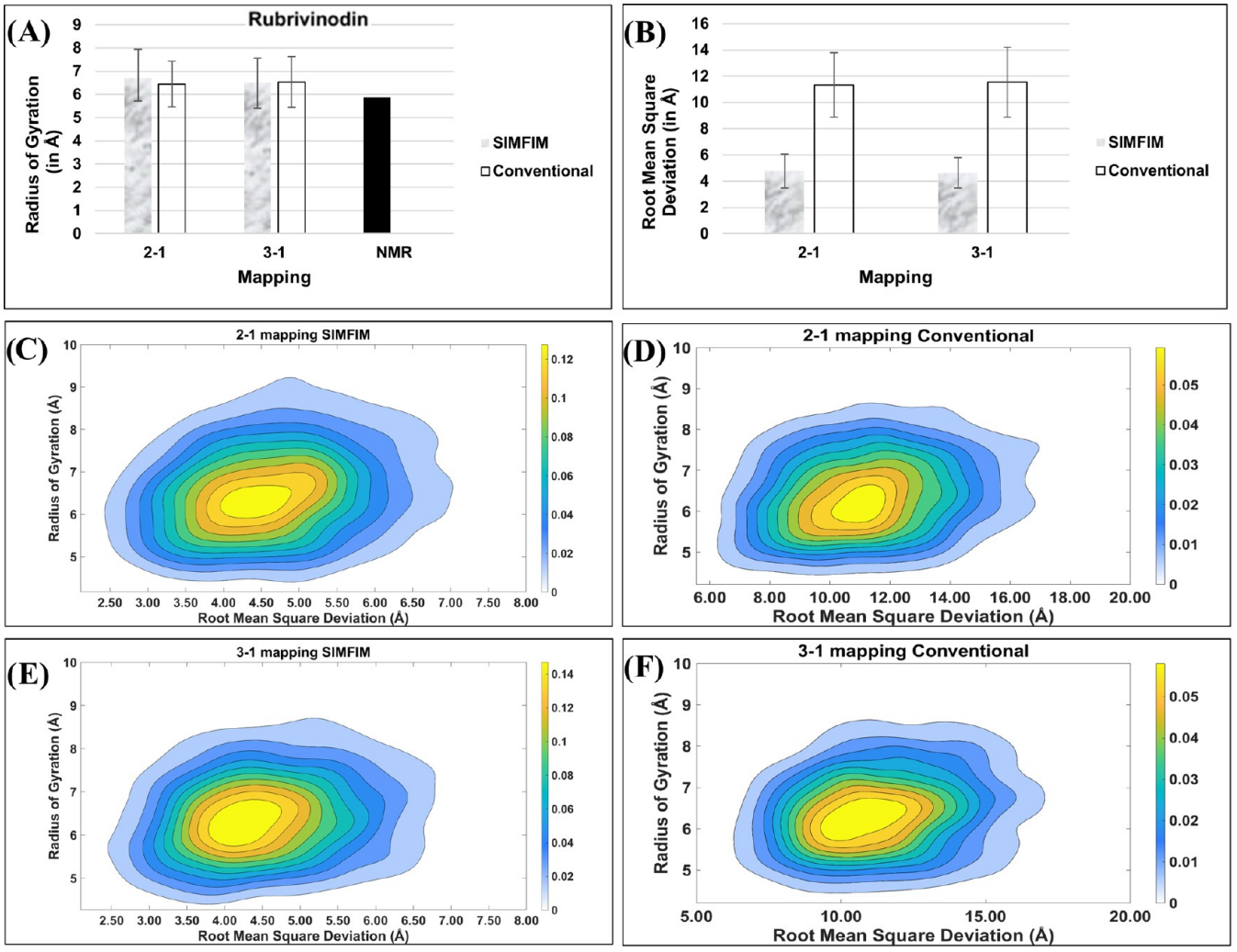


Figure 9. Comparison of (A) $\langle R_g \rangle$ and (B) $\langle$RMSD$\rangle$ values of the Rubrivinodin peptide simulated with the conventional and SIMFIM models and ERFR functional form for electrostatic interaction for 2−1 and 3−1 mappings; the simulated $\langle R_g \rangle$ and $\langle$RMSD$\rangle$ values are compared with the actual value from the PDB database. The R_g versus RMSD contour plots simulated with (C) the conventional and (D) SIMFIM models for 2−1 mapping; the R_g versus RMSD contour plots simulated with (E) the conventional and (F) SIMFIM models for 3−1 mapping. The error bars show SD for R_g and RMSD.

interacting with water, which is "A". The molecular fragment "KA" is a hydrophobic hydrocarbon, whereas "KB" is a hydrophilic positively charged amine group. Table 1 shows χ_{ij} values for the interaction of KA/A and KB/A pairs calculated by using the conventional and SIMFIM models. According to Table 1, the conventional model predicts that KA/A and KB/A interactions are both hydrophobic ($\chi_{ij} > 0$), whereas the SIMFIM model predicts that the KA/A interaction is hydrophobic ($\chi_{ij} = +11.60$) and KB/A is hydrophilic [$\chi_{ij} = -0.40$]. This is attributed to the fact that the solubility parameter method is based on differences in occupied volume by the two molecular fragments, which results in "like" interacting with "like" and a positive α_{ij} for any two interacting fragments. Conversely, the SIMFIM method is based on specific energetic interactions like polar−polar, hydrogen-bonding, or charge−charge interactions between two different fragments, which can result in a positive (repulsive) or a negative (attractive) χ_{ij}.

Figures 8−11 compare R_g and RMSD values simulated using the conventional and SIMFIM models for 2−1 and 3−1 mappings, respectively. Figure 8A,B compares $\langle R_g \rangle$ and $\langle$RMSD$\rangle$ values of the TrpZip2 peptide, respectively, simulated with the two models. The two models predicted similar $\langle R_g \rangle$ for 2−1 and 3−1 mappings, and the simulated values were close to the actual value of 5.83 Å (Figure 8A). SIMFIM model predicted a much lower $\langle$RMSD$\rangle$ for both 2−1 and 3−1 mappings as compared to the conventional model (Figure 8B). mpRMSD ranges for 2−1 mapping for SIMFIM and the conventional models were 4.12−5.14 Å (Figure 8C) and 10.60−13.02 Å (Figure 8D), respectively, whereas for 3−1 mapping, the ranges were 4.35−5.38 Å (Figure 8E) and 10.27−12.81 Å (Figure 8F).

Figure 9A,B compares $\langle R_g \rangle$ and $\langle$RMSD$\rangle$ values of the Rubrivinodin peptide, respectively, simulated with the conventional and SIMFIM models. The two models predicted $\langle R_g \rangle$ values for Rubrivinodin close to the actual value of 5.87 Å for both 2−1 and 3−1 mappings (Figure 9A). For 2−1 mapping, $\langle$RMSD$\rangle$ simulated with SIMFIM was 4.78 Å, which was much lower than the value of 11.35 Å for the conventional model (Figure 9A), and for 3−1 mapping, $\langle$RMSD$\rangle$ for SIMFIM and the conventional models were 4.64 and 11.55 Å, respectively (Figure 9B). mpRMSD ranges for 2−1 mapping for SIMFIM and the conventional models were 3.95−4.90 Å (Figure 9C) and 10.05−11.85 Å (Figure 9D), respectively, whereas for 3−1

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

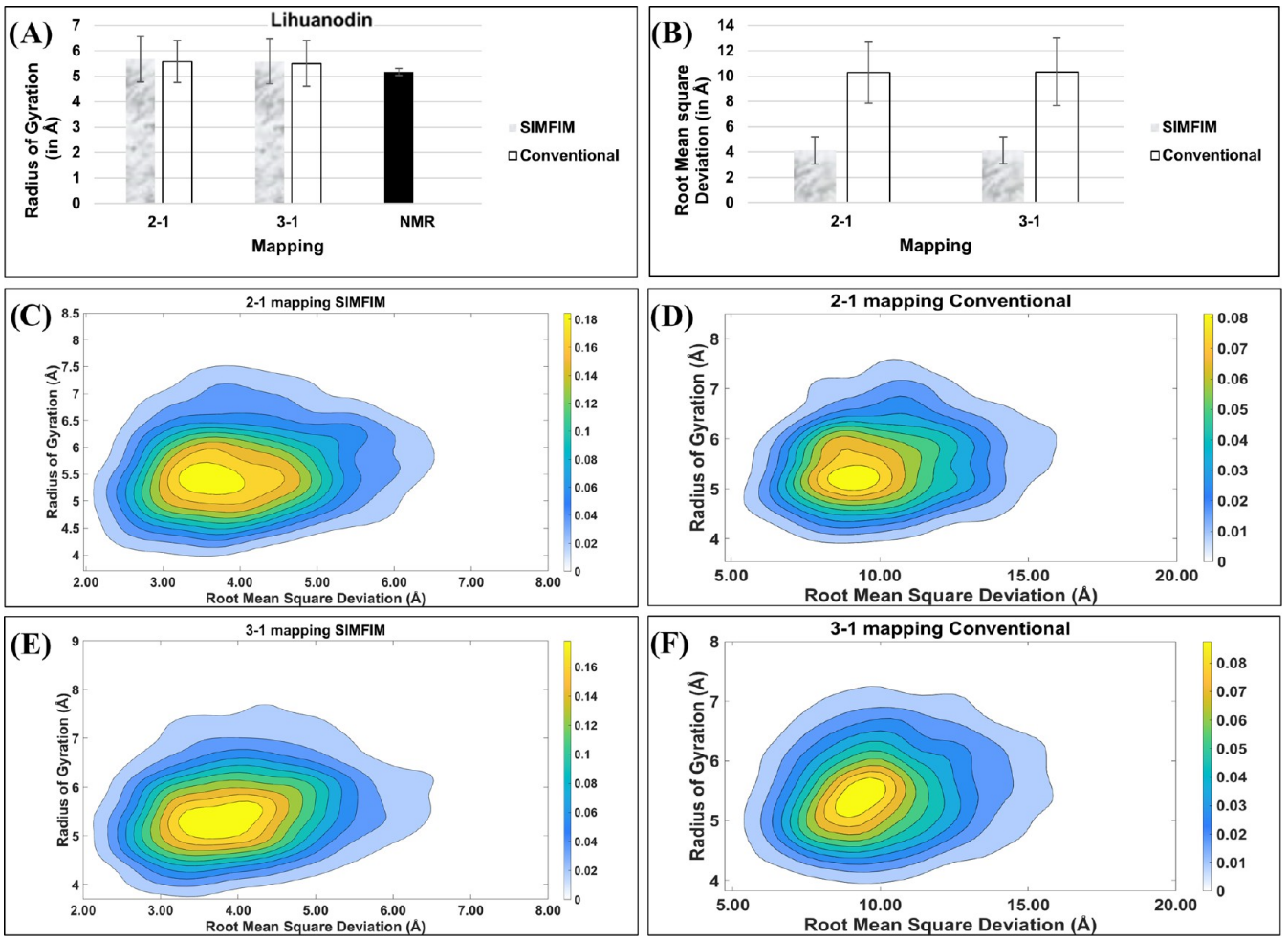


Figure 10. Comparison of (A) $\langle R_g \rangle$ and (B) ⟨RMSD⟩ values of the Lihuanodin peptide simulated with the conventional and SIMFIM models and ERFR functional form for electrostatic interaction for 2−1 and 3−1 mappings; the simulated $\langle R_g \rangle$ and ⟨RMSD⟩ values are compared with the actual value from the PDB database. The R_g versus RMSD contour plots simulated with (C) the conventional and (D) SIMFIM models for 2−1 mapping. The R_g versus RMSD contour plots simulated with (E) the conventional and (F) SIMFIM models for 3−1 mapping. The lowest (5.04 Å) and highest (5.32 Å) experimental R_g values among the 20 structures from the PDB database are shown as whiskers under the "NMR" column. The error bars show SD for R_g and RMSD.

mapping, the ranges were 3.78−4.70 Å (Figure 9E) and 9.29−12.27 Å (Figure 9F).

Figure 10A,B compares $\langle R_g \rangle$ and ⟨RMSD⟩ values of the Lihuanodin peptide, respectively, simulated with the conventional and SIMFIM models. Like Trpzip2 and Rubrivinodin peptides, the two models predicted $\langle R_g \rangle$ values for Lihuanodin close to the actual value of 5.17 Å for both 2−1 and 3−1 mappings (Figure 10A). For 2−1 mapping, ⟨RMSD⟩ simulated with SIMFIM was 4.21 Å, which was much lower than the value of 10.27 Å for the conventional model (Figure 10B), and for 3−1 mapping, ⟨RMSD⟩ for SIMFIM and the conventional models were 4.13 and 10.33 Å, respectively (Figure 10B). mpRMSD ranges for 2−1 mapping for SIMFIM and the conventional models were 3.29−4.06 Å (Figure 10C) and 8.24−9.99 Å (Figure 10D), respectively, whereas for 3−1 mapping, the ranges were 3.23−4.32 Å (Figure 10E) and 8.46−10.18 Å (Figure 10F).

Figure 11A,B compares $\langle R_g \rangle$ and ⟨RMSD⟩ values of the IC3-CB1/Gai peptide, respectively, simulated with the conventional and SIMFIM models. Like the other peptides, the two models predicted $\langle R_g \rangle$ values close to the actual value of 4.37 Å for 2−1 and 3−1 mappings (Figure 11A). For 2−1 mapping, ⟨RMSD⟩ simulated using SIMFIM was 7.79 Å, which was lower than the value of 11.58 Å for the conventional model (Figure 11B), and for 3−1 mapping, ⟨RMSD⟩ values for SIMFIM and the conventional models were 7.23 and 10.78 Å, respectively (Figure 11B). mpRMSD ranges for 2−1 mapping for SIMFIM and the conventional models were 6.05−7.35 Å (Figure 11C) and 8.20−10.01 Å (Figure 11D), respectively, whereas for 3−1 mapping, the ranges were 5.78−6.90 Å (Figure 11E) and 8.24−9.84 Å (Figure 11F).

Statistical analysis of the data in Figures 8−11 revealed that the R_g and RMSD values of the four peptides simulated with the SIMFIM model were significantly different from those simulated with the conventional model and they were closer to the actual R_g of the peptides. These simulation results demonstrate that the SIMFIM model predicted narrower distributions of RMSD and R_g values for the four peptides as compared to the conventional model. Further, ⟨RMSD⟩ values simulated with SIMFIM were much lower than those simulated with the conventional model, suggesting that the peptide structures predicted by SIMFIM were closer to the actual structures.

Effect of the Functional Form of Charge−Charge Interaction. The effect of the functional form (ERFR versus

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

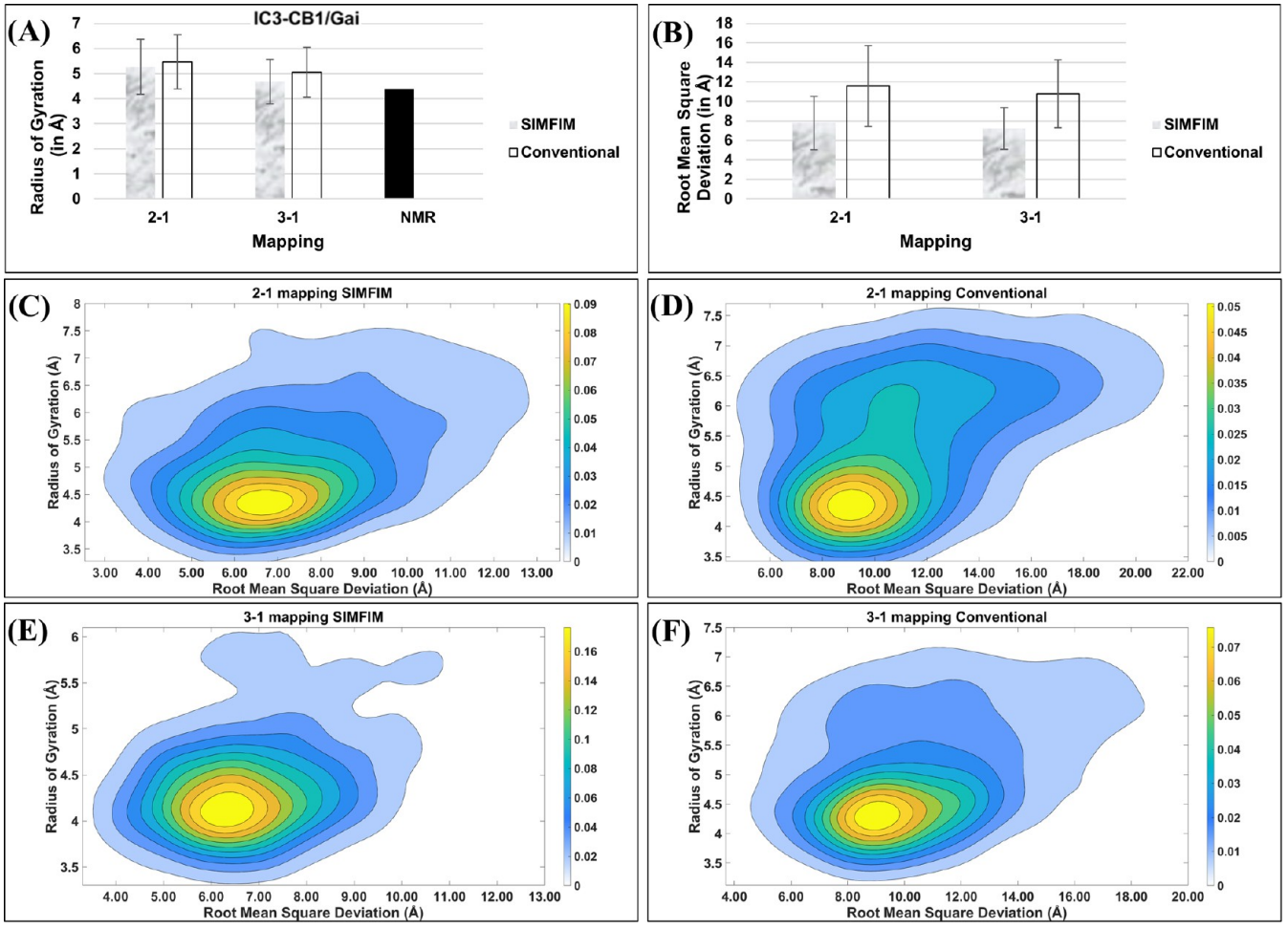


Figure 11. (A) Comparison of (A) ⟨R_g⟩ and (B) ⟨RMSD⟩ values of the IC3-CB1/Gai peptide simulated with the conventional and SIMFIM models and ERFR functional form for electrostatic interaction for 2−1 and 3−1 mappings; the simulated ⟨R_g⟩ and ⟨RMSD⟩ values are compared with the actual value from the PDB database. The R_g versus RMSD contour plots simulated with (C) the conventional and (D) SIMFIM models for 2−1 mapping. The R_g versus RMSD contour plots simulated with (E) the conventional and (F) SIMFIM models for 3−1 mapping. The error bars show SD for R_g and RMSD.

Coulombic) for electrostatic interaction for charged amino acids on ⟨RMSD⟩ and ⟨R_g⟩ values of Trpzip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai peptides simulated with SIMFIM model is shown in Figures 12−15, respectively, for 2−1 and 3−1 mappings. The peptides simulated without the electrostatic interaction, hereafter referred to as "None", were used as the reference group. For the TrpZip2 peptide (Figure 12) with 2−1 mapping, the ⟨R_g⟩ values simulated with ERFR, Coulombic, and None were 5.74, 9.25, and 7.98 Å, respectively, and for 3−1 mapping, the values were 6.02, 15.82, and 7.88 Å (Figure 12A). For 2−1 mapping, ⟨RMSD⟩ values simulated with ERFR, Coulombic, and None were 5.11, 17.95, and 7.98 Å, respectively, and for 3−1 mapping, the values were 5.28, 26.74, and 7.88 Å (Figure 12B). For 2−1 mapping, mpRg ranges for ERFR, Coulombic, and None were 4.85−5.67, 5.44−6.930, and 6.84−8.97 Å, respectively (Figure 12C,E,G), and for 3−1 mapping, the ranges were 5.25−5.80, 35.82−42.24, and 6.86−8.16 Å (Figure 12D,F,H). For 2−1 mapping, mpRMSD ranges for ERFR, Coulombic, and None were 4.12−5.14, 10.39−13.05, and 9.11−13.40 Å, respectively (Figure 12C,E,G), and for 3−1 mapping, the ranges were 4.38−5.38, 35.82−42.24, and 9.56−13.30 Å (Figure 12D,F,H).

For the Rubrivinodin peptide (Figure 13) with 2−1 mapping, ⟨R_g⟩ values simulated with ERFR, Coulombic, and None were 6.71, 12.53, and 10.14 Å, respectively, and for 3−1 mapping, the values were 6.51, 15.92, and 10.23 Å (Figure 13A). For 2−1 mapping, ⟨RMSD⟩ values for ERFR, Coulombic, and None were 4.78, 12.53, and 10.14 Å, respectively, and for 3−1 mapping, the values were 4.64, 26.27, and 10.23 Å (Figure 13B). For 2−1 mapping, mpRg ranges for ERFR, Coulombic, and None were 5.91−6.68, 6.73−8.02, and 8.51−10.44 Å, respectively (Figure 13C,E,G), and for 3−1 mapping, mpRg ranges were 5.75−6.82, 9.52−10.93, and 8.43−10.38 Å (Figure 13D,F,H). For 2−1 mapping, mpRMSD ranges for ERFR, Coulombic, and none were 3.95−4.90, 10.35−12.00, and 10.03−13.99 Å, respectively (Figure 13C,E,G), and for 3−1 mapping, mpRMSD ranges were 3.78−4.70, 18.34−21.30, and 10.26−13.47 Å (Figure 13D,F,H).

For the Lihuanodin peptide (Figure 14) with 2−1 mapping, ⟨R_g⟩ values simulated with ERFR, Coulombic, and None were 5.67, 13.48, and 9.13 Å, respectively, and for 3−1 mapping, the values were 5.59, 20.00, and 9.01 Å (Figure 14A). For 2−1 mapping, ⟨RMSD⟩ values simulated with ERFR, Coulombic, and None were 4.12, 21.96, and 12.85 Å, respectively, and for 3−1 mapping, ⟨RMSD⟩ values were 4.13, 24.61, and 12.53 Å (Figure 14B). For 2−1 mapping, mpRg ranges for ERFR, Coulombic, and None were 5.12−5.72, 6.66−8.18, and 7.29−9.84 Å, respectively (Figure 14C,E,G), and for 3−1 mapping, the

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

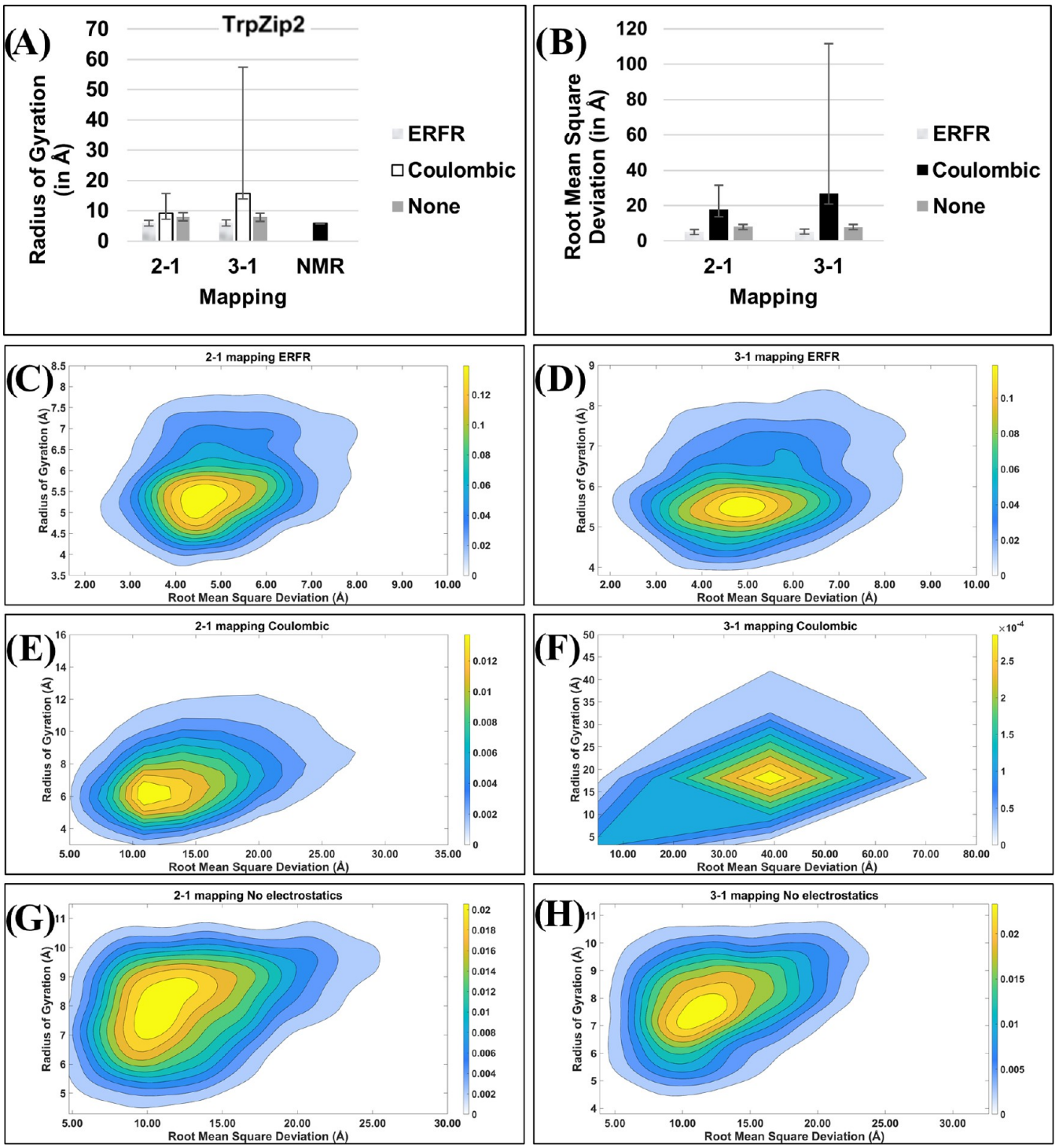


Figure 12. Comparison of (A) $\langle R_g \rangle$ and (B) ⟨RMSD⟩ values of the TrpZip2 peptide simulated with SIMFIM model and ERFR, Coulombic, or no electrostatic interaction, for 2−1 and 3−1 mappings; the simulated $\langle R_g \rangle$ and ⟨RMSD⟩ values are compared with the actual R_g from the PDB database. The R_g versus RMSD contour plots simulated with (C) ERFR, (E) Coulombic, and (G) None for 2−1 mapping. The R_g versus RMSD contour plots simulated with (D) ERFR, (F) Coulombic, and (H) None for 3−1 mapping. The lowest (5.69 Å) and highest (6.03 Å) experimental R_g values among the 20 structures from the PDB database are shown as whiskers under the "NMR" column. The error bars show SD for R_g and RMSD.

values were 4.95−5.71, 8.95−10.43, and 7.29−9.84 Å (Figure 14D,F,H). For 2−1 mapping, mpRMSD ranges for ERFR, Coulombic, and None were 3.29−4.06, 9.82−11.73, and 8.94−12.65 Å, respectively (Figure 14C,E,G), and for 3−1 mapping, the ranges were 3.23−4.32, 9.98−11.73, and 8.24−11.12 Å (Figure 14D,F,H).

For the IC3-CB1/Gai peptide (Figure 15) with 2−1 mapping, $\langle R_g \rangle$ values simulated with ERFR, Coulombic, and None were 7.79, 13.41, and 10.06 Å, respectively, and for 3−1 mapping, the values were 7.23, 17.85, and 9.85 Å (Figure 15A). For 2−1 mapping, ⟨RMSD⟩ values simulated with ERFR, Coulombic, and None were 7.79, 13.41, and 10.06 Å, respectively, and for 3−1 mapping, the values were 7.23, 17.85, and 9.85 Å (Figure 15B).

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

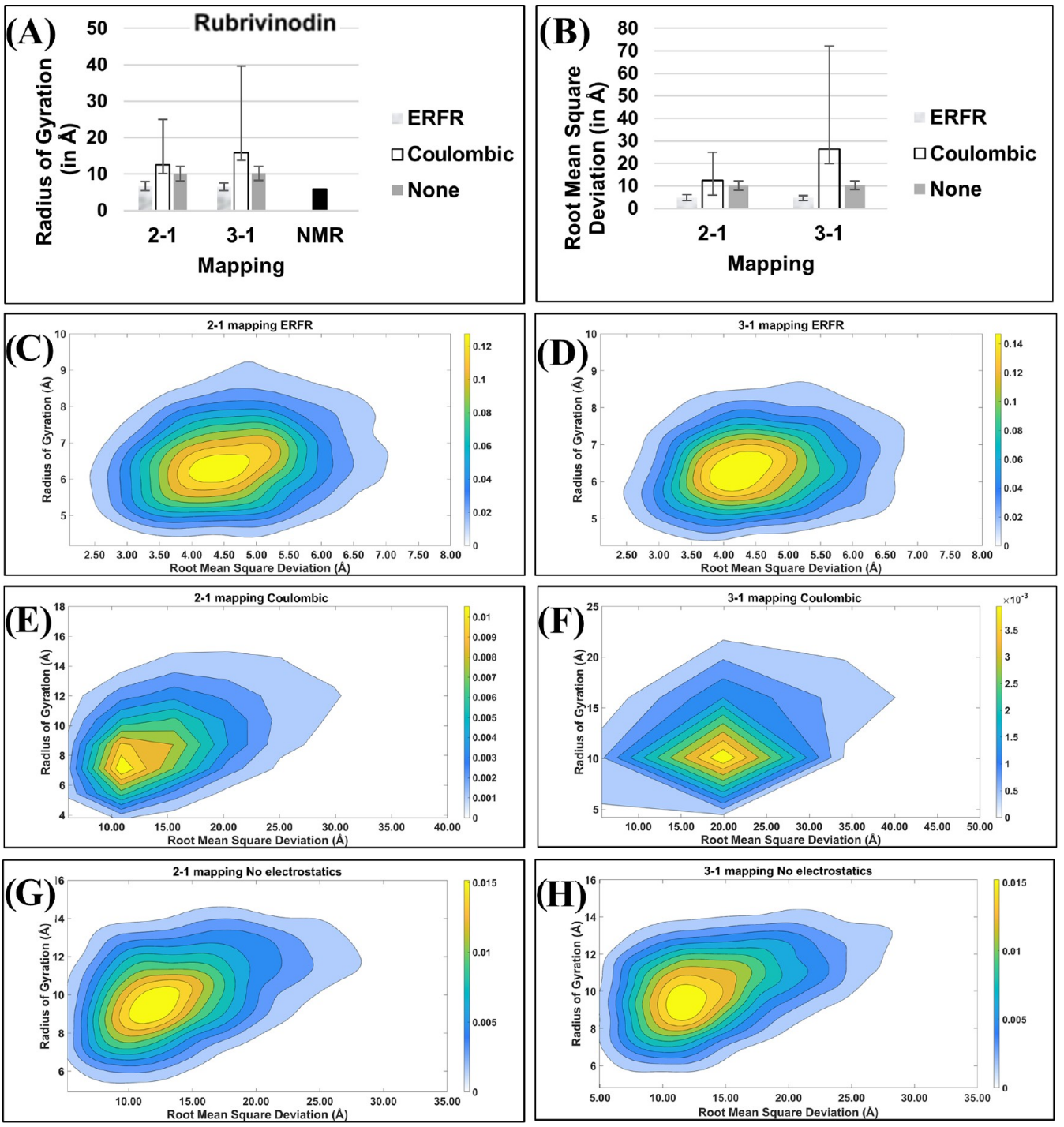


Figure 13. Comparison of (A) $\langle R_g \rangle$ and (B) ⟨RMSD⟩ values of the Rubrivinodin peptide simulated with the SIMFIM model and ERFR, Coulombic, or no electrostatic interaction, for 2−1 and 3−1 mappings; the simulated $\langle R_g \rangle$ and ⟨RMSD⟩ values are compared with the actual R_g from the PDB database. The R_g versus RMSD contour plots simulated with (C) ERFR, (E) Coulombic, and (G) None for 2−1 mapping. The R_g versus RMSD contour plots simulated with (D) ERFR, (F) Coulombic, and (H) None for 3−1 mapping. The error bars show SD for R_g and RMSD.

For 2−1 mapping, mpRg ranges with ERFR, Coulombic, and None were 4.12−4.58, 4.91−6.06, and 5.85−7.09 Å, respectively (Figure 15C,E,G), and with 3−1 mapping, the ranges were 3.94−4.32, 5.80−7.89, and 6.24−6.98 Å (Figure 15D,F,H). For 2−1 mapping, mpRMSD ranges for ERFR, Coulombic, and None were 6.05−7.35, 7.54−8.78, and 7.73−10.37 Å (Figure 15C,E,G), and for 3−1 mapping, the ranges were 5.78−6.90, 7.37−10.52, and 7.40−10.57 Å (Figure 15D,F,H).

As charged beads in the Coulombic functional form of electrostatic interaction are point charges, their overlap in the DPD simulation of soft beads resulted in the release of large amounts of energy and overstretching of the peptide chain beyond the fully stretched limit, which was physically unrealistic. As the $\langle R_g \rangle$ and ⟨RMSD⟩ values in the Coulombic form were skewed to very large positive values by the overstretching, the minus side of the standard deviations was calculated by

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

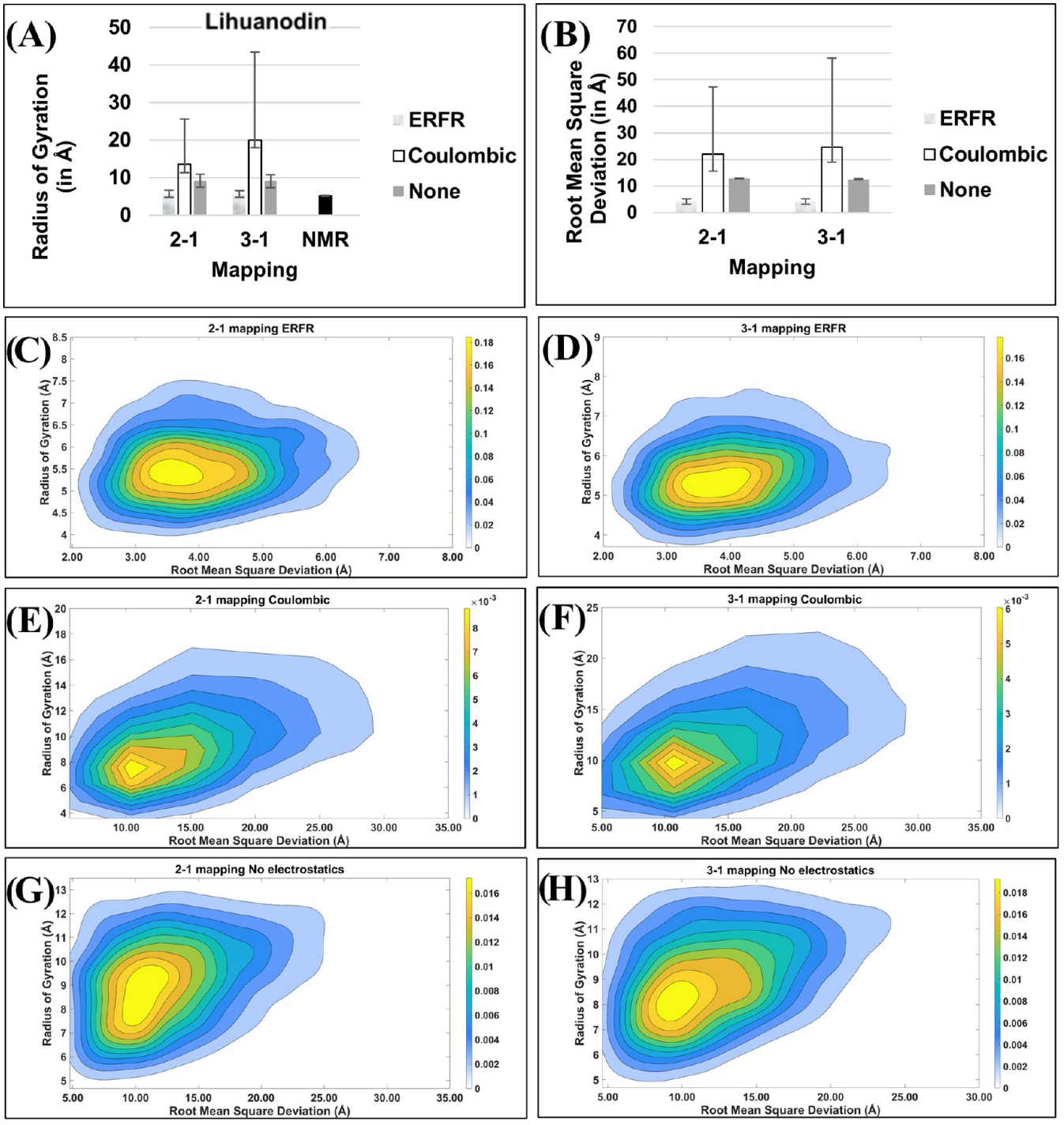


Figure 14. Comparison of (A) $\langle R_g \rangle$ and (B) ⟨RMSD⟩ values of the Lihuanodin peptide simulated with the SIMFIM model and ERFR, Coulombic, or no electrostatic interaction, for 2−1 and 3−1 mappings; the simulated $\langle R_g \rangle$ and ⟨RMSD⟩ values are compared with the actual R_g from the PDB database. The R_g versus RMSD contour plots simulated with (C) ERFR, (E) Coulombic, and (G) None for 2−1 mapping. The R_g versus RMSD contour plots simulated with (D) ERFR, (F) Coulombic, and (H) None for 3−1 mapping. The lowest (5.04 Å) and highest (5.32 Å) experimental R_g values among the 20 structures from the PDB database are shown as whiskers under the "NMR" column. The error bars show SD for R_g and RMSD.

removing the unphysically stretched data points to avoid negative values falling within the range of standard deviation.

Figure S5 in the Supporting Information File shows snapshots of the backbone of each peptide simulated using the Coulombic form at simulation times that the peptide experienced physically unrealistic overstretching. The R_g and RMSD values of each peptide were extremely large under overstretching (Figure S5). Figure 16A−D compares $\langle R_g \rangle$ values of TrpZip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai peptides, respectively, as a function of mapping simulated with ERFR, Coulombic, and no electrostatic interaction. For all peptides, $\langle R_g \rangle$ values simulated with ERFR form were closer to the actual $\langle R_g \rangle$ of the peptides as compared to Coulombic or None. For all peptides, mpRg values simulated with ERFR contained the actual $\langle R_g \rangle$ of the peptide, implying that the simulated peptides equilibrated closer to the actual structure with ERFR.

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

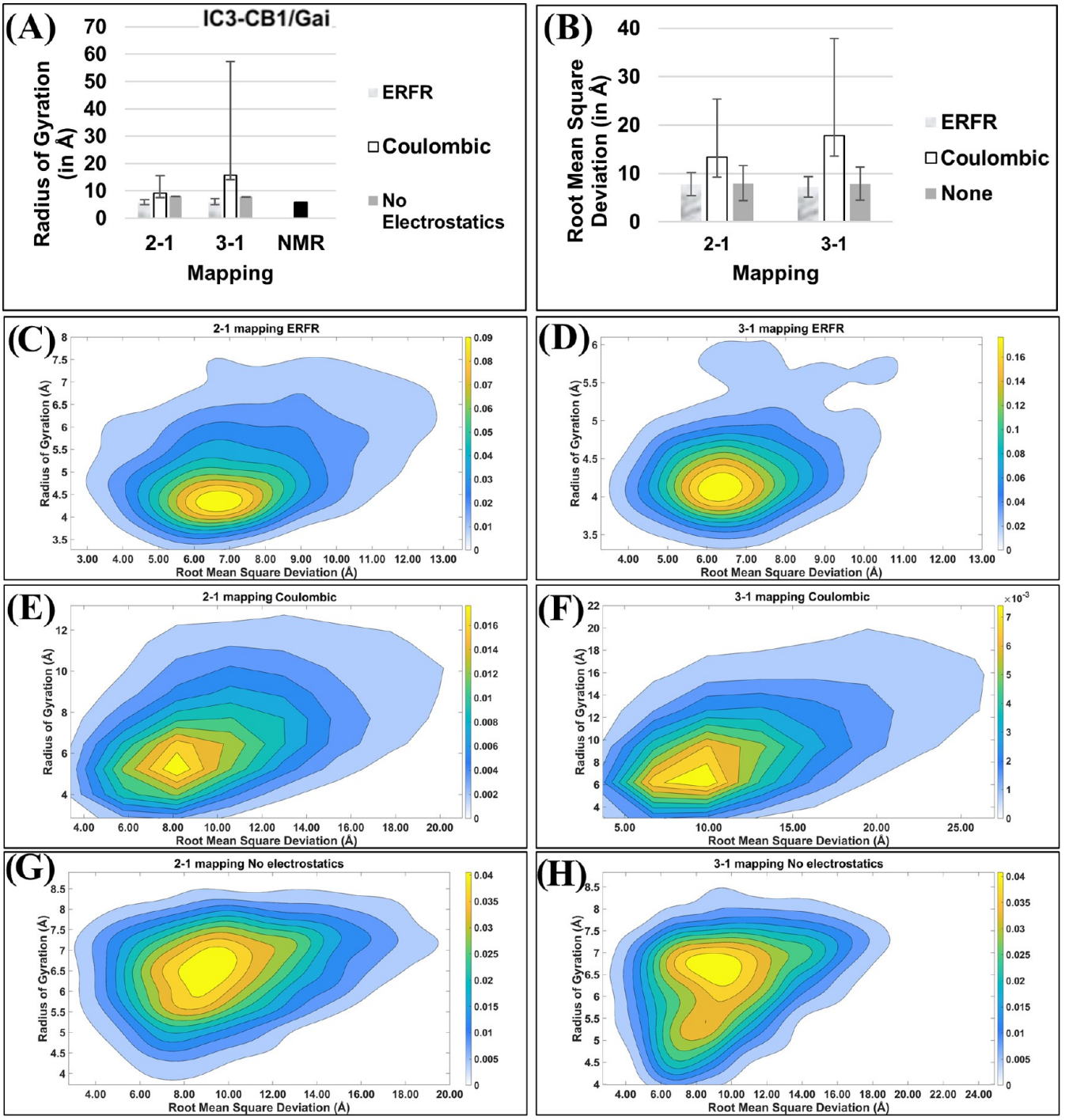


Figure 15. Comparison of (A) $\langle R_g \rangle$ and (B) ⟨RMSD⟩ values of the IC3-CB1/Gai peptide simulated with the SIMFIM model and ERFR, Coulombic, or no electrostatic interaction for 2−1 and 3−1 mappings; the simulated $\langle R_g \rangle$ and ⟨RMSD⟩ values are compared with the actual R_g from the PDB database. The R_g versus RMSD contour plots simulated with (C) ERFR, (E) Coulombic, and (G) None for 2−1 mapping. The R_g versus RMSD contour plots simulated with (D) ERFR, (F) Coulombic, and (H) None for 3−1 mapping. The error bars show SD for R_g and RMSD.

Figure 17 compares snapshot images of the simulated structures of the peptides' backbones with their actual structures from NMR spectroscopy. The ⟨RMSD⟩ and $\langle R_g \rangle$ values of the simulated structures in the snapshots are provided in the caption of Figure 17.

CONCLUSIONS

We studied the effect of mapping, model for determination of the Flory−Huggins interaction parameter between the beads, and the functional form of the electrostatic interaction on the DPD simulation of model peptides. The peptides with known structures were TrpZip2 (SWTWENGKWTWK), Rubrivinodin (GAPSLINSEDNPAFPQRD), Lihuanodin (GSKYSDTADESSYRW), and IC3-GB1/Gai (DIRLAKTLV) with 12, 18, 15, and 9 amino acids, respectively. The MARTINI model was used for coarse-graining groups of atoms or molecular fragments into beads with mapping ranging from 2−1 to 6−1. The interaction parameter matrix for pairwise bead interactions was determined

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

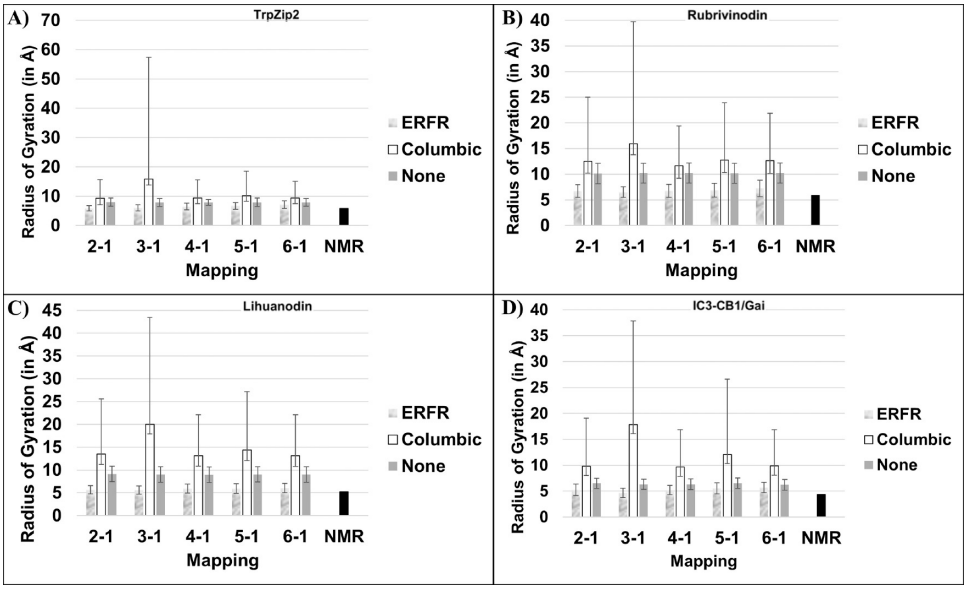


Figure 16. Comparison of $\langle R_g \rangle$ and ⟨RMSD⟩ values of (A) TrpZip2, (B) Rubrivinodin, (C) Lihuanodin, and (D) IC3-CB1/Gai peptides simulated with the SIMFIM model and ERFR, Coulombic, or no electrostatic interaction for 2−1 to 6−1 mappings; the simulated $\langle R_g \rangle$ values are compared with the actual R_g from the PDB database. For TrpZip2 and Lihuanodin, the lowest and highest experimental R_g values among the 20 structures from the PDB database are shown as whiskers under the "NMR" column. The error bars show SD for R_g.

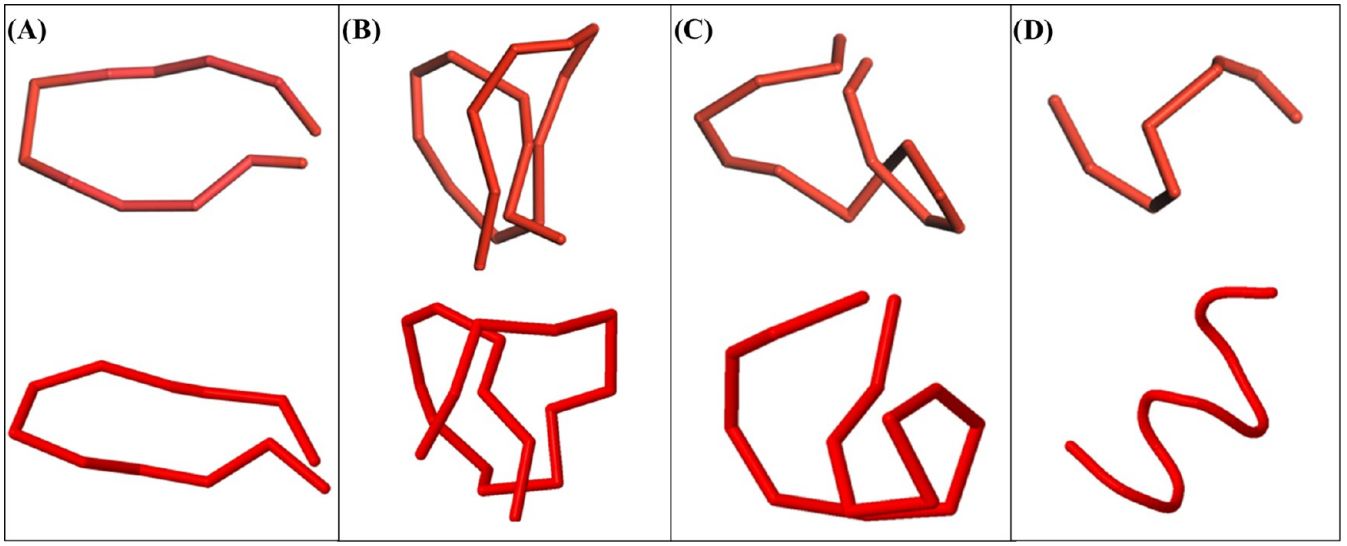


Figure 17. Comparisons of simulated (top) and actual (bottom) structures of (A) TrpZip2 (RMSD = 3.13 Å, R_g = 5.86 Å), (B) Rubrivinodin (RMSD = 3.51 Å, R_g = 5.79 Å), (C) Lihuanodin (RMSD = 4.9 Å, R_g = 5.13 Å), and (D) IC3-CB1/Gai (RMSD = 5.24 Å, R_g = 4.75 Å) peptides. The structures were simulated with the SIMFIM model, ERFR for electrostatic interaction, and mapping of 3−1.

using the conventional model, based on solubility parameter, and our new structure independent molecular fragment interfuse model (SIMFIM). The SIMFIM model accounts for specific interactions like hydrogen-bonding, polar−polar, and polar−nonpolar interactions, which are abundantly present in peptides and proteins. The effect of electrostatic interaction between amino acids with charged groups was modeled with ERFR and Coulombic functional forms of the electrostatic interaction. The simulated peptide structures were compared with the actual structures as measured by NMR spectroscopy. Lower mappings of 2−1 and 3−1 predicted structural properties closer to the actual properties as compared to higher mappings of 4−1 to 6−1, which was attributed to the higher atomic resolution of the simulated peptides at lower mappings. Further, the mass of individual beads in 2−1 and 3−1 mappings was 36 and 54 amu, respectively, which was closer to the average mass of all molecular fragments represented by the beads (45 amu), as compared to the mass of individual beads at higher mappings. The ⟨RMSD⟩ values, which represented the deviation of the spatial coordinates of the simulated beads from those of the actual structure, were lower at lower mappings of 2−1 and 3−1. The $\langle R_g \rangle$ values of the peptides were smaller at lower mappings and closer to the actual value as compared to higher mappings. The distribution of $\langle R_g \rangle$ values of the peptides was narrower at lower mappings. As the simulation time of the peptides with 2−1 mapping was relatively long, 3−1 mapping provided predictions sufficiently close to the actual properties at practical simulation times. The conventional model calculated repulsive Flory−

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

Huggins interaction parameters between water and all amino acids, which was unrealistic for some amino acids like lysine, glutamic acid, arginine, aspartic acid, serine, glutamine, and arginine. Conversely, our SIMFIM model predicted both attractive and repulsive interactions. Although the simulated $\langle R_g \rangle$ values with the conventional and SIMFIM models were close to the actual R_g of the peptides, the simulated ⟨RMSD⟩ values for SIMFIM were much lower than those of the conventional model. Therefore, the SIMFIM model predicted structures that were closer to the actual peptide structures. The $\langle R_g \rangle$ values of the peptides simulated using the ERFR functional form for electrostatic interaction were closer to the actual value as compared to the Coulombic function, which was attributed to the physically unrealistic overstretching of the chains and the release of excessively large quantities of energy. In the case of the IC3-GB1/Gai peptide, the ⟨RMSD⟩ values were slightly higher compared to the other peptides, which was attributed to the imposition of no constraints on the stability of the α-helical structure of IC3-GB1/Gai in the simulation. Unlike the other peptides, all amino acids in the IC3-CB1/Gai peptide participated in the secondary structure formation. Despite the lack of any constraints, the range of $\langle R_g \rangle$ values at maximum probability for the IC3-GB1/Gai peptide simulated with SIMFIM and ERFR functions for electrostatic interaction contained the actual value at 3−1 mapping.

The DPD simulation of peptides with the new SIMFIM model is potentially useful for predicting the physical properties of synthetic peptides with unknown structures. This methodology could potentially be used in the future to improve our ability to predict by simulation the formation of self-assembled peptide structures at the nanoscale in an aqueous medium. Further, it could potentially be used in the future to simulate the aggregation of peptides and predict their critical concentration for micelle formation.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c09534.

Molecular fragments in the amino acids that are represented by the coarse-grained beads (Table S1); Tukey's post hoc test results for the effect of mapping on R_g values of simulated TrpZip2 (Table S2); Tukey's post hoc test results for the effect of mapping on R_g values of simulated Rubrivinodin (Table S3); Tukey's post hoc test results for the effect of mapping on R_g values of simulated Lihuanodin (Table S4); Tukey's post hoc test results for the effect of mapping on R_g values of simulated IC3-CB1/Gai (Table S5); matrix interaction parameter values between beads i and j (α_{ij}) using SIMFIM model at 2−1 mapping (Table S6); matrix interaction parameter values between beads i and j (α_{ij}) using SIMFIM model at 3−1 mapping (Table S7); matrix interaction parameter values between beads i and j (α_{ij}) using SIMFIM model at 4−1 mapping (Table S8); matrix interaction parameter values between beads i and j (α_{ij}) using SIMFIM model at 5−1 mapping (Table S9); matrix interaction parameter values between beads i and j (α_{ij}) using SIMFIM model at 6−1 mapping (Table S10); matrix interaction parameter values between beads i and j (α_{ij}) using the conventional model at 2−1 mapping (Table S11); matrix interaction parameter values between beads i and j (α_{ij}) using the conventional model at 3−1 mapping (Table S12); matrix interaction parameter values between beads i and j (α_{ij}) using the conventional model at 4−1 mapping (Table S13); matrix interaction parameter values between beads i and j (α_{ij}) using the conventional model at 5−1 mapping (Table S14); matrix interaction parameter values between beads i and j (α_{ij}) using the conventional model at 6−1 mapping (Table S15); steps in determining the RMSD of a simulated peptide (Figure S1); 1D normalized distribution of end-to-end vector for TrpZip2 peptide in the x-direction obtained from the simulation (blue) and the fit of the data by MATLAB to the normal distribution function (red) (Figure S2); 1D normalized distribution of end-to-end vector for TrpZip2 peptide in the y-direction obtained from the simulation (blue) and the fit of the data by MATLAB to the normal distribution function (red) (Figure S3); 1D normalized distribution of end-to-end vector for TrpZip2 peptide in the z-direction obtained from the simulation (blue) and the fit of the data by MATLAB to the normal distribution function (red) (Figure S4); and snapshots of the backbone of each peptide (TrpZip2, Rubrivinodin, Lihuanodin, and IC3-CB1/Gai) simulated using the Coulombic form of electrostatic interaction at simulation times that the peptide experienced physically unrealistic overstretching (Figure S5) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Esmaiel Jabbari − *Biomimetic Materials and Tissue Engineering Laboratory, Chemical Engineering Department, University of South Carolina, Columbia, South Carolina 29208, United States;* orcid.org/0000-0001-6548-5422; Phone: (803) 777-8022; Email: jabbari@cec.sc.edu; Fax: (803) 777-0973

Author

Ricky Anshuman Dash − *Biomimetic Materials and Tissue Engineering Laboratory, Chemical Engineering Department, University of South Carolina, Columbia, South Carolina 29208, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c09534

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. Seyedsina Moeinzadeh for consultation in using the Materials Studio software. R.A.D. was supported by a teaching assistantship from the Chemical Engineering Department and the department supported the license for Materials Studio software. The authors thank Dr. Ting Ge in the Department of Chemistry and Biochemistry (U. South Carolina) for consultation regarding determination of the distribution of end-to-end distance of the peptide chains.

■ REFERENCES

(1) Wang, L.; Wang, N. X.; Zhang, W. P.; Cheng, X. R.; Yan, Z. B.; Shao, G.; Wang, X.; Wang, R.; Fu, C. Y. Therapeutic peptides: current applications and future directions. *Signal Transduction Targeted Ther.* **2022**, *7* (1), No. 48, DOI: 10.1038/s41392-022-00904-4.

(2) Fosgerau, K.; Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discovery Today* **2015**, *20* (1), 122−128.

(3) Koutsopoulos, S. Self-assembling Peptides in Biomedicine and Bioengineering: Tissue Engineering, Regenerative Medicine, Drug Delivery, and Biotechnology. In *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*; Koutsopoulos, S., Ed.; Elsevier, 2018; pp 387−408.
(4) Huan, Y.; Kong, Q.; Mou, H.; Yi, H. Antimicrobial peptides: classification, design, application and research progress in multiple fields. *Front. Microbiol.* **2020**, *11*, No. 582779.
(5) Del Genio, V.; Bellavita, R.; Falanga, A.; Hervé-Aubert, K.; Chourpa, I.; Galdiero, S. Peptides to overcome the limitations of current anticancer and antimicrobial nanotherapies. *Pharmaceutics* **2022**, *14* (6), No. 1235, DOI: 10.3390/pharmaceutics14061235.
(6) Gori, A.; Lodigiani, G.; Colombarolli, S. G.; Bergamaschi, G.; Vitali, A. Cell penetrating peptides: classification, mechanisms, methods of study, and applications. *ChemMedChem* **2023**, *18* (17), No. e202300236.
(7) Malonis, R. J.; Lai, J. R.; Vergnolle, O. Peptide-based vaccines: current progress and future challenges. *Chem. Rev.* **2020**, *120* (6), 3210−3229.
(8) Cao, Y.; Zhou, L.; Fang, Z. K.; Zou, Z. H.; Zhao, J.; Zuo, X. L.; Li, G. X. Application of functional peptides in the electrochemical and optical biosensing of cancer biomarkers. *Chem. Commun.* **2023**, *59* (23), 3383−3398.
(9) Levin, A.; Hakala, T. A.; Schnaider, L.; Bernardes, G. J. L.; Gazit, E.; Knowles, T. P. J. Biomimetic peptide self-assembly for functional materials. *Nat. Rev. Chem.* **2020**, *4* (11), 615−634.
(10) Sorolla, A.; Wang, E.; Golden, E.; Duffy, C.; Henriques, S. T.; Redfern, A. D.; Blancafort, P. Precision medicine by designer interference peptides: applications in oncology and molecular therapeutics. *Oncogene* **2020**, *39* (6), 1167−1184.
(11) Alshehri, S.; Susapto, H. H.; Hauser, C. A. E. Scaffolds from self-assembling tetrapeptides support 3d spreading, osteogenic differentiation, and angiogenesis of mesenchymal stem cells. *Biomacromolecules* **2021**, *22* (5), 2094−2106.
(12) Seon, J. K.; Kuppa, S. S.; Kang, J. Y.; Lee, J. S.; Park, S. A.; Yoon, T. R.; Park, K. S.; Kim, H. K. Peptide derived from stromal cell-derived factor 1δ enhances the expression of osteogenic proteins bone marrow stromal cell differentiation and promotes bone formation in models. *Biomater. Sci.* **2023**, *11* (19), 6587−6599.
(13) Khavinson, V.; Linkova, N.; Diatlova, A.; Trofimova, S. Peptide regulation of cell differentiation. *Stem Cell Rev. Rep.* **2020**, *16* (1), 118−125.
(14) Linkova, N.; Khavinson, V.; Diatlova, A.; Myakisheva, S.; Ryzhak, G. Peptide regulation of chondrogenic stem cell differentiation. *Int. J. Mol. Sci.* **2023**, *24* (9), No. 8415, DOI: 10.3390/ijms24098415.
(15) Kader, S.; Monavarian, M.; Barati, D.; Moeinzadeh, S.; Makris, T. M.; Jabbari, E. Plasmin-cleavable nanoparticles for on-demand release of morphogens in vascularized osteogenesis. *Biomacromolecules* **2019**, *20* (8), 2973−2988.
(16) Tong, Z.; Guo, J. X.; Glen, R. C.; Morrell, N. W.; Li, W. A bone morphogenetic protein (BMP)-derived peptide based on the type I receptor-binding site modifies cell-type dependent BMP signalling. *Sci. Rep.* **2019**, *9*, No. 13446.
(17) Moeinzadeh, S.; Barati, D.; Sarvestani, S. K.; Karimi, T.; Jabbari, E. Experimental and computational investigation of the effect of hydrophobicity on aggregation and osteoinductive potential of BMP-2-derived peptide in a hydrogel matrix. *Tissue Eng., Part A* **2015**, *21* (1−2), 134−146.
(18) Basith, S.; Manavalan, B.; Hwan, T. H.; Lee, G. Machine intelligence in peptide therapeutics: a next-generation tool for rapid disease screening. *Med. Res. Rev.* **2020**, *40* (4), 1276−1314.
(19) Zhao, G. Z.; Zhang, L. Q.; Che, L. F.; Li, H. Z.; Liu, Y.; Fang, J. Revisiting bone morphogenetic protein-2 knuckle epitope and redesigning the epitope-derived peptides. *J. Pept. Sci.* **2021**, *27* (6), No. e3309.
(20) Liu, Q.; Peng, X. L.; Liu, X. J.; Mou, X. H.; Guo, Y. Y.; Yang, L. H.; Chen, Y. F.; Zhou, Y. Y.; Shi, Z. Y.; Yang, Z. L.; Chen, Z. Advances in the application of bone morphogenetic proteins and their derived peptides in bone defect repair. *Composites, Part B* **2023**, *262*, No. 110805.
(21) Ajeeb, B.; Acar, H.; Detamore, M. S. Chondroinductive peptides for cartilage regeneration. *Tissue Eng., Part B* **2022**, *28* (4), 745−765.
(22) Guigas, G.; Morozova, D.; Weiss, M. Exploring Membrane and Protein Dynamics with Dissipative Particle Dynamics. In *Advances in Protein Chemistry and Structural Biology*; Elsevier, 2011; Vol. *85*, pp 143−182.
(23) Vendruscolo, M.; Dobson, C. M. Protein dynamics: Moore's law in molecular biology. *Curr. Biol.* **2011**, *21* (2), R68−R70, DOI: 10.1016/j.cub.2010.11.062.
(24) Jabbari, E.; Yang, X.; Moeinzadeh, S.; He, X. Drug release kinetics, cell uptake, and tumor toxicity of hybrid VVVVVVKK peptide-assembled polylactide nanoparticles. *Eur. J. Pharm. Biopharm.* **2013**, *84* (1), 49−62.
(25) Vishnyakov, A.; Talaga, D. S.; Neimark, A. V. DPD simulation of protein conformations: from α-helices to β-structures. *J. Phys. Chem. Lett.* **2012**, *3*, 3081−3087.
(26) Banerjee, A.; Lu, C. Y.; Dutt, M. A hybrid coarse-grained model for structure, solvation and assembly of lipid-like peptides. *Phys. Chem. Chem. Phys.* **2022**, *24* (3), 1553−1568.
(27) Li, C.; Fu, X.; Zhong, W.; Liu, J. Dissipative particle dynamics simulations of a protein-directed self-assembly of nanoparticles. *ACS Omega* **2019**, *4* (6), 10216−10224.
(28) Chen, R.-Y.; Wu, L.-Y.; Liao, J.-M.; Chen, C.-L. Dissipative particle dynamics simulations to investigate aggregation of peptide amphiphile nanofibers. *J. Chin. Chem. Soc.* **2007**, *54* (4), 861−868.
(29) Peter, E. K.; Lykov, K.; Pivkin, I. V. A polarizable coarse-grained protein model for dissipative particle dynamics. *Phys. Chem. Chem. Phys.* **2015**, *17* (37), 24452−24461.
(30) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* **2008**, *4* (5), 819−834.
(31) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theory Comput.* **2013**, *9* (1), 687−697.
(32) Basdevant, N.; Borgis, D.; Ha-Duong, T. Modeling protein-protein recognition in solution using the coarse-grained force field SCORPION. *J. Chem. Theory Comput.* **2013**, *9* (1), 803−813.
(33) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The power of coarse graining in biomolecular simulations. *WIREs Comput. Mol. Sci.* **2014**, *4* (3), 225−248.
(34) Pieczywek, P. M.; Płaziński, W.; Zdunek, A. Dissipative particle dynamics model of homogalacturonan based on molecular dynamics simulations. *Sci. Rep.* **2020**, *10* (1), No. 14691.
(35) Groot, R. D.; Warren, P. B. Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* **1997**, *107* (11), 4423−4435.
(36) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-grained protein models and their applications. *Chem. Rev.* **2016**, *116* (14), 7898−7936.
(37) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812−7824.
(38) Poma, A. B.; Cieplak, M.; Theodorakis, P. E. Combining the MARTINI and structure-based coarse-grained approaches for the molecular dynamics studies of conformational transitions in proteins. *J. Chem. Theory Comput.* **2017**, *13* (3), 1366−1374.
(39) Vaiwala, R.; Ayappa, K. G. A generic force field for simulating native protein structures using dissipative particle dynamics. *Soft Matter* **2021**, *17* (42), 9772−9785.
(40) Choudhury, C. K.; Kuksenok, O. Native-based dissipative particle dynamics approach for α-Helical folding. *J. Phys Chem. B* **2020**, *124* (50), 11379−11386, DOI: 10.1021/acs.jpcb.0c08603.
(41) Periole, X.; Cavalli, M.; Marrink, S. J.; Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **2009**, *5* (9), 2531−2543.

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022

(42) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589, DOI: 10.1038/s41586-021-03819-2.
(43) Xiao, Y.; Chen, C. J.; He, Y. Folding mechanism of beta-hairpin Trpzip2: heterogeneity, transition state and folding pathways. *Int. J. Mol. Sci.* **2009**, *10* (6), 2838−2848.
(44) Xiu, H. H.; Wang, M. J.; Fage, C. D.; He, Y. L.; Niu, X. G.; Han, M.; Li, F.; An, X. P.; Fan, H. H.; Song, L. H.; et al. Discovery and characterization of Rubrinodin provide clues into the evolution of lasso peptides. *Biochemistry* **2022**, *61* (7), 595−607.
(45) Cao, L.; Beiser, M.; Koos, J. D.; Orlova, M.; Elashal, H. E.; Schröder, H. V.; Link, A. J. Cellulonodin-2 and Lihuanodin: lasso peptides with an aspartimide post-translational modification. *J. Am. Chem. Soc.* **2021**, *143* (30), 11690−11702.
(46) Shim, J. Y.; Ahn, K. H.; Kendall, D. A. Molecular basis of cannabinoid CB1 receptor coupling to the G protein heterotrimer G$\alpha i\beta\gamma$. *J. Biol. Chem.* **2013**, *288* (45), 32449−32465.
(47) Zhu, T.; He, X.; Zhang, J. Z. H. Fragment density functional theory calculation of NMR chemical shifts for proteins with implicit solvation. *Phys. Chem. Chem. Phys.* **2012**, *14* (21), 7837−7845.
(48) Santo, K. P.; Neimark, A. V. Dissipative particle dynamics simulations in colloid and Interface science: a review. *Adv. Colloid Interface Sci.* **2021**, *298*, No. 102545.
(49) Jarray, A.; Gerbaud, V.; Hemati, M. Structure of aqueous colloidal formulations used in coating and agglomeration processes: mesoscale model and experiments. *Powder Technol.* **2016**, *291*, 244−261.
(50) Travis, K. P.; Bankhead, M.; Good, K.; Owens, S. L. New parametrization method for dissipative particle dynamics. *J. Chem. Phys.* **2007**, *127* (1), No. 014109.
(51) Español, P.; Warren, P. Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.* **1995**, *30* (4), No. 191, DOI: 10.1209/0295-5075/30/4/001.
(52) Moeinzadeh, S.; Jabbari, E. Mesoscale simulation of the effect of a lactide segment on the nanostructure of star poly(ethylene glycol-*co*-lactide)-acrylate macromonomers in aqueous solution. *J. Phys. Chem. B* **2012**, *116* (5), 1536−1543.
(53) Hu, C.; Lu, T.; Guo, H. Mesoscale modeling of sulfonated polyimides copolymer membranes: effect of sequence distributions. *J. Membr. Sci.* **2018**, *564*, 146−158.
(54) Wang, C.; Paddison, S. J. Mesoscale modeling of hydrated morphologies of sulfonated polysulfone ionomers. *Soft Matter* **2014**, *10* (6), 819−830.
(55) Bawendi, M. G.; Freed, K. F.; Mohanty, U. A lattice model for self-avoiding polymers with controlled length distributions. II. Corrections to Flory−Huggins mean field. *J. Chem. Phys.* **1986**, *84*, 7036−7047.
(56) Truszkowski, A.; van den Broek, K.; Kuhn, H.; Zielesny, A.; Matthias Epple, M. Mesoscopic simulation of phospholipid membranes, peptides, and proteins with molecular fragment dynamics. *J. Chem. Inf. Model.* **2015**, *55* (5), 983−997.
(57) Warren, P. B.; Vlasov, A. Screening properties of four mesoscale smoothed charge models with application to dissipative particle dynamics. *J. Chem. Phys.* **2014**, *140* (8), No. 084904.
(58) Andress, J.; Linn, R.Introduction to Perl. In *Coding for Penetration Testers*; Elsevier, 2017; pp 81−110.
(59) Cheng, F.; Guan, X.; Cao, H.; Su, T.; Cao, J.; Chen, Y.; Cai, M.; He, B.; Gu, Z.; Luo, X. Characteristic of core materials in polymeric micelles effect on their micellar properties studied by experimental and DPD simulation methods. *Int. J. Pharm.* **2015**, *492* (1−2), 152−160.
(60) Sun, M.; Li, B.; Li, Y.; Liu, Y.; Liu, Q.; Jiang, H.; He, Z.; Zhao, Y.; Sun, J. Experimental observations and dissipative particle dynamic simulations on microstructures of pH-sensitive polymer containing amorphous solid dispersions. *Int. J. Pharm.* **2017**, *517* (1−2), 185−195.
(61) Ley, K.; Christofferson, A.; Penna, M.; Winkler, D.; Maclaughlin, S.; Yarovsky, I. Surface-water interface induces conformational changes critical for protein adsorption: implications for monolayer formation of EAS hydrophobin. *Front. Mol. Biosci.* **2015**, *2*, No. 64, DOI: 10.3389/fmolb.2015.00064.
(62) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; OUP Oxford, 2017.
(63) Fayaz-Torshizi, M.; Müller, E. A. Coarse-grained molecular simulation of polymers supported by the use of the SAFT-gamma Mie equation of state. *Macromol. Theory Simul.* **2021**, *31* (1), No. 2100031, DOI: 10.1002/mats.202100031.
(64) Jia, Y.-B. *Rotation in the Space*; Iowa State University, 2022.
(65) Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press, 2003.

https://doi.org/10.1021/acsomega.3c09534
ACS Omega 2024, 9, 18001−18022