ACS OMEGA

Open Access

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

Article

Computational Discovery of Aggregation-Mediated Dipeptide Inhibitors Targeting PHLDA1 for Cardiovascular Therapy

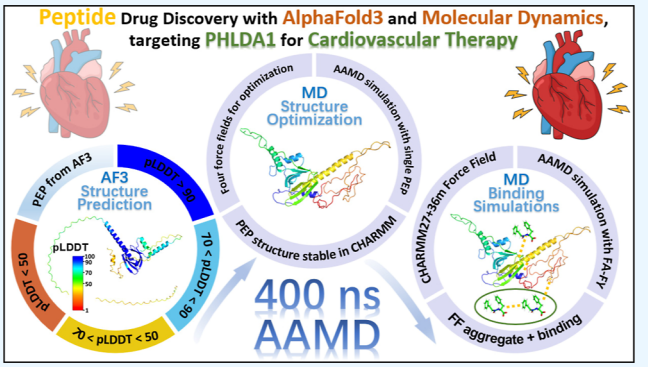
Shujia Liu,[¶] Haojin Zhou,[¶] Weihua Bian, Kanghui Huan, Meng Zhao,* and Jiaqi Wang*

Cite This: *ACS Omega* 2025, 10, 50088−50102

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Cardiovascular disease remains a leading cause of global mortality, underscoring the urgent need for novel therapeutic strategies. This study focuses on the pleckstrin homology-like domain family A member 1-encoded protein (PEP), a regulator of cardiomyocyte apoptosis and a promising yet underexplored therapeutic target. Leveraging an integrated computational approach, we combined AlphaFold3-based structure prediction with extensive molecular dynamics simulations to characterize PEP's dynamic architecture and identify high-affinity dipeptide inhibitors. Our results reveal that PEP possesses significant intrinsic disorder, with molecular dynamics simulations refining its conformation and highlighting force field-dependent behaviors, particularly the superior performance of CHARMM-class force fields in preserving functional secondary structures. High-throughput screening of phenylalanine-based dipeptides identified FF as the strongest binder, exhibiting a unique aggregation-mediated binding mechanism that engages both primary and secondary sites on PEP through multivalent interactions. Concentration-dependent simulations further confirmed the robustness of FF binding and revealed residue-specific interaction hotspots. Notably, we demonstrate that low-confidence regions in AF3 predictions (pLDDT <50) frequently participate in functional binding, challenging the conventional lock-and-key paradigm. These findings not only establish PEP as a tractable drug target but also provide a novel framework for designing aggregation-prone peptide therapeutics against cardiovascular diseases.



1. INTRODUCTION

Cardiovascular disease (CVD) remains the leading cause of death worldwide, accounting for approximately 17.9 million deaths annually.[1] Myocardial ischemia, myocardial infarction,[2] hypertrophy, and heart failure[3] are major contributors to CVD-related mortality, highlighting the critical need for timely therapeutic interventions to preserve cardiac function. Cardiomyocytes are the basic functional units of the heart, responsible for contraction and pumping functions, which are essential for maintaining systemic circulation.[4] Cardiomyocytes are also capable of generating and transmitting electrical impulses, which are the triggering signals for heart contraction and relaxation.[5] They are tightly interconnected by intercalated discs, which contain desmosomes, adherens junctions, and gap junctions, ensuring the heart contracts as a functional syncytium to pump blood effectively.[6] Damage or dysfunction of cardiomyocytes impairs the heart's pumping capacity, often leading to heart failure and life-threatening complications.[5] Consequently, research targeting cardiomyocyte-specific proteins with pharmacological potential is crucial for the development of novel therapeutic strategies.

Peptides are one such promising therapeutic targeting cardiomyocyte-specific proteins. Composed of 2−50 amino acid residues linked by peptide bonds (−CO−NH−), they are particularly effective in treating complex diseases due to their potential for high specificity, which minimizes off-target effects.[7] Compared to conventional small-molecule drugs, peptide therapeutics exhibit superior selectivity, rapid systemic clearance, and a reduced risk of toxicity or accumulation.[8] Their advantages also include low immunogenicity (unlike larger biologics), lower required dosages, and higher potency per unit mass.[9] These properties make them a compelling alternative to current mainstays of cardiovascular pharmacotherapy, such as beta-blockers, ACE inhibitors, statins, and antiplatelet agents, which are limited by systemic off-target effects.[1,2,6] Peptide-based therapeutics offer the potential to overcome these shortcomings by enabling high-affinity, cell-specific targeting of previously undruggable targets, allowing for precise inhibition of key apoptotic drivers in cardiac injury.[7,10]

Received: July 7, 2025
Revised: September 18, 2025
Accepted: September 25, 2025
Published: October 17, 2025

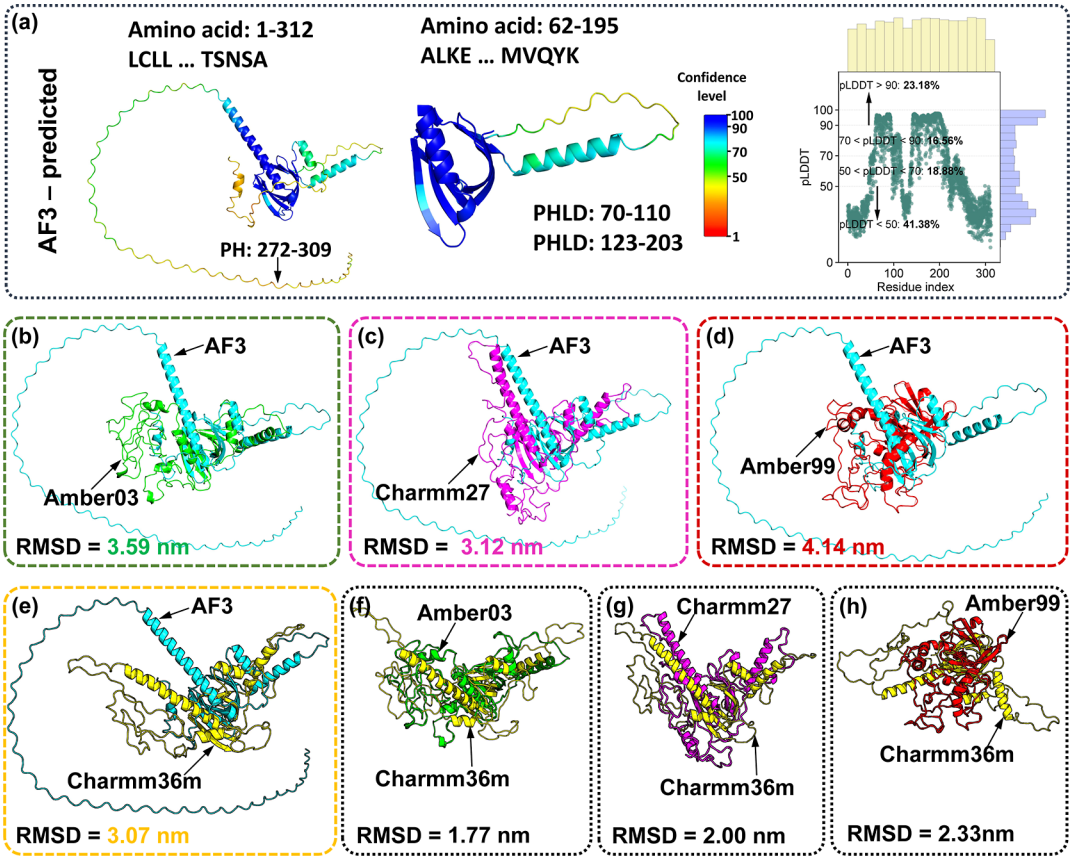


Figure 1. Structural comparison between AF3-predicted and MD-simulated conformations of PEP. (a) AF3-predicted PEP structure colored by pLDDT confidence scores, showing the full structure (left), the isolated PHLD region (middle), and the per-residue pLDDT distribution (right; see also Table S1). (b–e) Structural alignments between the initial AF3 prediction (cyan) and MD-relaxed conformations (400 ns, *NPT* ensemble) simulated using (b) Amber03 (green), (c) CHARMM27 (purple), (d) Amber99 (red), and (e) CHARMM36m (yellow). (f–h) Pairwise comparisons of CHARMM36m with the other three force fields. The RMSD values shown in (b–h) between two structures were computed using PyMOL to quantify structural differences after alignment, with "cutoff" set to 6 Å using 5 outlier rejection cycles. "cutoff" defines the maximum allowed distance in Ångströms between equivalent atoms to be considered a "match" during alignment.

The pleckstrin homology-like domain family A member 1 encoded-protein (PHLDA1-encoded protein, abbreviated as PEP) is a regulatory protein involved in cell growth, survival, and apoptosis and has been associated with various diseases such as obesity, diabetes, cancer, and so forth.[11] In cardiomyocytes, the expression and functional roles of PEP have received significant research interest. Studies have demonstrated that PEP overexpression in H9c2 cardiomyocytes leads to decreased cell viability.[12] Mechanistically, PEP modulates the stability of Bax, a key mediator of the mitochondrial apoptosis pathway, through protease-dependent regulation,[13] and PHLDA1 knockout has been shown to attenuate Bax upregulation during myocardial ischemia-reperfusion injury.[10] Other than that, one of the most well-documented interaction partners of PHLDA1 is Akt, which it inhibits by competitively binding phosphatidylinositol (3,4,5)-trisphosphate (PIP3) via its pleckstrin homology (PH) domain, thereby suppressing the pro-survival PI3K/Akt pathway.[14] In the nucleus, PHLDA1 functions as a transcriptional coactivator for CREB through a direct interaction mediated by its C-terminal domain, enhancing the expression of target genes.[10,11] Targeted screening of short peptide ligands (e.g., dipeptides) capable of binding to PEP may yield novel therapeutic candidates[15] inhibiting the aforementioned protein–protein interactions, offering a viable strategy to mitigate myocardial ischemia and related cardiovascular conditions, as well as other diseases such as obesity and cancer.

Notably, while PEP has been implicated in cardiomyocyte apoptosis,[12] due to its intrinsically disordered protein (IDP) nature,[16] no experimental structures are currently available. Consequently, no prior studies have systematically explored its potential as a drug target. In this work, we adopted an in silico approach, which combines AlphaFold3 (AF3) structure prediction and molecular dynamics (MD) simulations with biophysical analyses for efficient ligand discovery. This work represents the first comprehensive effort to (1) obtain an accurate three-dimensional structure of PEP, (2) map PEP's druggable binding sites, and (3) identify optimal dipeptide ligands. MD simulation is a powerful computational methodology rooted in Newtonian mechanics[17] and is widely employed in chemistry, physics, biology, materials science, and pharmaceutical research.[18] Its capacity to efficiently explore molecular trajectories, interactions, and structural dynamics offers a time- and cost-effective alternative to experimental approaches (i.e., cryo-EM), while providing atomic-level insights that complement empirical studies.[19] Our strategy combines computational efficiency with mechanistic precision, offering a foundation for developing novel peptide-based therapies against PEP-driven cardiovascular pathologies.

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088–50102

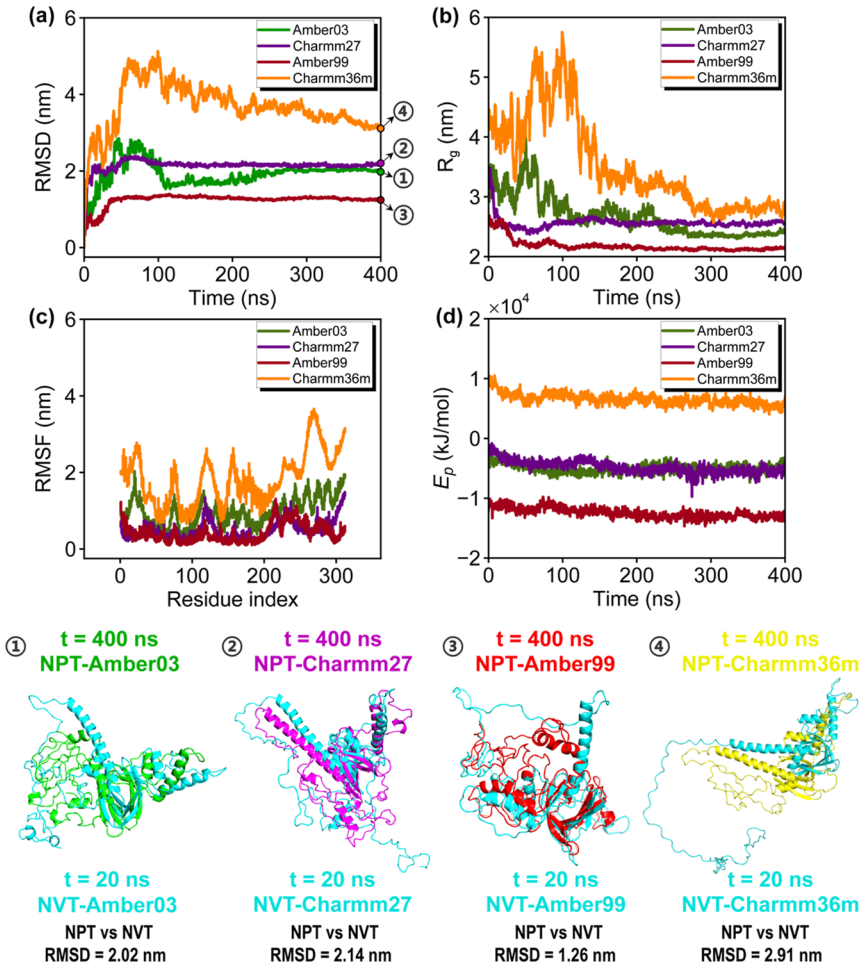


Figure 2. Structural stability evaluations of PEP from MD simulations. (a) RMSD evolutions; (b) R_g dynamics; (c) per-residue RMSF; and (d) E_p profile. Structural alignment (①–④) compares the initial *NPT* configuration (i.e., final *NVT*-relaxed structure after 20 ns) with the final *NPT*-equilibrated structure (t = 400 ns). CHARMM36m exhibits the most substantial conformational changes during *NPT* equilibration, as evidenced by its largest RMSD values calculated both by PyMOL (i.e., 2.91 nm) and MD. The RMSD shown besides of the structure ① - ④ is calculated by PyMOL with "cutoff" of 6 Å using 5 outlier rejection cycles, same as in Figure 1b–h. Conversely, Amber99 shows minimal structural variation during *NPT* but undergoes significant reorganization during the initial *NVT* phase. These differential behaviors highlight force field-dependent kinetic and thermodynamic profiles of each equilibration stage (i.e., *NVT* and *NPT*).

2. RESULTS AND DISCUSSION

2.1. Structure of PEP: AF3 vs MD. The determination of an accurate and dynamic protein structure is the foundation for a rational drug design. In this study, we retrieved the complete amino acid sequence of PEP from the UniProt database (https://www.uniprot.org/uniprotkb/A2BDE7/entry). Residues 62–195 of this sequence correspond to residues 1–134 in the construct characterized by Yousof et al.[11] This full-length sequence served as input for AF3, which generated a three-dimensional structural prediction of PEP (Figure 1a). To assess the reliability of the predicted structure, three confidence metrics were used:[20] per-residue pLDDT (predicted local distance difference test) scores, ipTM (interface predicted temperate modeling) score, and PAE (predicted aligned error). PHLDA1 contains three main functional regions:[11] two pleckstrin homology-like domains (PHLD: residues 70–110; and another PHLD region: residues 123–203) and one proline–histidine (PH)-rich region (residues 272–309). The full predicted structure is displayed in the first panel in Figure 1a, the truncated structure encompassing the two PHLD regions is shown in the middle panel, and the per-residue pLDDT distribution is presented in the final panel. A detailed breakdown of residues within each pLDDT confidence regime is provided in Table S1.

Regions with pLDDT scores above 90 (23.18%), colored dark blue, indicate high-confidence predictions with near-atomic accuracy. Scores between 70 and 90 (16.56%), shown in light blue, correspond to moderately confident regions that likely represent true structural elements but may exhibit local flexibility. Segments scoring 50–70 (18.88%), colored green, suggest low-confidence regions requiring cautious interpretation. Regions with pLDDT <50 (41.38%), depicted in orange to red, represent very low-confidence predictions typically associated with intrinsic disorder or unreliable structural features. Only a small fraction of the sequence (39.74% = 16.56% + 23.18%, pLDDT >70), primarily exhibiting the α-helical conformation (dark and light blue), demonstrates high prediction confidence. These high-confidence regions correspond to the core PHLD functional domains of PEP, supporting their suitability for subsequent dipeptide interaction studies. In contrast, the PH-rich region displays predominantly low prediction confidence (pLDDT <50), as indicated in the distribution figure (last panel of Figure 1a). The low overall model confidence is further evidenced by an

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088–50102

ipTM score of 0.48 and substantial PAE errors, with many residue pairs exhibiting distance errors exceeding 25 Å. These metrics suggest that the initial AF3-predicted structure requires further refinement, particularly as the low-confidence regions (i.e., PH region) are likely highly dynamic and prone to conformational changes during MD simulations.

The structural comparison between the AF3-predicted PEP conformation and the MD-simulated structures reveals critical insights into the protein's dynamic behavior. Through PyMOL-based structural alignment,[19] we observed significant conformational changes following MD simulation, particularly in regions that initially exhibited low prediction confidence (pLDDT <70). These flexible regions, primarily corresponding to the (PH)-rich domain, demonstrated substantial structural plasticity during simulation. In contrast, the high-confidence core domains (pLDDT >70), which encompass the PHLD regions, maintained greater structural stability (except Amber99 force field). Quantitative RMSD analysis revealed pronounced force field-dependent behavior of PEP,[21] indicated by the comparison between MD structure and predicted structure (Figure 1b−e), as well as between various force fields (Figure 1f−h). All four force fields indicated structural reoptimization of the flexible PH loop region during MD simulations. However, CHARMM27 and CHARMM36m better preserved the α-helical secondary structure elements, while Amber03 and Amber99 exhibited partial loss of these structural features. This differential performance is reflected in the RMSD values, with CHARMM27 and CHARMM36m producing the most stable simulations (RMSD ≈3.0 nm), suggesting their parameters are better suited for modeling PEP's structural dynamics, while the Amber-series force fields indicate these force fields may introduce excessive conformational sampling or insufficient restraints for this specific protein system.

The superior preservation of α-helices by the CHARMM force fields, particularly in the core PHLD domains, can be attributed to their parametrization strategy. CHARMM force fields are known for their stronger torsional potentials and optimized backbone parameters (e.g., ϕ/ψ dihedral terms) that explicitly stabilize canonical secondary structures like α-helices.[22,23] In contrast, the Amber03 and Amber99SB-ILDN force fields, while excellent for general protein dynamics, may have slightly under-restrained backbone dihedrals or a different balance between hydrogen bonding and torsional terms, which can permit excessive sampling of nonhelical conformations in inherently flexible or marginally stable regions.[24,25] This effect is particularly pronounced for IDPs like PEP, which contain large intrinsically disordered regions (e.g., the PH-rich domain) that can exert a dynamic pull on the structured domains. The stronger restraints in the CHARMM force field appear to provide sufficient resilience against this dynamic propagation, thereby maintaining the integrity of the functional PHLD domains. This finding is critical because the preservation of these helical bundles is essential for the PHLD domain's function in lipid binding and molecular recognition. The excessive conformational sampling induced by the Amber force fields, while useful for exploring conformational space, may introduce artifactual unfolding that could mislead the interpretation of PEP's true functional dynamics and binding mechanisms.

These results collectively demonstrate that while AF3 provides a valuable starting point for PEP structural studies, subsequent MD refinement is essential, particularly for accurate modeling of flexible regions. The CHARMM-series force field emerges as the most appropriate choice for such simulations, balancing structural maintenance with necessary conformational sampling. This work establishes a structural foundation for future investigations of PEP's interactions with potential therapeutic compounds, while also highlighting the need for continued development of force field parameters optimized for proteins with significant intrinsically disordered regions or dynamic domains.

2.2. Structural Stability Evaluations in MD. To comprehensively evaluate the structural evolution of the AF3-predicted PEP conformation during MD simulations, we calculated the RMSD, radius of gyration (R_g),[26] root-mean-square fluctuation (RMSF),[27] and potential energy (E_p).

As shown in Figure 2a, the RMSD values suggest distinct stabilization kinetics. CHARMM27 and Amber99 reach equilibrium rapidly (within 100 ns), suggesting efficient relaxation during the initial *NVT* phase, while Amber03 and CHARMM36m exhibit delayed stabilization: RMSD surges followed by rapid decreases, indicating substantial conformational reorganization. These patterns correlate with the radius of gyration profiles (Figure 2b): Amber03 and CHARMM36m show pronounced R_g fluctuations during the first 200 ns, while CHARMM27 and Amber99 reach equilibrium more quickly. Structural comparison (structures ① - ④) between *NVT* end points (i.e., *NPT* starting point) and final *NPT* configurations (t = 400 ns) reveals significant force field-dependent rearrangements during *NPT* equilibration, with PyMOL-calculated RMSD values of 2.02 nm (Amber03), 2.14 nm (CHARMM27), 1.26 nm (Amber99), and 2.91 nm (CHARMM36m). These values show excellent agreement with the continuous MD-derived RMSD evolutions in Figure 2a, validating the consistency between various analytical approaches and underscoring the substantial structural reorganization that occurs during production simulations.

To assess residue-specific structural variability across the four force fields, we analyzed the RMSF values for each residue (Figure 2c). The results reveal distinct dynamic profiles: CHARMM36m exhibits the most substantial atomic fluctuations overall, with pronounced mobility in seven key regions (residues 1−25, 70−90, 115−130, 150−170, 220−230, 265−275, and 295−300). Amber03 (from residues 1 to 250) and CHARMM27 show similar fluctuation patterns to CHARMM36m, though with reduced amplitude. The Amber03 exhibits different structural changes from residues 250−300, while Amber99 produces markedly different results due to partial loss of core helical structure during simulations. Although the absence of experimental structural data precludes definitive assessment of force field accuracy, comparison with the AF3-predicted structure suggests that CHARMM-class force fields provide more reliable representations, as they better preserve regions with high prediction confidence while permitting appropriate flexibility in low-confidence regions.

Finally, we examined the evolution of E_p across all four force fields (Figure 2d). CHARMM-36m exhibits consistently positive E_p values throughout the simulations, while Amber99 demonstrates the lowest E_p. Amber03 and CHARMM27 show intermediate and similar energy profiles, despite producing distinct structural conformations, highlighting the complex relationship between energy landscapes and final configurations in MD simulations. These energetic differences reflect fundamental variations in force field parametrization. The significantly lower E_p in Amber99

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

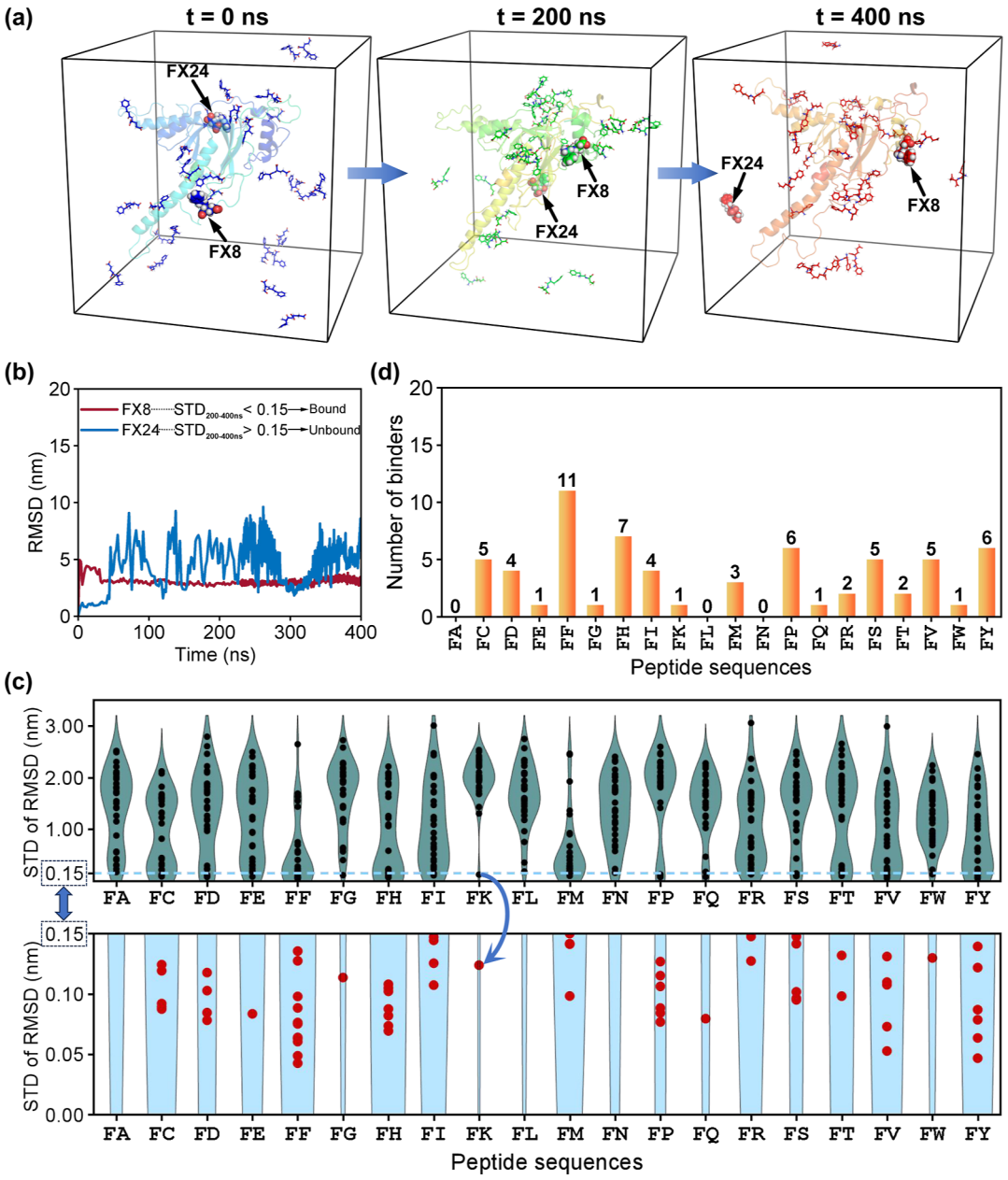


Figure 3. Identification of binding peptides through RMSD analysis. (a) Trajectories of representative binding (FX8) and nonbinding (FX24) peptides during 400 ns MD simulations. The binding peptide FX8 maintains a stable position with minor fluctuations, while FX24 exhibits random diffusion throughout the simulation box. (b) RMSD profiles demonstrate distinct behaviors: FX8 shows minimal deviation (STD <0.15 nm), whereas FX24 displays large fluctuations (STD >0.15 nm). (c) Statistical analysis of 600 dipeptides (20 types × 30 replicates) showing the distribution of STD values (upper panel) and the subset meeting the binding criterion (STD <0.15 nm, lower panel). (d) Binding frequency (out of 30 simulated) across 20 dipeptide types, revealing sequence-dependent interactions. FF dipeptides show the highest binding propensity (11/30 binders), while FA, FL, and FN exhibit no detectable binding under the defined STD threshold.

suggests stronger nonbonded interactions or different van der Waals parameters that favor more compact structures. CHARMM36m's consistently positive E_p may indicate either less favorable electrostatic interactions or different torsion potentials that maintain structural integrity at higher energy states. While absolute E_p values are not directly comparable between force fields due to their different functional forms and parameter sets, the relative reduction in E_p from initial to minimized structures demonstrates successful energy minimization and system relaxation. The preservation of the secondary structure in CHARMM-class force fields, despite higher E_p values, suggests that their parameters better balance energy minimization with biological relevance for this specific system.

2.3. Identifying Interacting Dipeptides with PEP.

Subsequently, we employed 400 ns MD simulations to systematically screen dipeptides capable of binding to PEP, using the final protein structure obtained from CHARMM27 force field simulations. Our screening library comprised 20 distinct FX dipeptides (where X represents all natural amino acids), each featuring an N-terminal phenylalanine residue maintained in the zwitterionic form without terminal capping. The dipeptide scaffold was selected for its favorable physicochemical properties, including synthetic accessibility,

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

Table 1. Dipeptides (Identified in Figure 3D)—PEP Residue Interaction Profiles[a]

dipeptides	nearby residues in PEP within 3.5 Å distance of the dipeptides
FC 5	Q107(3), Q110(3), Q111(2)
FC 9	**L**55(11), *L*99(7), **C**60(10), A62(4), **T**87(10), *E*88(7), Q102(8), Q103(2), *Q*106(8)
FC 17	G59(1), **C**60(10), K61(4), *E*88(7)
FC 26	**C**60(10), Q103(2), *Q*106(8), Q107(3), Q110(3)
FC 27	*I*85(7), **T**87(10), **L**92(12), *L*99(7), *L*140(8), *Q*102(8), *Q*106 (8), *V*138(6), *K*139(7), **E**142(9)
FD 2	K64(3), Q181(4), G182(4), *A*185(6), E186(4), L189(5), L305 (3), *R*306(8)
FD 19	R17(2), R303(2), **P**299(9), L304(5), L305(3), T308(1)
FD 24	A51(3), *M*54(8), **L**55(11), *L*99(7), *S*58(6), **C**60(10), *I*85(7), *I*94(6), **E**89(9), *L*99(7), *Q*102(8)
FD 29	*S*58(6), *Q*106(8), Q109(3), Q110(3)
FE 14	N184(4), *A*185(6), T188(2), L189(5), L304(5), P254(1), **P**299 (9), H300(4), G301(3), *R*306(8)
FF 2	H278(1), H280(2), H282(1), P279(1), P281(2), P285(1)
FF 9	L3(3), V6(1), *S*7(3), S297(4), C10(1), H268(1), H272(2), P293(3), P295(3)
FF 10	*M*54(8), *Q*102(8), Q105(2), *Q*106(8), Q109(3), *K*139(7)
FF 14	S146(1), N147(1), K149(1), K195(1), A168(1), Q198(5)
FF 17	L4(2), L8(2), L28(3), L29(2), R27(1)
FF 19	**T**87(10), **E**89(9), **E**142(9), *K*139(7)
FF 21	**L**55(11), L63(2), **C**60(10), K61(4), A62(4), **T**87(10), *E*88(7), **E**89(9), *R*197(6)
FF 22	L8(2), L12(2), L28(3), A11(1), W21(3), G22(3), E23(2)
FF 23	L1(1), L3(3), L4(2), L28(3), L29(2), S7(3), H280(2), P281(2), P293(3)
FF 24	A51(3), *M*54(8), **L**55(11), L86(2), **L**92(12), *L*99(7), *L*140(8), *S*58(6), **C**60(10), *I*85(7), *I*94(6), **T**87(10), G90(3), *Q*106(8), *K*139(7), **E**142(9)
FF 25	*E*88(7), **E**89(9), Y194(5), *R*197(6), Q198(5), L201(4)
FG 6	M161(1), Q179(5), Q181(4), *D*180(6), G182(4), N184(4), *A*185(6), L304(5), *R*306(8), S307(1)
FH 2	Q100(2), Q104(1), Q107(3), Q108(1), Q111(2), P129(1), A130(1), A132(1), V131(1), S133(1), L134(1)
FH 7	L3(3), *S*7(3), S297(4), R156(4), Q179(5), H270(3), P273(3), P293(3), P295(3), **P**299(9)
FH 12	E43(4), K159(4)
FH 18	*D*180(6)
FH 22	R35(1), A36(1), A37(2), A42(1), G38(1), G40(1), N39(1), E41(1), E43(4), K159(4)
FH 24	E23(2), E155(1), R71(3), W79(2), V154(2), K157(1), Y162 (2), D174(2)
FH 30	P273(3), H274(1)
FI 14	**T**87(10), **L**92(12)
FI 21	**T**87(10), **E**89(9), **E**142(9), G90(3), L91(4), **L**92(12), *L*140(8), L143(4), H144(2)
FI 24	*M*54(8), **L**55(11), **L**92(12), *L*99(7), *L*140(8), *I*85(7), *I*94(6), *Q*102(8), K133(1)
FI 25	R156(4), K159(4), H270(3), H272(2), H276(1), S271(3), S297 (4), P295(3)
FK 5	*R*197(6), I200(2), L201(4), S311(2), A312(1)
FM 22	Y194(5), Q198(5), Q233(1), L201(4), A202(2), S205(1)
FM 23	L63(2), K64(3), *E*88(7), E186(4), Q190(2), Q193(2), Y194(5), *R*197(6), *R*306(8)
FM 29	*M*54(8), **L**55(11), **L**92(12), **C**60(10), K61(4), *K*139(7), A62 (4), *I*85(7), **T**87(10), **E**89(9), **E**142(9)
FP 8	S49(1), R53(1), E65(1)
FP 11	Q179(5), *D*180(6), G182(4), N184(4), *A*185(6), E186(4), L189(5), L304(5), **P**299(9), *R*306(8)
FP 17	*M*54(8), **L**55(11), **L**92(12), *L*140(8), *I*94(6), Q98(3), *Q*102(8), V138(5), *K*139(7)
FP 21	A34(1), A37(2), R71(3), Q77(1), W79(2), Y162(2)
FP 23	**L**55(11), L86(2), **L**92(12), **C**60(10), A62(4), *I*85(7), **T**87(10), G90(3), **E**142(9)
FP 29	R156(4), Q179(5), *D*180(6), **P**299(9)
FQ 18	K70(2), K82(3), K141(2), L91(4), **L**92(12), L93(2), *L*140(8), L143(4), *I*94(6), P95(2), Q98(3)
FR 12	D73(1), K82(3), K97(2), P95(2), Q98(3), E135(1), V138(5), *L*140(8)
FR 25	K64(3), *R*306(8)
FS 4	**L**55(11), L91(4), L143(4), **C**60(10), **T**87(10), **E**89(9), **E**142 (9), H144(2)
FS 5	V203(1), K204(2), R207(1), Q211(1), Q227(1), Q229(1), L228(1), P230(1)
FS 16	K61(4), *E*88(7), **E**89(9), Q190(2), Q193(2), Y194(5), *R*197(6)
FS 27	**L**55(11), **L**92(12), *S*58(6), **C**60(10), V138(5), *K*139(7), **E**142 (9)
FS 30	R20(3), W21(3), G22(3), D24(1), G25(1), L30(1), L31(1), P32(1), P33(1), V154(2)
FT 7	L12(2), R13(1), R20(3), A14(1), G18(1), G22(3), S19(1), W21(3), H266(1)
FT 23	*A*185(6), T188(2), L189(5), V192(1), R303(2), **P**299(9)
FV 10	E43(4), P44(1), P96(1), S45(1), S127(1), G46(1), V67(1), K80 (1), K81(2), K97(2), C83(1), Q100(2)
FV 15	K82(3), K70(2), K141(2), L91(4), **L**92(12), L93(2), L143(4), **E**142(9), I173(1), F175(1)
FV 23	RR20(3), R71(3), T164(1), E172(1), D174(2)
FV 25	Y50(1), *M*54(8), S57(1), *S*58(6), Q102(8), *Q*106(8), P137(1), V138(5), *L*140(8)
FV 29	*E*88(7), **E**89(9), Y194(5), *R*197(6), Q198(5), Q231(1), I200 (2), L201(4), K204(2), P232(1)
FW 24	A51(3), *M*54(8), **L**55(11), **L**92(12), *L*99(7), *S*58(6), *I*85(7), *I*94(6), *Q*102(8), Q105(2), *Q*106(8), Q109(3)
FY 2	Q198(5), Q257(1), A199(1), A202(2)
FY 7	E43(4), K81(2), K159(4), Y160(1), P178(1), Q181(4)
FY 14	R156(4), Q179(5), *D*180(6), S271(3), P273(3), **P**299(9)
FY 23	*D*180(6), Q181(4), Q298(2), G182(4), N184(4), *A*185(6), E186(4), L189(5), L304(5), **P**299(9), H300(4), *R*306(8)
FY 28	H300(4), G301(3), L305(3), *R*306(8), S311(2)
FY 29	R17(2), P269(1), **P**299(9), H270(3), S271(3), S297(4), Q298 (2), H300(4), G301(3)

[a]Occurrence counts of residues across all peptide-PEP complexes shown in this table are indicated in parentheses. Residues with ≥9 occurrences are formatted in bold, while those with 6−8 occurrences appear in italics.

ease of purification, and potential for high bioavailability, making it particularly suitable for therapeutic development. The strategic incorporation of phenylalanine at the N-terminus was based on its aromatic side chain, which facilitates critical noncovalent interactions with target proteins, possibly through $\pi-\pi$ stacking and hydrophobic forces.[28,29] The design of placing the phenylalanine residue at the N-terminus was initially aimed to minimize aggregation propensity,[30] although we observed that certain phenylalanine-containing dipeptides, particularly FF, exhibited aggregation-mediated binding behavior, which unexpectedly enhanced their target engagement through multivalent interactions.

For each dipeptide variant, we constructed a simulation system comprising one PEP molecule and 30 identical dipeptide copies, followed by 400 ns production runs in the *NPT* ensemble. The peptide concentration of approximately ∼0.05 mol/L (30 peptides in a 10 × 10 × 10 nm^3 simulation box) was intentionally designed to exceed typical dipeptide solubility limits (<0.002 mol/L) to enhance sampling efficiency within the limited simulation time frame. Representative binding behavior is illustrated in Figure 3a, which

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

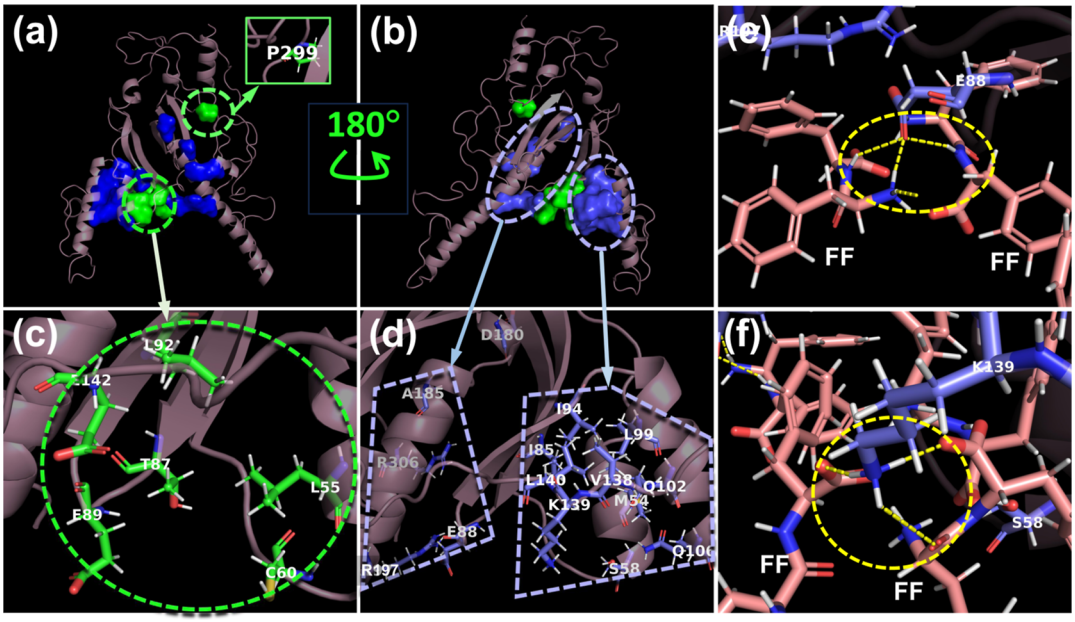


Figure 4. Spatial mapping of active interacting residues in PEP. (a,b) 180° inverted cartoon representations of PEP, with residues exhibiting interaction frequencies ≥9 (bolded in Table 1) colored green with surface representation, and those with frequencies of 6−8 displayed as blue surface. (c,d) Zoomed views of the primary binding region, highlighting high-frequency residues (≥9; green: L55, C60, T87, E89, L92, and E142) and moderate-frequency residues (6−8; blue: M54, S58, I85, E88, I94, L99, Q102, Q106, V138, K139, L140, D180, A185, R197, and R306). Residue P299 (≥9 frequency, green) is shown in the distal region of panel (a). (e,f) Structural validation of key dipeptide-PEP interactions: Hydrogen bonds (yellow dashes) between FF dipeptides (in pink) and catalytic residues E88 and K139, both of which are multivalent binding, demonstrating cooperative stabilization of peptide clusters.

contrasts stable binding (FX8) with nonbinding (FX24) interactions: the binding dipeptide maintains consistent positioning with minimal conformational drift, whereas the nonbinder exhibits unrestricted diffusion throughout the simulation box over the entire 400 ns trajectory.

To establish quantitative binding criteria, we calculated RMSD values from 200 to 400 ns for all 600 dipeptide trajectories (20 types × 30 replicates) relative to the protein structure (Figures S1−S20). It should be noted that the RMSD calculated here does not represent the structural stability of peptides. Binding peptides were rigorously identified based on a single rigorous criterion for binding: peptides with STD <0.15 nm were classified as binders, reflecting their stable positional maintenance. This approach reliably distinguishes true binders, such as FX8, which shows stable RMSD trajectories with minimal fluctuations (STD <0.15 nm, Figure 3b), from nonbinders like FX24 that display large RMSD variations (STD >0.15 nm, Figure 3b). The complete STD analysis for all 600 dipeptide trajectories (20 types × 30 replicates) is presented in Figure 3c, with upper panel showing the STD for all peptides while lower panel exhibiting the peptides with STD <0.15 nm. Figure 3d summarizes the binder number for each dipeptide type out of 30 dipeptides of the same type. The true binder number of 20 dipeptide types revealed distinct interaction patterns with the target protein. Among all tested sequences, FF exhibited the strongest binding propensity, with 11 out of 30 replicates (36.7%) meeting the binding criterion (STD <0.15 nm). This was followed by FH (23.3%), FP and FY (20.0% each), and FC, FS, and FV (16.7% each). In contrast, FA, FL, and FN showed no detectable binding under the current threshold. These results demonstrate clear sequence-dependent binding preferences, with aromatic residues (F, Y, H) generally showing higher affinity compared to aliphatic or polar residues, possibly due to $\pi-\pi$ stacking and hydrophobic interactions. While STD <0.15 nm was chosen as a relatively strict criterion for this study, it is important to recognize that alternative thresholds would yield different results (although they may not be significant). A more permissive cutoff (e.g., STD <0.20 nm) would classify more transient interactions as binders, while a stricter threshold (e.g., STD <0.10 nm) would identify only the most stable complexes.

The methodology employed in this work for identifying true binders presents several key advantages for investigating peptide−protein interactions. The high-throughput screening approach enables concurrent evaluation of multiple binding candidates, while the quantitative RMSD standard deviation criterion provides an objective alternative, yielding valuable thermodynamic information regarding binding stability while simultaneously offering kinetic insights into interaction persistence through the analysis of temporal RMSD fluctuations. Compared to static docking methods, our MD-based strategy captures the dynamic aspects of peptide binding, including potential aggregation behavior and competitive binding effects in a multipeptide environment. Furthermore, while maintaining a more cost-effective framework than that of rigorous free energy calculations, our method still provides meaningful quantitative metrics for binding characterization. These features collectively establish our protocol as a versatile and reliable approach for studying peptide−protein interactions with potential applications across diverse biological systems and binding scenarios.

2.4. Hotspots Mapping of PEP. To identify the hotspots of PEP, we systematically identified PEP residues within 3.5 Å of bound dipeptides (Table 1). Interaction frequencies are denoted in parentheses, with residues occurring ≥9 times in bold and those appearing 6−8 times in italics. Spatial mapping (Figure 4a,b) visualizes these binding sites using surface

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

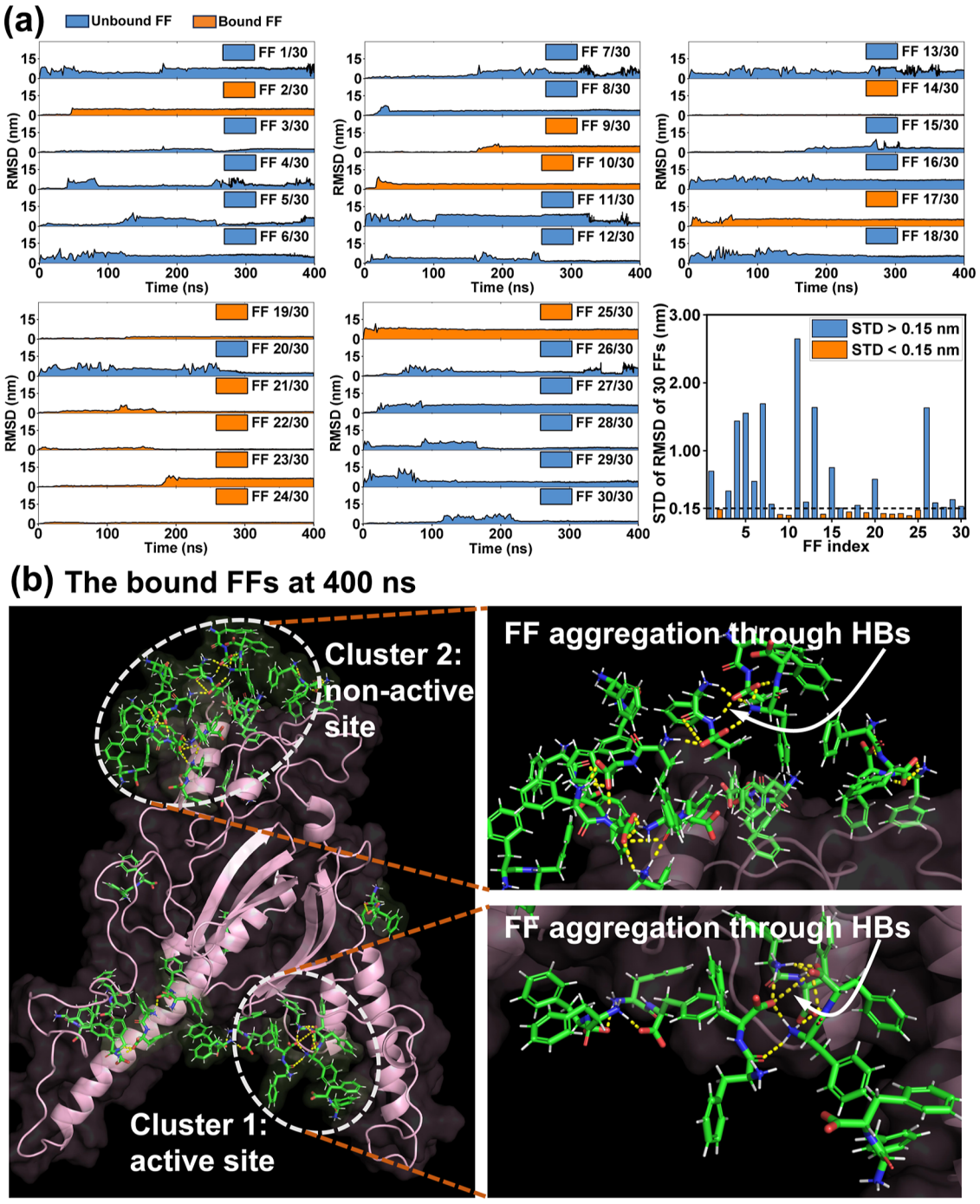


Figure 5. RMSD analysis and structural basis of cooperative FF aggregation and binding. (a) RMSD trajectories of 30 FF dipeptides during 400 ns MD simulations. STD calculated from 200 to 400 ns identifies 11 stable binders exhibiting equilibrium RMSD in this window. (b) Aggregation mediated by hydrogen bonds (yellow dashes) occurs at two sites: Cluster 1 occupies the primary binding groove, while Cluster 2 engages a secondary surface region. These structural observations reveal that (i) concurrent peptide aggregation enhances binding affinity through multivalent interactions, and (ii) surfaces of peptide aggregates can develop new binding sites through aggregation-induced structural complementarity, enabling interactions inaccessible to monomeric ligands.

representation: high-frequency residues (≥9 occurrences) are colored green, while moderate-frequency residues (6−8 occurrences) appear in blue. This analysis revealed 7 high-activity residues, i.e., L55, C60, T87, E89, L92, E142, and P299 (Figure 4c) and 15 moderately active residues, i.e., M54, S58, I85, E88, I94, L99, Q102, Q106, V138, K139, L140, D180, A185, R197, and R306 (Figure 4d) among PEP's 312 residues. Collectively, these residues form a continuous binding groove (Figure 4a,b), with high-frequency residues clustering at the interaction hotspot. Direct evidence for the stabilizing role of identified residues is provided by the FF-PEP interaction (Figure 4e,f). Residues E88 and K139 participate in forming three distinct hydrogen bonds with separate FF molecules, demonstrating a multivalent binding capacity. These specific interactions exemplify how the high-frequency residues

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

characterized in Table 1 and Figure 4a−d mechanistically contribute to peptide stabilization.

2.5. Aggregation-Mediated Binding between FF and PEP. The FF dipeptide, extensively characterized as an aggregation core in amyloid-Aβ pathologies associated with neurodegenerative diseases, emerges as the highest-affinity binder for PEP in this study. Of the 30 simulated FF peptides, 11 met the binding criterion (STD <0.15 nm during 200−400 ns), exhibiting stable RMSD trajectories indicative of persistent target engagement (Figure 5a). Structural analysis reveals that these bound FF peptides simultaneously undergo aggregation primarily into two hydrogen-bonded clusters (white dashes, Figure 5b), demonstrating that peptide aggregation and target binding occur cooperatively rather than sequentially.

Two major FF clusters occupy spatially distinct regions on PEP's surface: cluster 1 localizes precisely within the previously identified high-activity binding groove (Figure 4c,d), while cluster 2 engages a secondary site (i.e., nonactive binding site). The groove-localized cluster 1 shows particularly stable interactions, with intermolecular hydrogen bond persistence exceeding 85% occupancy during the production phase, suggesting that this region serves as the primary affinity hotspot.

The simultaneous aggregation of FF peptides during target engagement reveals new mechanistic principles distinct from those of conventional docking paradigms. Namely, aggregation actively constructs complementary binding surfaces through collective hydrogen bonding (yellow dashes, Figure 5b), enhancing FF-PEP binding affinity via multivalent avidity, a capability unattainable by isolated dipeptides. In detail, PEP's spatially dispersed active residues (e.g., K139 and E142 separated by >1.8 nm) cannot be effectively engaged by individual FF molecules, but peptide aggregates bridge these discontinuous epitopes through hydrogen bonding, essentially "stapling" distant sites into a cohesive binding interface. This aggregation-mediated binding mechanism fundamentally diverges from mainstream docking approaches, which typically assume monomeric ligands binding preformed pockets, by demonstrating how dynamic oligomerization generates novel interfacial geometries that overcome topological mismatches. This is also corroborated by the fact that additional binding surfaces beyond primary active sites can be generated by peptide aggregation (i.e., cluster 2). Therapeutically, this suggests that PEP (and possibly other protein targets as well) inhibition might adopt three strategies: one is targeting the primary active groove, and the second is employing aggregating/self-assembling peptides for targeting the nonactive surface, or cooperatively targeting the primary site and the nonactive sites induced by aggregation.

2.6. Concentration Effect on the Binding of FF to PEP. To quantitatively evaluate the influence of ligand concentration and force field parametrization on the binding of FF dipeptides to the target protein PEP, we conducted seven additional all-atom MD simulations using the CHARMM36m force field at varied concentrations (Table 2). In our earlier simulations employing the CHARMM27 force field, a concentration of ∼0.05 mol/L was set, which exceeds the solubility limit of the dipeptide (normally <0.002 mol/L) and also the critical aggregation concentration (CAC, <0.03 mol/L[31]). As both aggregation and binding are inherently concentration-dependent, a systematic assessment of concentration effects is essential to provide a rigorous basis for interpreting peptide−protein interactions.

Table 2. MD Simulations Performed with the CHARMM36m Force Field to Quantify the Effect of FF Ligand Concentration on Binding Behavior[a]

	L (nm)	N_{FF}	c_{FF} (mol/L)	rep (s)	time (ns)	N_{binder}	residues
1	10 × 10 × 10	10	0.0166	1	800	1	E142
2	10 × 10 × 10	20	0.0332	1	800	4	L1, E142, Q109, R156
3	10 × 10 × 10	30	0.0498	1	800	12	E142, E122, E89
4	18 × 18 × 18	1	0.0003	4	800	1/4	E142
5	18 × 18 × 18	10	0.0028	1	400	1	E142
6	18 × 18 × 18	20	0.0057	1	400	1	E122, Q125, S127
7	18 × 18 × 18	30	0.0085	1	400	0	

[a]L: simulation box length in nanometers; N_{FF}: number of FF peptides in the simulation box; c_{FF}: concentration in mol/L; rep (s): replicates of independent simulations; time: total production time per simulation in nanoseconds; N_{binder}: number of FF binders identified as stably bound; residue(s): PEP residues forming persistent interactions with FF, with E142 identified as the primary binding hotspot.

Simulations 1−3 in Table 2, performed with the CHARMM36m force field,[23] demonstrated that the number of stably bound FF peptides increased with an increasing concentration. A consistent finding across all concentrations was the interaction of FF with residue Glu142 on PEP (Table 2 and Figure 6a−c). At higher concentrations, additional binding epitopes were engaged, including residues Leu1, Gln109, and Arg156 (Figure 6b), as well as Glu122 (not shown) and Glu89 (Figure 6c). This coincides with our results of the CHARMM27 force field; i.e., Glu89 and Glu142 were identified as the most prominent binding hotspots, each observed in 9 independent binding events (bolded in Table 1). Residues Gln109 and Arg156 were also significant, appearing in 3 and 4 instances in previous simulations, respectively. Notably, at the highest concentration (simulation 3, 0.0498 mol/L), we observed the spontaneous aggregation of FF peptides into a large aggregate comprising approximately 15 monomers stabilized by a network of intermolecular hydrogen bonds (Figure 6c). This supramolecular complex formed stable, multipoint attachments to the PEP surface, as in the CHARMM27 force field. In summary, the MD simulations performed with the CHARMM36m force field corroborate our previous results obtained by the CHARMM27 force field.

To definitively identify the primary binding residue on PEP and decouple binding events from potential concentration-dependent aggregation effects, we conducted four additional MD simulations (simulations 4−7, Table 2) within a larger simulation volume (18 × 18 × 18 nm^3), maintaining FF concentrations below 0.01 mol/L (below the CAC). At the ultralow concentration of 0.0003 mol/L (a single FF peptide in the 18 × 18 × 18 nm^3 box), we performed four independent replicates with randomized initial velocities to ensure statistical significance. Binding to PEP was observed in only one of the four simulations, with the remaining three exhibiting purely diffusive peptide behavior, consistent with the expected rarity of spontaneous binding events at this concentration. In the binding replicate, the peptide's RMSD stabilized after ∼400 ns (Figure 6d). The STD of the RMSD (0.20 nm) was marginally

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

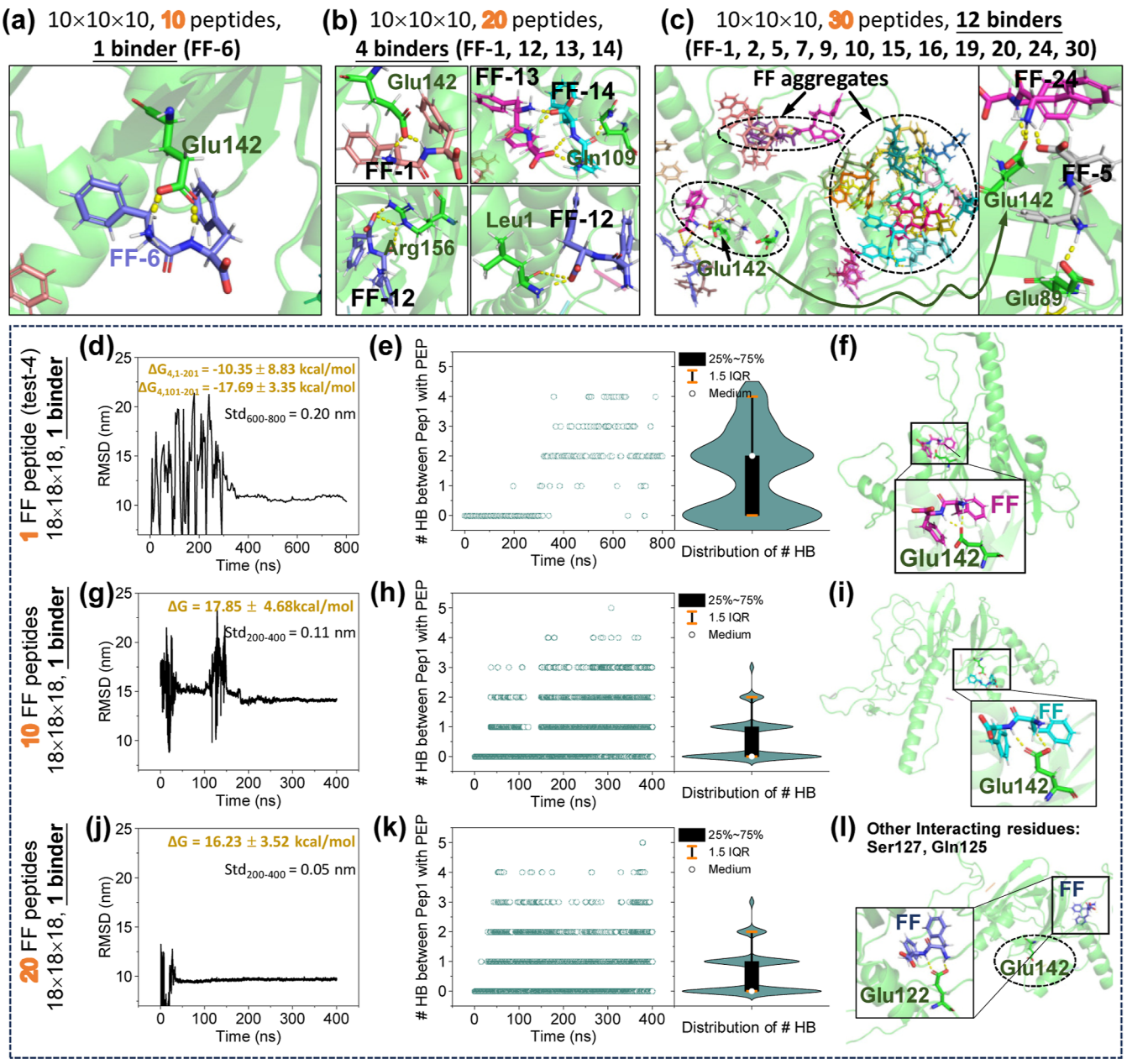


Figure 6. Concentration dependence of FF dipeptide binding to PEP simulated with the CHARMM36m force field. (a–c) Interactions between PEP and FF ligands within a 10 × 10 × 10 nm^3 simulation box at concentrations of 0.0166 mol/L (10 peptides), 0.0322 mol/L (20 peptides), and 0.0498 mol/L (30 peptides). The FF dipeptides consistently interact with Glu142 on PEP. At higher concentrations, additional interactions involve residues Leu1, Gln109, Arg156, Glu122, and Glu89. A large FF aggregate consisting of 15 monomers is observed at 0.0498 mol/L. (d–f) Results from simulations performed with a larger box of 18 × 18 × 18 nm^3, showing RMSD (and also STD and binding free energy ΔG), number of hydrogen bonds (# HBs), and residue-level interaction details. A single binding site is consistently occupied across all concentrations except 0.0085 mol/L, with no aggregation observed at all concentrations. Although the ΔG calculations (values exceeding −15 kcal/mol) indicate high-affinity binding, those data should be interpreted with caution due to the limited kinetic and concentration regime in MD simulations.

higher than our initial stringent threshold (0.15 nm), which can be attributed to the increased flexibility of the isolated peptide in the large box. Crucially, the binding free energy (ΔG) calculation using the gmx_mmpbsa method[32] yielded a value of $\Delta G_{4,101-201}$ = −17.69 ± 3.35 kcal/mol for frames between 400 and 800 ns (from 101 to 201), indicating strong, specific binding. Here, 4 in the subscript indicates the fourth binding replicate. In contrast, including the initial diffusive phase (0–400 ns) resulted in a less favorable and more uncertain energy ($\Delta G_{4,1-201}$ = −10.35 ± 8.83 kcal/mol), underscoring the importance of defining the bound state for accurate quantification. This interaction was stabilized by the formation of 2 (mostly) to 4 persistent hydrogen bonds (Figure 6e) between the FF ligand and PEP, and was unequivocally anchored by residue Glu142 (Figure 6f). At intermediate concentrations of 0.0028 mol/L (10 peptides in the 18 × 18 × 18 nm^3 box) and 0.0057 mol/L (20 peptides in the 18 × 18 × 18 nm^3 box), a single binding event was observed per simulation, primarily involving E142 alongside secondary interactions with E122, Q125, and S127. All metrics, including binding free energy, RMSD stability, hydrogen bond count, and direct interaction analysis (Figure 6g–l) confirmed

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

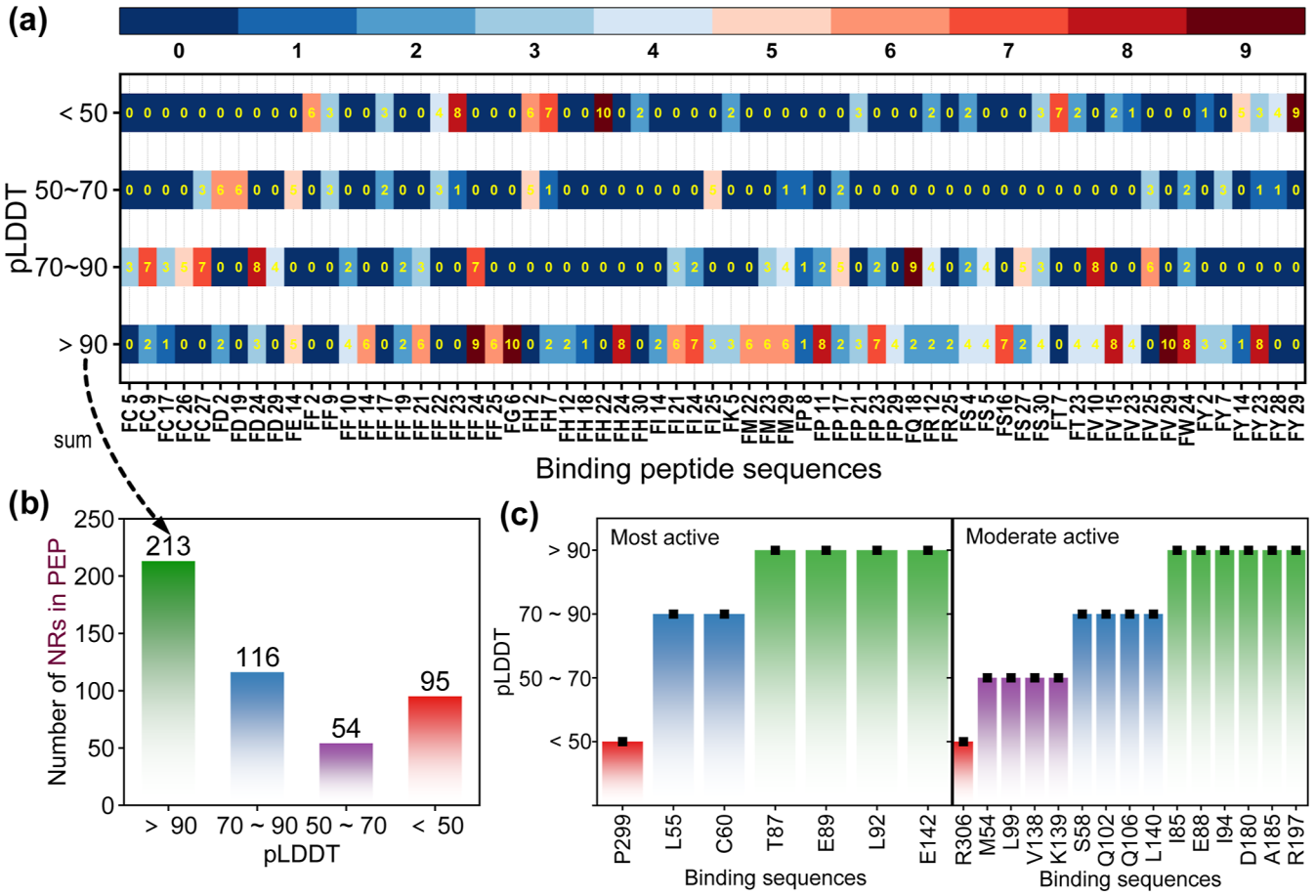


Figure 7. Correlation between pLDDT values and active residues. (a) Number of PEP residues located within each pLDDT range that is within 3.5 Å of each dipeptide. The color bar above indicates the number of residues, and the number is also directly shown. (b) Total number of nearby residues within each pLDDT range. The acronym "NRs" in the *y* axis stands for nearby residues. The residues with pLDDT values smaller than 50 also actively participate in binding. (c) pLDDT values of the most active and moderately active PEP residues.

these as high-affinity binding events. Interestingly, at the highest concentration in this low-range series (0.0085 mol/L, 30 peptides), neither binding nor aggregation was observed. This suggests that while specific, high-affinity binding to residues like E142 can occur at picomolar-to-nanomolar concentrations, the probability of such an event within a practical simulation time frame (e.g., 400 ns) remains low. The absence of binding at 0.0085 mol/L, in contrast to the pronounced effects seen at higher concentrations (~0.05 mol/L) in smaller boxes, highlights a concentration-dependent threshold for observable binding and aggregation. This aligns with the concept that molecular recognition can operate through both monomeric binding and collective aggregation-mediated mechanisms, with the dominant pathway dictated by ligand availability.

At elevated concentrations, the propensity for FF dipeptides to self-assemble increases, a well-documented phenomenon driven by hydrophobic effects and intermolecular van der Waals forces.[33,34] To isolate monomeric binding behavior, we deliberately employed concentrations above the theoretical CAC. Crucially, however, our data reveal that peptide aggregation is not necessarily deleterious to target engagement (as also in another work[35]). On the contrary, the self-assembled aggregate observed in our simulations forms a stable, multivalent complex with PEP, suggesting a mechanism of aggregation-mediated binding. This phenomenon aligns with an emerging paradigm in drug discovery where controlled supramolecular assembly enhances pharmacological function.[36−39] An aggregate-based drug delivery strategy can offer significant advantages, notably prolonged systemic circulation and increased residency time at the target site due to altered pharmacokinetics.[40] This approach is being actively investigated for applications such as targeting the unique microenvironment of tumors, where factors like pH or enzyme activity can be leveraged to trigger precise, localized aggregation from systemically administered precursors.[41,42] Therefore, the FF aggregation observed here may not be an artifact to be mitigated but a feature to be harnessed, potentially transforming a simple dipeptide into a novel supramolecular therapeutic agent with enhanced efficacy and tunable release properties.

2.7. Correlation between pLDDT and Binding Sites.

AF3 is a prevalent protein structure prediction model, with pLDDT serving as a key metric for evaluating the per-residue prediction confidence. This study investigates the spatial correlation between pLDDT values and ligand-binding sites to assess whether pLDDT can reliably indicate the functional regions involved in molecular interactions. Figure 7a illustrates the pLDDT distribution of PEP residues interacting with dipeptides, revealing that active residues span the entire pLDDT spectrum, though a significant proportion exhibit high confidence (pLDDT >90). Notably, residues binding dipep-

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

tides FF, FH, FT, and FY frequently display low confidence (pLDDT <50). Figure 7b quantifies this trend, showing residues with pLDDT <50 outnumber those in the 50−70 range, suggesting disordered regions (e.g., PH regions of PEP) frequently participate in ligand binding. Figure 7c further highlights pLDDT values of high- and moderate-activity residues (shown in Figure 4c,d), reinforcing that both high-confidence (pLDDT >90) and low-confidence (pLDDT <50) regions can constitute functional binding sites.

The prevalence of binding activity in low-pLDDT regions (pLDDT <50) challenges the rigid "lock-and-key" paradigm,[43,44] indicating these likely represent intrinsically disordered regions (IDRs) or flexible loops that adopt structure upon ligand binding via induced fit or conformational selection. This aligns with dynamic binding mechanisms, where conformational plasticity enables interactions with diverse ligands. Our data suggest that pLDDT may underestimate the functional relevance of such flexible regions due to its emphasis on static structures. Conversely, the concentration of binding residues in high-pLDDT regions (pLDDT >90) underscores the role of well-defined pockets in molecular recognition, reinforcing pLDDT's utility for identifying structured sites while highlighting that high confidence alone cannot define functionality, as many such residues are noninteracting.

Collectively, pLDDT operates as a context-dependent indicator: High values (>70) strongly signal structurally ordered regions where binding typically follows conventional pocket-based mechanisms, while low values (<50) suggest structural disorder yet do not preclude function, i.e., these regions may mediate transient, dynamic, or allosteric interactions critical for signaling or promiscuous binding. The bimodal distribution in Figure 6b implies that pLDDT reliably identifies ordered binding pockets but may overlook functional IDRs. Consequently, relying solely on pLDDT for binding site prediction risks missing flexible interfaces; integrating it with metrics of evolutionary conservation, physicochemical pocket properties, or dynamics-focused algorithms (e.g., MD simulations) could improve functional annotation. The prominence of low-pLDDT binding in FF/FH/FT/FY interactions warrants investigation into whether aromatic dipeptides preferentially target disordered regions through an aggregation-induced hydrophobicity mechanism (AIHB).[35]

3. CONCLUSIONS

In this study, we present a comprehensive computational strategy to explore the structure and druggability of PEP, a key apoptotic regulator in cardiomyocytes. By integrating AF3 predictions with multiforce-field MD simulations, we resolved PEP's dynamic architecture and identified its force field-sensitive regions, establishing CHARMM-class force fields as the most suitable for simulating this intrinsically disordered protein. We further demonstrated that dipeptides, particularly FF, can bind PEP with high affinity and specificity and uncovered an unconventional aggregation-mediated binding mechanism that leverages multivalency to engage discontinuous epitopes. This mechanism not only enhances binding affinity but also opens new avenues for targeting proteins with complex surface topologies. The concentration-dependent binding behavior of FF dipeptides and their robust interaction with residues such as Glu142 support their potential as lead candidates for inhibiting PEP-mediated apoptosis. These findings provide a foundational framework for future experimental validation and rational design of peptide-based therapeutics targeting PEP in cardiovascular diseases and beyond. The methodologies and insights described here may also be extended to other challenging targets with significant intrinsically disordered regions, offering a generalizable strategy for discovering and optimizing peptide inhibitors where conventional small molecules have struggled.

4. EXPERIMENTAL SECTION

4.1. Structure Prediction by AF3. The amino acid sequence of the human PHLDA1 protein was retrieved from the UniProt database (web site: https://www.uniprot.org/uniprotkb/A2BDE7/entry). The complete protein sequence contains 312 amino acids. This sequence was subsequently submitted to AF3 to generate a predicted three-dimensional protein structure including the atomic coordinates of all constituent residues.

4.2. Structure Optimization by MD. To assess the structural stability and accuracy of the predicted PHLDA1 model, we conducted all-atom MD simulations using a GROMACS 2023.2. Simulations were performed in parallel using distinct force fields: AMBER03;[45] the CHARMM27 all-atom force field;[46] and AMBER99SB-ILDN;[47] and CHARMM36m.[23] For each system, the protein structure was solvated within a cubic periodic boundary box (18 × 18 × 18 nm^3) filled with explicit TIP3P water molecules. Before energy minimization, the systems were charge-neutralized by adding a proper amount of Cl^- or Na^+ counterions. Production simulations comprised sequential equilibration: 20 ns under *NVT* ensemble conditions (constant number of particles, volume, and temperature = 300 K) followed by 400 ns under *NPT* ensemble conditions (constant number of particles, temperature = 300 K, pressure = 1 bar). A periodic boundary condition (PBC) was applied in three dimensions. Trajectory after production was processed by applying GROMACS utilities to enforce consistent handling: the protein was centered in the box (-pbc center), translational jumps across PBCs were corrected (-pbc nojump), and overall system rotation/translation was removed via least-squares fitting to the initial structure (-fit rot + trans). Finally, "-pbc mol" was applied to maintain molecular continuity and ensure the protein remained unconstrained by box boundaries throughout the analyzed trajectory frames.

4.3. Structure Stability Assessment in MD. Given that an accurate structural model is fundamental for structure-based drug design, we assessed the optimized PEP stability by evaluating key kinetic and thermodynamic metrics: RMSD, R_g, E_p, and RMSF, which were calculated directly from the postprocessed trajectories using GROMACS' built-in commands. Specifically, the command "gmx rms -s ref.gro -f traj.xtc -o rmsd.xvg" computed the RMSD relative to the initial structure. RMSF per residue was generated via "gmx rmsf -s ref.gro -f traj.xtc -o rmsf.xvg", while R_g was calculated using "gmx gyrate -s topol.tpr -f traj.xtc -o R_g.xvg". To isolate the protein's potential energy contribution from the solvated system, a multistep extraction was performed: First, a protein-only trajectory (prolig.xtc) was created with "gmx trjconv -f full.xtc -s full.tpr -o prolig.xtc". The corresponding protein topology file (prolig.tpr) was then generated using "gmx convert-tpr -s full.tpr -n prolig.ndx -o prolig.tpr". This enabled energy recalculation specifically for the protein conformation along the trajectory using "gmx mdrun -s prolig.tpr -rerun

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

prolig.xtc -e prolig.edr". Finally, the time series of protein potential energy was extracted via "gmx energy -f prolig.edr -o prolig-energy.xvg". All quantitative plots were generated using OriginPro, while molecular structure figures were rendered in PyMOL.

4.4. MD Simulations of Peptide−Protein Interaction. To investigate dipeptide-PEP binding interactions, MD simulations were performed with the CHARMM27 force field using the CHARMM27 force field-optimized PEP structure. The protein was solvated in a cubic periodic boundary box (10 × 10 × 10 nm) with 30 identical dipeptides (FA, FC, FD, FE, FF, FG, FH, FI, FK, FL, FM, FN, FP, FQ, FR, FS, FT, FV, FW, and FY), which were added per simulation system. To investigate the concentration and force field effect, additional simulations with 1, 10, 20, 30 identical dipeptides were simulated in a 18 nm × 18 nm × 18 nm box with CHARMM36m force field, and also 10, 20, 30 identical dipeptides were simulated in a 10 nm × 10 nm × 10 nm box with CHARMM36m force field (see Table 2 for details). Each system underwent the established protocol: energy minimization followed by 20 ns of *NVT* equilibration and 400 ns (or 800 ns) of *NPT* production simulation at 300 K and 1 bar. Trajectories were postprocessed using identical periodic boundary condition corrections as applied during structural assessment. The ligand binding event was quantified by calculating RMSD values of dipeptides relative to the protein structure. This involved: (1) generating peptide-specific index groups (gmx make_ndx -f sys.gro -o peptides.ndx); (2) extracting the t = 0 ns reference structure of protein-peptide (gmx trjconv -f sys.xtc -s sys.tpr -o ref.gro -b 0 -e 0); (3) iteratively computing RMSD for each peptide group relative to the protein target using gmx rms -s ref.gro -f sys.xtc -n peptides.ndx -o rmsd-pepX.xvg. The standard deviation of RMSD values during the 200−400 ns trajectory segment was calculated for each dipeptide. Peptides exhibiting a standard deviation (STD) of RMSD below 0.15 nm were identified as stable binders for subsequent analysis.

4.5. Binding Free Energy Calculation Using gmx_MMPBSA. For the binding free energy calculations presented in Figure 6d,g,j, we employed the gmx_MMPBSA[32] tool using the following parameter set: startframe = 1, endframe = 201, interval = 1, igb = 2, saltcon = 0.15. In Figure 6d, the trajectory was extracted from 0 to 800 ns with an output interval of 4 ns, resulting in a total of 201 frames. To evaluate the influence of nonequilibrium conformations on the calculated binding energy, we performed calculations using two distinct time windows: frames 1−201 (0−800 ns), which include the initial equilibration phase, and frames 101−201 (400−800 ns), which represent the fully converged equilibrium segment of the simulation. For Figure 6g,j, the binding free energies were derived exclusively from the equilibrated region of the trajectory. Specifically, 101 frames between 200 and 400 ns were used, ensuring that the energy values reflect stable, equilibrium binding interactions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.5c06570.

Table summarizing the residues within each pLDDT confidence regime and figures showing the RMSD values for all 600 dipeptides (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Meng Zhao − *Frontier Innovation Center, Department of Physiology and Pathophysiology, School of Basic Medical Sciences, Fudan University, Shanghai 200032, P. R. China*; Email: zhaomeng@fudan.edu.cn

Jiaqi Wang − *Wisdom Lake Academy of Pharmacy, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, P. R. China; Jiangsu Province Higher Education Key Laboratory of Cell Therapy Nanoformulation (Construction), Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, P. R. China; Department of Chemistry, State Key Laboratory of Synthetic Chemistry, The University of Hong Kong, Hong Kong SAR 999077, P. R. China*; orcid.org/0000-0002-5045-1497; Email: benwang@hku.hk

Authors

Shujia Liu − *Wisdom Lake Academy of Pharmacy, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, P. R. China; Jiangsu Province Higher Education Key Laboratory of Cell Therapy Nanoformulation (Construction), Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, P. R. China; Suzhou Industrial Park Center of Excellence for Single-Molecule and Single-Cell Imaging and Analysis, Xi'an Jiaotong-Liverpool University, Suzhou 215123, P. R. China*

Haojin Zhou − *Wisdom Lake Academy of Pharmacy, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, P. R. China*

Weihua Bian − *Department of Pharmacy, Binzhou Medical University, Yantai, Shandong 264003, P. R. China*

Kanghui Huan − *Frontier Innovation Center, Department of Physiology and Pathophysiology, School of Basic Medical Sciences, Fudan University, Shanghai 200032, P. R. China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.5c06570

Author Contributions

¶Shujia Liu and Dr. Haojin Zhou are co-first authors.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.W. acknowledges the funding support of the National Natural Science Foundation of China (Grant No. 52101023) and Basic Research Program of Jiangsu—General Program of Jiangsu Provincial Department of Science and Technology (Grant No. BK20241816). M.Z. acknowledges the funding support of the National Natural Science Foundation of China (Grant No. 82170238). W.B. acknowledges the funding support of the Natural Science Foundation of Shandong Province (Grant No. ZR2024MH336) and National Natural Science Foundation of China (Grant No. 82270305). The authors also acknowledge the high-performance computing platform at Xi'an Jiaotong-Liverpool University and Beijing Paratera Tech Corp., LTD.

■ REFERENCES

(1) Nabel, E. G.; Braunwald, E. A Tale of Coronary Artery Disease and Myocardial Infarction. *N. Engl. J. Med.* **2012**, *366*, 54−63.

(2) Ambrose, J. A. Myocardial Ischemia and Infarction. *J. Am. Coll. Cardiol.* **2006**, *47*, D13−D17.

(3) Heinzel, F. R.; Hohendanner, F.; Jin, G.; Sedej, S.; Edelmann, F. Myocardial Hypertrophy and Its Role in Heart Failure with Preserved Ejection Fraction. *J. Appl. Physiol.* **2015**, *119*, 1233−1242.

(4) He, L.; Nguyen, N. B.; Ardehali, R.; Zhou, B. Heart Regeneration by Endogenous Stem Cells and Cardiomyocyte Proliferation: Controversy, Fallacy, and Progress. *Circulation* **2020**, *142*, 275−291.
(5) Li, X.; Wu, F.; Günther, S.; Looso, M.; Kuenne, C.; Zhang, T.; Wiesnet, M.; Klatt, S.; Zukunft, S.; Fleming, I. Inhibition of Fatty Acid Oxidation Enables Heart Regeneration in Adult Mice. *Nature* **2023**, *622*, 619−626.
(6) Braunwald, E. The War against Heart Failure: the Lancet Lecture. *Lancet* **2015**, *385*, 812−824.
(7) Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discovery Today* **2015**, *20*, 122−128.
(8) Otvos, L. *Peptide-based Drug Design: Here and Now*; Peptide-based Drug Design, 2008; Vol. *1*−*8*, pp 1−8.
(9) Evers, A.; Pfeiffer-Marek, S.; Bossart, M.; Heubel, C.; Stock, U.; Tiwari, G.; Gebauer, B.; Elshorst, B.; Pfenninger, A.; Lukasczyk, U.; et al. Peptide Optimization at the Drug Discovery-Development Interface: Tailoring of Physicochemical Properties Toward Specific Formulation Requirements. *J. Pharm. Sci.* **2019**, *108*, 1404−1414.
(10) Guo, Y.; Jia, P.; Chen, Y.; Yu, H.; Xin, X.; Bao, Y.; Yang, H.; Wu, N.; Sun, Y.; Jia, D. PHLDA1 is a New Therapeutic Target of Oxidative Stress and Ischemia Reperfusion-induced Myocardial Injury. *Life Sci.* **2020**, *245*, 117347.
(11) Yousof, T.; Byun, J. H.; Chen, J.; Austin, R. C. Pleckstrin Homology-like Domain, Family A, Member 1 (PHLDA1): A Multifaceted Cell Survival Factor that Drives Metabolic Disease. *Engineering* **2023**, *20*, 9−18.
(12) Nagai, M. A. Pleckstrin Homology-like Domain, Family A, Member 1 (PHLDA1) and Cancer. *Biomed. Rep.* **2016**, *4*, 275−281.
(13) Park, C. G.; Lee, S. Y.; Kandala, G.; Lee, S. Y.; Choi, Y. A Novel Gene Product that Couples TCR Signaling to Fas (CD95) Expression in Activation-induced Cell Death. *Immunity* **1996**, *4*, 583−591.
(14) Chen, Y.; Takikawa, M.; Tsutsumi, S.; Yamaguchi, Y.; Okabe, A.; Shimada, M.; Kawase, T.; Sada, A.; Ezawa, I.; Takano, Y.; et al. PHLDA1, another PHLDA family protein that inhibits Akt. *Cancer Sci.* **2018**, *109*, 3532−3542.
(15) Zhao, J.; Zhang, Q.; Cheng, W.; Dai, Q.; Wei, Z.; Guo, M.; Chen, F.; Qiao, S.; Hu, J.; Wang, J.; et al. Heart−gut microbiota communication determines the severity of cardiac injury after myocardial ischaemia/reperfusion. *Cardiovasc. Res.* **2023**, *119*, 1390−1402.
(16) Nguyen, P. H.; Ramamoorthy, A.; Sahoo, B. R.; Zheng, J.; Faller, P.; Straub, J. E.; Dominguez, L.; Shea, J.-E.; Dokholyan, N. V.; De Simone, A.; et al. Amyloid Oligomers: A Joint Experimental/Computational Perspective on Alzheimer's Disease, Parkinson's Disease, Type II Diabetes, and Amyotrophic Lateral Sclerosis. *Chem. Rev.* **2021**, *121*, 2545−2647.
(17) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: from Algorithms to Applications*; Elsevier, 2023.
(18) AlRawashdeh, S.; Barakat, K. H. Applications of Molecular Dynamics Simulations in Drug Discovery. *Computational Drug Discovery and Design* **2024**, *2714*, 127−141.
(19) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129−1143.
(20) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630*, 493−500.
(21) Kumar, D. T.; Susmita, B.; Judith, E.; Priyadharshini Christy, J.; George Priya Doss, C.; Zayed, H. Elucidating the Role of Interacting Residues of the MSH2-MSH6 Complex in DNA Repair Mechanism: A Computational Approach. *Adv. Protein Chem. Struct. Biol.* **2019**, *115*, 325−350.
(22) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671−690.
(23) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71−73.
(24) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(25) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(26) Rampogu, S.; Shaik, B.; Kim, J. H.; Jung, T. S.; Ha, M. W.; Lee, K. W. Explicit Molecular Dynamics Simulation Studies to Discover Novel Natural Compound Analogues as Mycobacterium Tuberculosis Inhibitors. *Heliyon* **2023**, *9*, No. e13324.
(27) Margreitter, C.; Oostenbrink, C. MDplot: Visualise Molecular Dynamics. *R J.* **2017**, *9*, 164.
(28) Wang, J.; Liu, Z.; Zhao, S.; Xu, T.; Wang, H.; Li, S. Z.; Li, W. Deep Learning Empowers the Discovery of Self-Assembling Peptides with over 10 Trillion Sequences. *Adv. Sci.* **2023**, *10*, 2301544.
(29) Xu, T.; Wang, J.; Zhao, S.; Chen, D.; Zhang, H.; Fang, Y.; Kong, N.; Zhou, Z.; Li, W.; Wang, H. Accelerating the Prediction and Discovery of Peptide Hydrogels with Human-in-the-Loop. *Nat. Commun.* **2023**, *14*, 3880.
(30) Frederix, P. W.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T. Exploring the Sequence Space for (Tri-) Peptide Self-Assembly to Design and Discover New Hydrogels. *Nat. Chem.* **2015**, *7*, 30−37.
(31) Roy, S. S.; Ghosh, R.; Mukherjee, D.; Biswas, S.; Pal, S. K.; Banerjee, R. F. F. FF Activates Nucleation for Aggregation in Aqueous Medium: A Key Insight of Amyloid Plaque Formation. *ChemistrySelect* **2024**, *9*, No. e202304127.
(32) Valdés-Tresanco, M. S.; Valdés-Tresanco, M. E.; Valiente, P. A.; Moreno, E. gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17*, 6281−6291.
(33) Gazit, E. Self-Assembled Peptide Nanostructures: the Design of Molecular Building Blocks and Their Technological Utilization. *Chem. Soc. Rev.* **2007**, *36*, 1263−1269.
(34) Wang, J.; Liu, Z.; Zhao, S.; Zhang, Y.; Xu, T.; Li, S. Z.; Li, W. Aggregation Rules of Short Peptides. *JACS Au* **2024**, *4*, 3567−3580.
(35) Chen, X.; Zhou, H.; He, X.; Wang, J. Computational Design of Next-Gen Peptide Biopesticides: Targeting the Nicotinic Acetylcholine Receptor in Rice Pests Class. *bioRxiv* **2025**, 661460.
(36) Du, X.; Zhou, J.; Shi, J.; Xu, B. Supramolecular Hydrogelators and Hydrogels: from Soft Matter to Molecular Biomaterials. *Chem. Rev.* **2015**, *115*, 13165−13307.
(37) Webber, M. J.; Appel, E. A.; Meijer, E.; Langer, R. Supramolecular Biomaterials. *Nat. Mater.* **2016**, *15*, 13−26.
(38) Zhou, X.-R.; Cao, Y.; Zhang, Q.; Tian, X.-B.; Dong, H.; Chen, L.; Luo, S.-Z. Self-Assembly Nanostructure Controlled Sustained Release, Activity and Stability of Peptide Drugs. *Int. J. Pharm.* **2017**, *528*, 723−731.
(39) Habibi, N.; Kamaly, N.; Memic, A.; Shafiee, H. Self-Assembled Peptide-based Nanostructures: Smart Nanomaterials toward Targeted Drug Delivery. *Nano Today* **2016**, *11*, 41−60.
(40) Webber, M. J.; Langer, R. Drug Delivery by Supramolecular Design. *Chem. Soc. Rev.* **2017**, *46*, 6600−6620.
(41) Zhou, J.; Du, X.; Berciu, C.; He, H.; Shi, J.; Nicastro, D.; Xu, B. Enzyme-instructed Self-Assembly for Spatiotemporal Profiling of the Activities of Alkaline Phosphatases on Live Cells. *Chem.* **2016**, *1*, 246−263.
(42) Liu, J.; Liu, J.; Chu, L.; Zhang, Y.; Xu, H.; Kong, D.; Yang, Z.; Yang, C.; Ding, D. Self-Assembling Peptide of D-Amino Acids Boosts Selectivity and Antitumor Efficacy of 10-hydroxycamptothecin. *ACS Appl. Mater. Interfaces* **2014**, *6*, 5558−5565.
(43) Tripathi, A.; Bankaitis, V. A. Molecular Docking: from Lock and Key to Combination Lock. *J. Mol. Med. Clin. Appl.* **2017**, *2*, 10.16966/2575-0305.106.

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102

(44) Jorgensen, W. L. Rusting of the Lock and Key Model for Protein-Ligand Binding. *Science* **1991**, *254*, 954−955.
(45) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(46) MacKerell, A. D., Jr; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(47) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950−1958.

https://doi.org/10.1021/acsomega.5c06570
ACS Omega 2025, 10, 50088−50102