JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

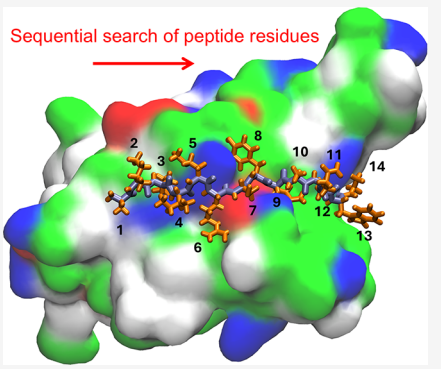
On-the-Fly Sequential Design of Simple Peptides

Francesco Coppola and Petr Král*

Cite This: J. Chem. Inf. Model. 2025, 65, 9738−9746 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We introduce a simple and efficient computational methodology capable of designing peptides lacking higher order structures. The method is based on the sequential modification of residues in a simple peptide attached to a substrate of interest. The peptide design can start at an arbitrary point of the substrate and proceed into a direction chosen spontaneously or guided by an external potential. The decision about using a certain residue is based on its binding free energy to the substrate, as evaluated in molecular dynamics simulations. Here, this approach is tested on the design of peptides binding to spike proteins in SARS-CoV-2. The methodology can be easily modified according to actual needs and extended to other molecules.



Peptides, peptide-based hybrids, and their supramolecular constructs are becoming a popular alternative to small molecule therapeutics.[1,2] Therapeutic peptides could be relatively simple, despite being highly selective and biocompatible, with minimal side effects.[3−10] Typically, peptide-based inhibitors have been designed by extracting and modifying fragments from known protein binders,[11−15] by screening of sequence motifs based on phage display experiments,[16] and lately by various *in silico* methods.[17−19]

Despite a continuous progress, it can be relatively challenging to design from scratch simple peptides binding to specific receptors.[20] The phase space of peptides grows exponentially with their lengths, and peptide characteristics are correlated among different peptide residues and segments. These complex issues have been addressed by advanced computational approaches based on machine learning methods.[21−25] In these methods, which are mostly focused on proteins, the residues are searched and optimized *simultaneously* based on a thorough experience gained by training the machines on realistic systems.[26]

Short peptides lacking higher order structures might be sufficient in many therapeutic applications. However, the question is how to design such peptides and hybrid molecules based on them. Natural evolution of proteins often proceeded in a gradual manner, where individual regions evolved more or less sequentially, sometimes very fast, as in the search of novel antibodies (somatic hypermutation).[27,28] In analogy, we could try to develop computational methods capable of designing simple peptides in a sequential manner. Therefore, at any moment, the residues would be searched in a much smaller phase space region, which could dramatically accelerate the speed of searching. Importantly, residue−residue coupling should be limited in such peptides lacking higher order structures. Although, the residues would remain loosely connected through the peptide backbone and locally act like multivalent binders, their cooperative effect on the peptide characteristics would be limited.

RESULTS AND DISCUSSIONS

Here, we introduce *sequential design of peptide* (SDP) methodology based on the stepwise optimization of peptide residues. Using classical atomistic molecular dynamics (MD) simulations in explicit solvents, each residue would be optimized by maximizing its binding energy to the substrate, while effectively (mean-field) including coupling of other residues to the substrate. We can develop different SDP methods, depending on the peptide binding position, its length, required affinity, and other criteria, such as good solvation. Here, we present two SDP methods (Figure 1) and demonstrate how they can design peptides in two areas of the RBD (top and side regions of spike), as shown in Figure S1.

SDP1 Method. In the first variant, called an SDP1 method, the peptide to be designed has initially a fixed length and only alanine residues. This template is placed above the substrate of interest, partially fixed at its two termini, and sequentially modified. Here, the SDP1 method (Figure 1 − top) is tested by designing 12 residues long (14 at start) peptides separately binding to two (top and side) regions on the receptor binding

Received: June 12, 2025
Revised: September 3, 2025
Accepted: September 4, 2025
Published: September 11, 2025

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738−9746

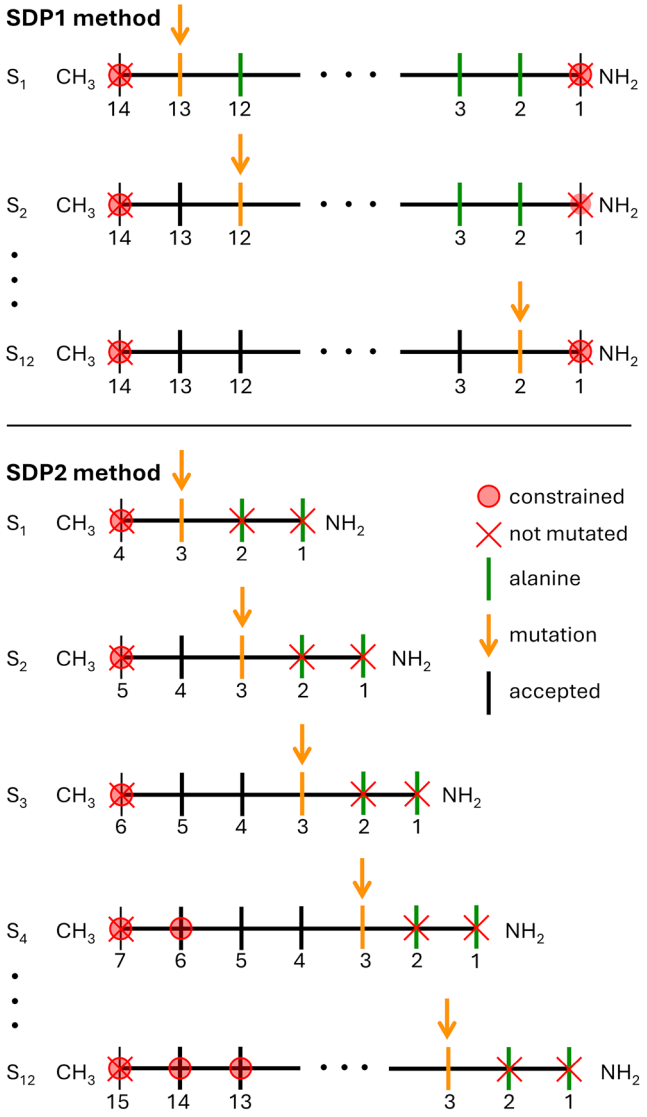


Figure 1. Schematic of two SDP methods with individual S_i steps described in the legend. The NH_2 and CH_3 termini are neutralized during the peptide design.

domain (RBD) of the Omicron (PDB 7WVN) spike protein of SARS-CoV-2.

During the design, highlighted in Figure 1, the two terminal (charged) groups of the peptide at residues 1 and 14, H_3N^+-R and R-COO^-, are converted into neutral groups, NH_2-R (amine) and R-CH_3 (methyl), respectively, called the modified N-terminus and C-terminus. This protection was introduced to avoid strong Coulombic coupling at the termini during the initial search of the residues. During the search, the terminal residues (1 and 14) are held close ($\approx$ 8 Å) to the chosen positions at the target protein by harmonic potentials of $U(r) = k\,(r - r_{ref})^2$, where r and r_{ref} are vectors of the residue and the reference point, respectively, and k = 0.5 kcal/mol. In the top position, residue 14 (at the C terminus) was placed between the Y446 and R490 residues of the RBD, and residue 1 (at the N terminus) was placed between the Q403 and R405 residues of the viral protein. In the side position, K455 and S456 (K353 and R354) are the related chosen pairs of residues. In Figure S1A–D, we show the relative position of the peptide with respect to that of the protein. If necessary, the design conditions can be easily modified (peptide position, length, stretching, etc.).

The remaining 12 residues are sequentially mutated (13 → 2), where each residue is optimized by gradually simulating all amino acids at the position of this residue, each for 5 ns, while holding only the two terminal residues (Figure 1). The MD simulations are performed in an explicit physiological solution (0.15 M NaCl) at standard conditions within an NpT ensemble (Methods). The trajectory obtained for each tested amino acid is used to calculate the free energy of binding (MMGB-SA), ΔG, of the whole peptide (fixed termini) to the receptor (Methods). After testing all 20 amino acids at the chosen residue, the method accepts the amino acid that provides the strongest binding of the peptide to the receptor. Then, the next residue is searched, etc. Upon completion of this run, the peptide is gradually reoptimized in the backward direction (2 → 13) to further improve the residue optimization. Then, residues 1 and 14 are replaced by regular (charged) termini, and the peptide residues are renumbered (1–12). The optimized peptide is released and simulated for 100 ns to test its overall stability. Given the fact that only a very short segment of the peptide (not considering unbound ALA residues) is free during the simulations where a residue is chosen, we believe that 5 ns could be sufficient to capture the binding possibilities around the tested residue; peptides were even redesigned by combined MD simulations with Monte Carlo methods,[15] where each new amino acid was only simulated for 1 ns.

Figure 2 shows snapshots of the SDP1-designed 12-residues peptides binding to the top (Figure 2A) and side (Figure 2F) positions of the RBD. Some details of this binding are highlighted on the side (Figure 2B–E, G–J) and discussed later. Each residue is denoted by one letter code name and position (number) in the sequence (I1, E2, and ···).

To visualize how the peptides change during optimization, Figure 3A,B shows the gradually increasing binding free energies, ΔG, between the SDP1-designed peptides and the two RBD (spike) sites (Methods). The residues obtained are shown in the two figures. One can see that, practically, all residues accepted in the forward run (1) are increasing $|\Delta G|$. In the top position (Figure 3A), one residue is optimized in backward run (2), while none is accepted in the side position (Figure 3B). The two peptides designed for the top and side positions eventually gain large binding free energies of ΔG = −66.68 and −65.03 kcal/mol, respectively. Despite the fact that these peptides do not have higher order structures, they reached similar binding energies like peptides with higher order structures which were extracted from ACE2.[14,15]

Next, we examine separate contributions of the 12 peptide residues to the total binding energy of the freely simulated peptide (charged termini). The average direct (electrostatic and van der Waals) binding energies between the peptide and the spike, averaged over the last 20 ns (100 ns trajectory), are split into contributions of individual residues (side chain and nearest groups of the peptidic backbone), as shown in Figure 3C–D (Methods).

Figure 2A–F and 2G–L illustrates that the residues designed closely follow the nature of the protein surface to which they bind. For example, in the top position (Figure 2A), the 4–8 residues are surrounding charged residues of the RBD (spike) and account for 77% of the total binding energy. In particular, the positive R5 residue has good binding to the negative D402 residue (Figure 2B). The polar C6 residue

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738–9746

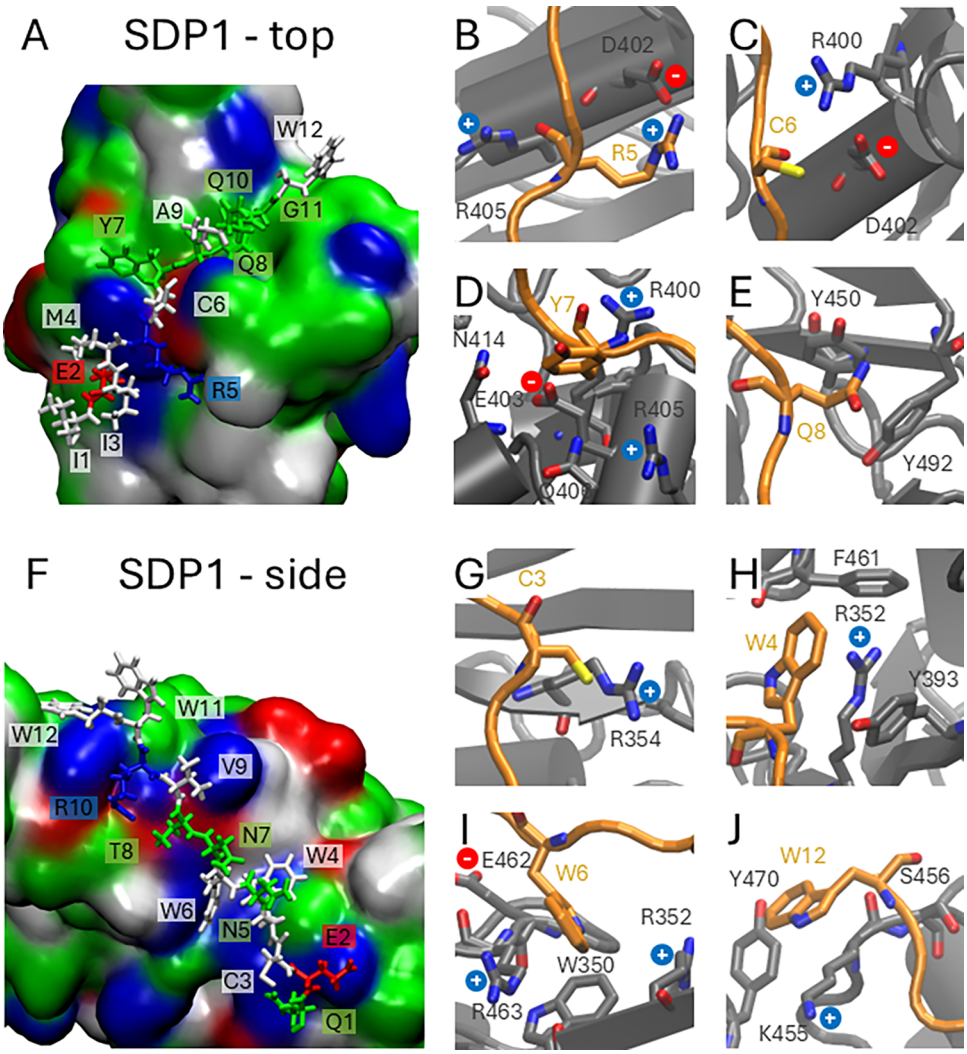


Figure 2. Conformations of the 12-residues peptides designed by the SDP1 method for binding to (A) top and (F) side regions of RBD (spike) in SARS-CoV-2 (Omicron) after 100 ns of simulations. Details of this binding are in (B–E) and (G–J), respectively.

binds with the positive R400 and negative D402 residues (Figure 2C). The polar Y7 and Q8 residues have the highest binding energies. The polar Y7 residue binds to the negative E403, polar Q406 and N414, and charged R400 and R405 residues (Figure 2D). The polar Q8 residue fits into a polar pocket formed at the top, where it develops multipolar interactions with the polar Y450 and Y492 residues (Figure 2E).

In the side position (Figure 2F), the polar C3 residue couples with the positive R354 residue (Figure 2G), while the hydrophobic W4 residue (20% of the total binding energy) couples to the positive R352, polar Y393, and hydrophobic F461 residues (Figure 2H). The hydrophobic W6 residue (26% of the total binding energy) couples with the positive R463 and R352, polar E462, and hydrophobic W350 residues (Figure 2I). The hydrophobic W12 residue (accounting for 18% of the binding energy) binds to the positive K456 and polar Y470 residues (Figure 2J).

We also examined the dynamics of these peptides when they are freely released in their binding positions. In 100 ns simulations, we calculate their root-mean-square deviations (RMSD), starting from their initial configurations obtained in the optimization. Figure 4A reveals that peptide backbones have stable configurations in both the top and side positions. To gain a better idea about the stability of individual residues, Figure 4B,C shows the relative motion of each residue in the top and side positions, respectively. The top-peptide residues (Figure 4B) show sizable motion in the central part, despite the fact that those residues bind pretty well (Figure 3C). The two termini, especially the hydrophobic W12 residue, show more flexibility. The side-peptide residues (Figure 4C) appear more stable, especially the 4–10 residues, despite the fact that only the 4 and 6 residues had large binding energies (Figure 3D). Apparently, the 12 residues (together with 4 and 6) help to stabilize the whole central region. However, the 1–3 residues are less stable in their binding positions.

Finally, in Figure S2A,B, we present the Surface Accessible Solvent Area (SASA) of the peptides during their binding to the RBD (spike). The results are obtained during the 100 ns simulations performed with a probe of radius 1.4 Å. The data show that the peptides bind to the RBD within 28–35% of their areas, and for each system this fraction fluctuates very little. Overall, we can see that the SDP1 method provides reasonably stable peptides, especially in the flatter side position.

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738–9746

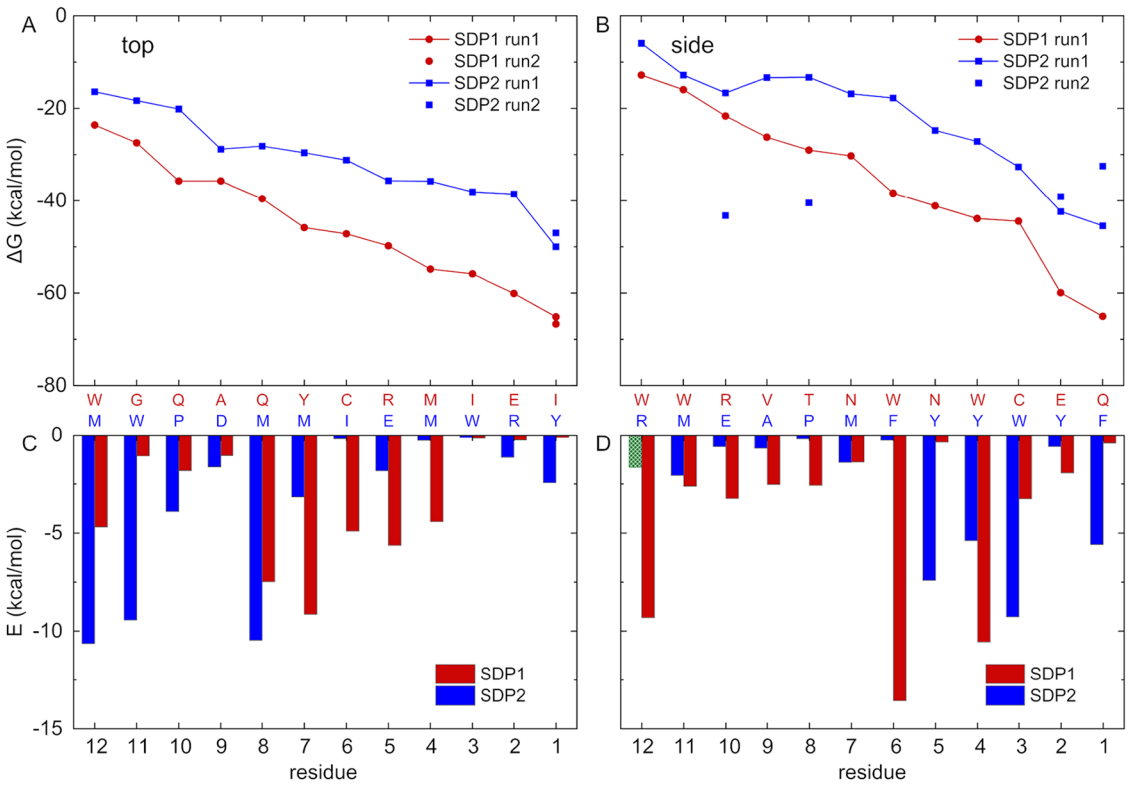


Figure 3. Peptides evolving with the SDP1 and SDP2 methods on the top (A) and side (B) positions of the Omicron SARS-CoV-2 variant (red, SDP1, blue - SDP2). Not connected by a line, we show, also in the same color, the second run optimization. In panels (C) and (D) are the averaged nonbonded energies for each residue interacting with the viral protein during 100 ns simulation with no constrains. The calculations are averaged over the last 20 ns of simulations. In panel (D), the green bar (W12 residue) is a positive value. The final sequences are shown between the plots.

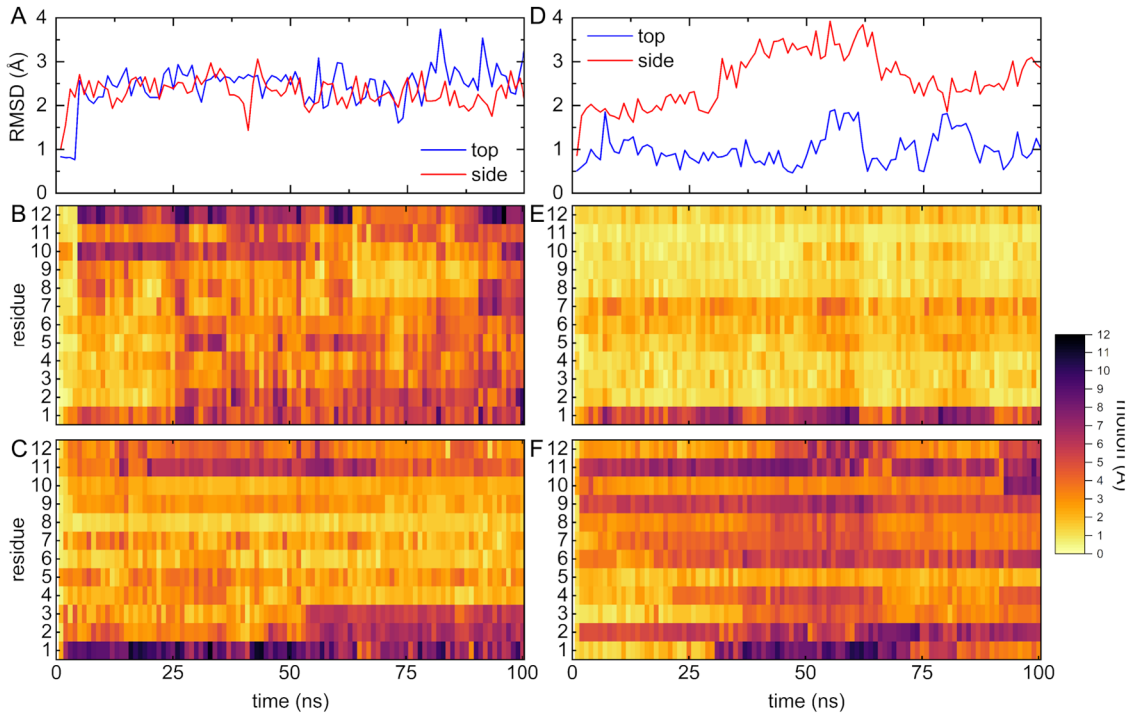


Figure 4. RMSD calculations of peptides designed by the (A−C) SDP1 and (D-F) SDP2 methods. (A) Peptide backbone designed on the top (blue) and side (red) positions on the RBD (spike) in 100 ns simulations. (B, C) RMSD calculations done for each residue (motion represented by colors). (D−F) The same was true for the SDP2 method.

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738−9746

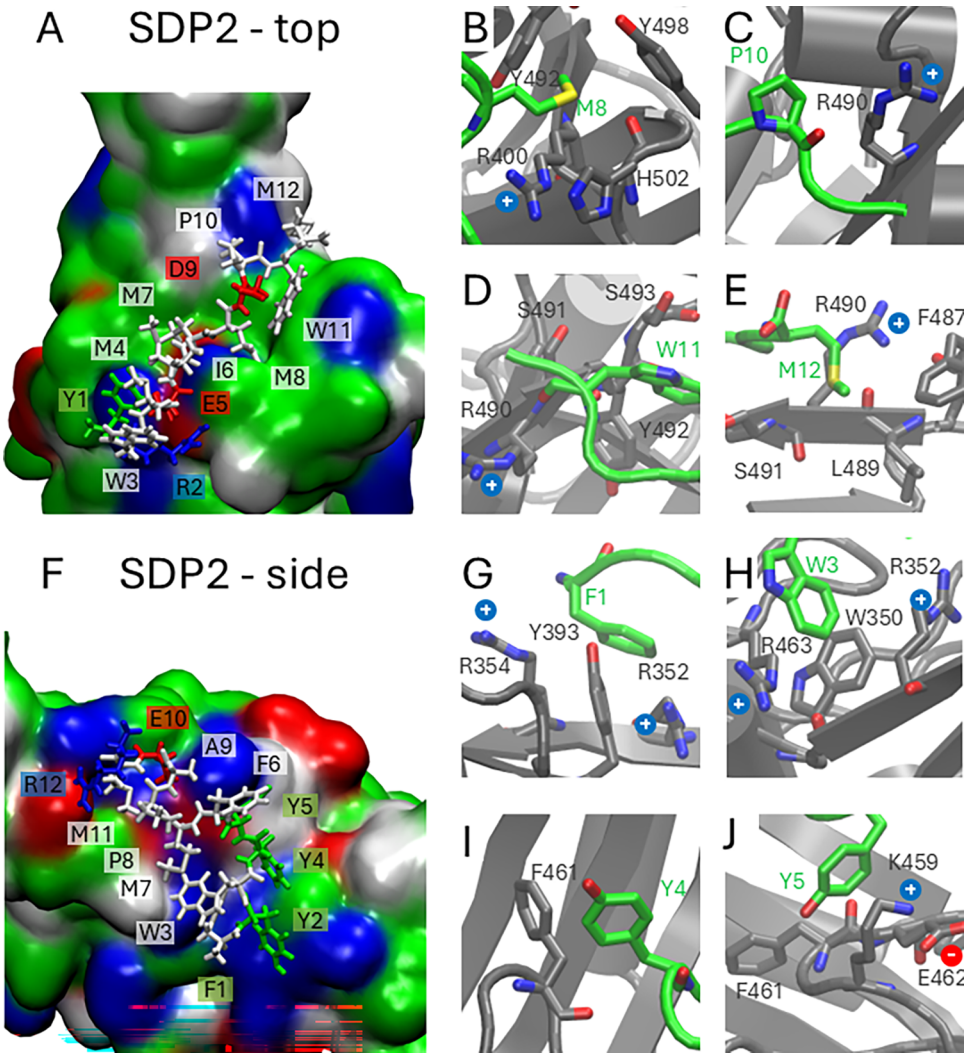


Figure 5. Conformations of the 12-residues peptides designed by the SDP2 method for binding to the (A) top and (F) side of the RBD (spike) in SARS-CoV-2 (Omicron), after 100 ns of simulations. Details of this binding are in (B–E) and (G–J), respectively.

SDP2 Method. Next, we discuss the second SDP approach, called an SDP2 method (Figure 1 – bottom). This method is nearly *translationally invariant* in the sense that it provides each residue almost the same designing conditions. In the SDP2 method, the peptide grows from an initial point on the receptor, while being assisted by an external guiding field. Here, the SDP2 method is tested by designing 12 residues long peptides binding to the above two (top and side) regions of the RBD (spike).

The design starts from a short peptide with just 4 alanine residues, as shown in Figure 1 (bottom). Its C-terminal residue (position 4) is held in a harmonic potential (k = 1 kcal/mol) above the receptor in the same position as in the SDP1 method, and it is not mutated. We thought about various options of terminating the peptide during its design. Eventually, we used just 4 ALA residues. Since these ALA residues are removed in the end, they should not affect much the output. The N-terminal residue and the next residue (positions 1 and 2) are free, and they are also not mutated. The residue in between (position 3) is free, mutated, and optimized. The simulations are done like in the SDP1 method, including neutralization of both termini, but here, only the C-terminus is fixed. The peptide-receptor binding free energies, ΔG, are again calculated for the whole (growing) peptide.

After the selection of one amino acid at position 3, a new alanine residue is added at the N-terminal end and is used as the new N-terminal residue (position 1), while all residues are renumbered (shifted by 1), as seen in Figure 1 (bottom). Then, the residue at the new position 3 is again optimized, while only the C-terminal residue (position 5) is fixed. In this manner, the peptide gradually continues to grow, and this growth can be performed in the presence of a guiding potential (Methods). When the mutated residue 3 becomes separated by more than 2 residues from the fixed C-terminal residue (Figure 1), the 6, 7, and further residues are also held fixed around their positions obtained in the previously designed residue. This arrangement is also held during the simulations and calculations of ΔG. Once the peptide reaches 15 residues, where 12 residues are optimized, the 1, 2, and 15 residues are removed, and the peptide ends are replaced by regular (charged) termini. Then, the 12-residues peptide undergoes backward screening, while only the two termini are fixed by external potentials (as in the SDP1 method), except when they are redesigned themselves (the last 3 residues are free). Finally, the designed peptide is released and simulated for 100 ns.

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738–9746

Snapshots of the designed peptides are visualized in the top (Figure 5A) and side (Figure 5F) positions (details in Figure 5B−E, G−J), in analogy to Figure 2.

Figure 3A,B shows again the changing binding free energies, ΔG, between the SDP2-designed peptides (current length of growing peptide) and the two sites in RBD (Methods). In backward run (2), only one residue is optimized in the top position (Figure 3A), while four residues are accepted in the side position (Figure 3B). The two peptides designed for the top and side positions gained relatively large binding free energies of $\Delta G = -47.65$ and -43.24 kcal/mol, respectively, although smaller than in the SDP1 method.

Figure 3C-D reveals again how much individual residues contribute to the overall peptide binding energy. In the top position (Figure 5A), the 7−12 residues account for 86% of the total binding energy. The nonpolar M8 residue binds to the positive R400 and H502 and polar Y492 and Y498 residues (Figure 5B). The nonpolar P10 residue binds to the positive R490 residue (Figure 5C). The hydrophobic W11 residue binds to polar S491, Y492, and S493, and positive R490 residues (Figure 5D). The terminal M12 residue binds to positive R490, polar S491, nonpolar L489, and hydrophobic F487 residues (Figure 5E).

In the side position (Figure 5F), the first 5 residues account for 88% of the total binding energy. The hydrophobic F1 residue binds to the positive R352 and R354, and the polar Y393 residues (Figure 5G). The hydrophobic W3 residue, shown also in (Figure 2I), binds to the positive R352, R465, and hydrophobic W350 residues (Figure 5H). The polar Y4 residue binds to the hydrophobic F461 residue (Figure 5I). The polar Y5 residue binds to positive K459, negative E462, and hydrophobic F461 residues (Figure 5J).

Finally, we again examine the free dynamics of these peptides above their binding positions. Figure 4D presents the RMSD of the peptide backbones, obtained in 100 ns simulations, confirming their stable configurations in both positions. Figure 4E,F shows the RMSD of each residue in the top and side positions, respectively. The residues in the top peptide (Figure 4E) are practically not moving except for the 1 residue (repelling in Figure 3D). However, the residues in the side peptide (Figure 4F) are significantly less stable, especially the 2, 6, 9, and 11 residues. SASA (Figure S2) again shows that binding accounts for 25−35% of the surface of each peptide. Overall, the SDP2 method provided relatively stable peptides in both positions, but the stability seems better in the top position, where the substrate is more curved.

FURTHER DISCUSSIONS

The above results show that the SDP1 and SPD2 methods, which operate under different conditions, could give rather different peptides on the same substrate. In the SDP1 method, except for the first and last residues, the peptide freely decides the path to follow according to the binding energy of the residues. In the SDP2 method, a user-defined potential guides the peptide's direction of motion on the substrate, where the residues are searched. In this way, the user can control the peptide design. Although the control potential added in the SDP2 method might destabilize the peptide in certain regions, once it is released, it can connect several regions of the substrate, which the peptide would not have a chance to find in the SDP1 method.

The length of peptide designed by the discussed methodology is not inherently limited, but the peptide should not significantly self-interact on the substrate, which could generate higher order structures. Realistically, peptides with fewer than 20 residues should be able to cover the substrate. When these peptides are freely solvated, they might form random coils without stable configurations that can be stabilized on the substrate by loosely binding to its charged residues. Nonpolar residues also tend to contribute a lot to the peptide binding, but their numbers should be limited to preserve its solubility in water. The binding peptide should not be too stretched on the substrate because entropic forces might destabilize it.

The presented methodology is highly versatile, and it could be executed in different versions depending on the peptide design strategy, position, fixation, direction of growth, simultaneous optimization of multiple residues (beyond 1-residue mean-field approach), branching, etc. For example, we have designed cyclic peptides that needed to be optimized around their circumference. The methodology can also use additional sources of information, such as experimental data, and deliberately implement them during peptide design. An interesting possibility would be to design peptides with non-natural residues or entirely different polymers, which might also be also branched. In such cases, the basis sets of possible residues would be different, and the backbone could be, possibly too. For example, we have designed hybrid molecules where peptides are connected to other molecular components (lipids, sugars, ...). The designed peptides can be further screened with advanced (AI) computational methods pretrained for these purposes.

CONCLUSIONS

We have developed a simple computational methodology, based on molecular dynamics simulations, that can design peptides lacking higher order structures capable of binding to protein receptors. The methodology is easily accessible, transparent, efficient, and able to provide many parameters of the designed peptides. The progress of the designed peptide can be controlled at any step, where the binding ability of the designed peptide is accessible. The proposed methodology could provide high-quality peptides for medical, pharmaceutical, and material applications, and be executed to design other types of molecules and nanomaterials.[29,30]

METHODS

Here, we present the theoretical basis of the SDP methods, the implemented numerical procedures, and other computational details. The systems were simulated by NAMD3[31,32] and the CHARMM36 protein force field.[33] The PME method was used for the evaluation of long-range Coulombic interactions.[34] The time step was set to 2 fs. The simulations were performed in the NpT ensemble (p = 1 bar and T = 310 K), using Langevin dynamics with a damping constant of 1 ps^{-1}. After energy minimization, the peptides were fixed in a harmonic potential of $U(r) = k\,(r - r_{ref})^2$ (r and r_{ref} are vectors) with different spring constants k, and the whole systems were equilibrated. All mutated systems were simulated for 2 ns to pre-equilibrate ions and for 3 ns to equilibrate the systems with the tested amino acid. The mutated residue in the SDP2 method is guided by a weak harmonic potential of $U(r) = k\,(r - r_{ref})^2$ (k = 0.01−0.30), which is separately applied to the N- and C-atoms of the backbone of this residue. The r_{ref} reference positions of these atoms are chosen in such a way that they provided guidance to the growing peptide. Upon designing the

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738−9746

peptide, the guiding potential is removed, and the peptide is freely simulated.

Molecular mechanics with a generalized Born model and surface area solvation (MMGB-SA) calculations have been performed to calculate binding free energies between peptides and the spike (Figure 3A,B).[35−38] The MMGB-SA free energies were calculated from the following terms averaged over a system trajectory,

$$G_{tot} = E_{MM} + G_{solv-p} + G_{solv-np} - T\Delta S_{conf} \quad (1)$$

Here, E_{MM} is the sum of bonded, van der Waals, and Coulombic energies, G_{solv-p} is a polar contribution to the solvation energy, $G_{solv-np}$ is a nonpolar contribution, and $T\Delta S_{conf}$ is an entropic term (neglected here). NAMD3 was used to calculate the first 3 terms of eq 1 in implicit water with a dielectric constant of $\varepsilon = 78.4$. The G_{solv-p} term was obtained using the generalized Born (GB) model. The $G_{solv-np}$ term for each system configuration was calculated as a linear function of the solvent-accessible surface area (SASA), determined using a probe radius of 1.4 Å, as $G_{solv-np} = \gamma$ SASA, where $\gamma = 0.00542$ kcal mol^{-1} Å^{-2} is the surface tension. The approximate free energies of binding for the studied complexes were obtained from three separate MMGB-SA calculations (peptide, spike, and the whole complex), which were isolated from the explicit solvent MD simulation

$$\Delta G_{MMGB-SA} = G_{tot}(\text{peptide} - \text{spike}) - G_{tot}(\text{peptide}) - G_{tot}(\text{spike}) \quad (2)$$

To evaluate the free energies, other methods and procedures could be implemented as well.

The MMGB-SA method is less precise than other (more explicit) free energy methods, such as umbrella sampling or related free energy perturbation (FEP) methods, but it is relatively easy to implement computationally, which is a necessary condition when many free energy calculations must be performed. The precision of MMGB-SA should not be a critical issue here given the limitations used in the choice of specific residues, where the peptide is largely frozen during the energy evaluation. When computational resources are sufficient, more precise methods could be used to redesign certain segments of the peptides or compare a limited number of peptides obtained by different means. The peptide design methodology is relatively cost-effective, and the costs grow linearly with the peptide length. By using a modern GPU processing unit on a single node, we can optimize one residue in 12 h, when each of the 20 tested amino acids is simulated for 5 ns at the modified residue site.

The calculation of (direct) binding energies was done with the NAMD Energy Plugin (version 1.4) and VMD1.9.4,[39] where the last 20 ns of each 100 ns long simulation (2000 frames) were used to calculate nonbonded energies using the NAMD energy tool. The nonbonded energies (Figure 3C,D) have been calculated as the sum of the electrostatic and vdW attractive (and short-range repulsive) contributions. The electrostatic contribution was given by

$$U_{elec} = \sum_{i=1}^{n}\sum_{j>i}^{n} \frac{1}{4\pi\varepsilon}\frac{q_i q_j}{|\vec{r}_i - \vec{r}_j|} \quad (3)$$

where $|\vec{r}_i - \vec{r}_j|$ is a distance between the charges, q_i and q_j, and $\varepsilon = 78.4$; pairwise interaction calculations are not performed beyond a cutoff distance.

The Lennard-Jones (LJ) 6−12 potential energies were used to describe the vdW attractions and close-distance atomic repulsions

$$U_{LJ} = \sum_{i=1}^{n}\sum_{j>i}^{n} \varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \quad (4)$$

where ε_{ij} is the maximum stabilization energy for the ith and the jth atoms, σ_{ij} is the distance between ith and jth atoms at the minimum of the potential, and r_{ij} is the actual distance between the two atoms. The LJ parameters between different atom types were calculated from $\sigma_{ij} = (\sigma_{ii} + \sigma_{jj})/2$ and $\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}$ (Lorentz−Berthelot rules).

■ ASSOCIATED CONTENT

Data Availability Statement

NAMD and VMD are distributed free of charge. The final peptide structures (.psf and .pdb) for each position, together with the configuration files and their parameter files, have been made available at https://github.com/PetrKral-group/On-the-fly-sequential-design-of-simple-therapeutic-peptides. Additionally, the files for the constrained referenced atoms have been included for each simulation.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.5c01340.

Placement of the 12-residue peptides designed by the SDP1 method on the (A) top and (B) side of the RBD in spike protein of SARS-CoV-2 (Omicron). (C) and (D) show the same for peptides designed by the SDP2 method (Figure S1); SASA for peptides designed by the SDP1 method and docked on the (A) top and (B) side of the RBD in spike protein of SARS-CoV-2 (Omicron). (C) and (D) show the same forpeptides designed by the SDP2 method. (red) SASA for peptides docked on spike, (green) SASA for peptides removed from spike (the same trajectory) (Figure S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Petr Král − *Department of Chemistry, University of Illinois Chicago, Chicago, Illinois 60607, United States; Departments of Physics, Pharmaceutical Sciences, and Chemical Engineering, University of Illinois Chicago, Chicago, Illinois 60607, United States;* orcid.org/0000-0003-2992-9027; Email: pkral@uic.edu

Author

Francesco Coppola − *Department of Chemistry, University of Illinois Chicago, Chicago, Illinois 60607, United States;* orcid.org/0000-0002-2429-204X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.5c01340

Author Contributions

F.C. has performed the simulations, prepared the figures, and wrote the manuscript. P.K. has designed the methodology, guided the research, and wrote the manuscript.

Notes

The authors declare no competing financial interest.

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738−9746

■ ACKNOWLEDGMENTS

P.K. acknowledges the support from NSF DMR 2212123.

■ REFERENCES

(1) Vanhee, P.; van der Sloot, A. M.; Verschueren, E.; Serrano, L.; Rousseau, F.; Schymkowitz, J. Computational design of peptide ligands. *Trends Biotechnol.* **2011**, *29*, 231−239.
(2) Wang, L.; Wang, N.; Zhang, W.; Cheng, X.; Yan, Z.; Shao, G.; Wang, X.; Wang, R.; Fu, C. Therapeutic peptides: current applications and future directions. *Signal Transduction Targeted Ther.* **2022**, *7*, No. 48.
(3) Wells, J. A.; McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450*, 1001−1009.
(4) Tsomaia, N. Peptide therapeutics: Targeting the undruggable space. *Eur. J. Med. Chem.* **2015**, *94*, 459−470.
(5) Fosgerau, K.; Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discovery Today* **2015**, *20*, 122−128.
(6) Russo, A.; Aiello, C.; Grieco, P.; Marasco, D. Targeting "Undruggable" Proteins: Design of Synthetic Cyclopeptides. *Curr. Med. Chem.* **2016**, *23*, 748−762.
(7) Xu, F.; Yuan, Y.; Wang, Y.; Yin, Q. Emerging peptide-based nanovaccines: From design synthesis to defense against cancer and infection. *Biomed. Pharmacother.* **2023**, *158*, No. 114117.
(8) Sharma, K.; Sharma, K. K.; Sharma, A.; Jain, R. Peptide-based drug discovery: Current status and recent advances. *Drug Discovery Today* **2023**, *28*, No. 103464.
(9) Subramanian, S.; Jain, M.; Misra, R.; Jain, R. Peptide-based therapeutics targeting genetic disorders. *Drug Discovery Today* **2024**, *29*, No. 104209.
(10) Jalil, A. T.; Abdulhadi, M. A.; Al-Ameer, L. R.; Taher, W. M.; Abdulameer, S. J.; Abosaooda, M.; Fadhil, A. A. Peptide-Based Therapeutics in Cancer Therapy. *Mol. Biotechnol.* **2024**, *66*, 2679−2696.
(11) Watt, P. M. Screening for peptide drugs from the natural repertoire of biodiverse protein folds. *Nat. Biotechnol.* **2006**, *24*, 177−183.
(12) Petsalaki, E.; Stark, A.; García-Urdiales, E.; Russell, R. B. Accurate prediction of peptide binding sites on protein surfaces. *PLoS Comput. Biol.* **2009**, *5*, No. e1000335.
(13) Jeong, W.-j.; Bu, J.; Han, Y.; Drelich, A. J.; Nair, A.; Král, P.; Hong, S. Nanoparticle Conjugation Stabilizes and Multimerizes β-Hairpin Peptides To Effectively Target PD-1/PD-L1 β-Sheet-Rich Interfaces. *J. Am. Chem. Soc.* **2020**, *142*, 1832.
(14) Han, Y.; Král, P. Computational Design of ACE2-Based Peptide Inhibitors of SARS-CoV-2. *ACS Nano* **2020**, *14*, 5143−5147.
(15) Chaturvedi, P.; Han, Y.; Král, P.; Vuković, L. Adaptive Evolution of Peptide Inhibitors for Mutating SARS-CoV-2. *Adv. Theory Simul.* **2020**, *3*, No. 2000156.
(16) Giordano, R. J.; Cardó-Vila, M.; Salameh, A.; Anobom, C. D.; Zeitlin, B. D.; Hawke, D. H.; Valente, A. P.; Almeida, F. C. L.; Nör, J. E.; Sidman, R. L.; Pasqualini, R.; Arap, W. From combinatorial peptide selection to drug prototype (I): targeting the vascular endothelial growth factor receptor pathway. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 5112−5117.
(17) Lin, E.; Lin, C.; Lane, H. De Novo Peptide and Protein Design Using Generative Adversarial Networks: An Update. *J. Chem. Inf. Model.* **2022**, *62*, 761−774.
(18) Xia, Y.; Zhao, K.; Liu, D.; Zhou, X.; Zhang, G. Multi-domain and complex protein structure prediction using inter-domain interactions from deep learning. *Commun. Biol.* **2023**, *6*, No. 1221.
(19) Banhos Danneskiold-Samsøe, N.; Kavi, D.; Jude, K. M.; Nissen, S. B.; Wat, L. W.; Coassolo, L.; Zhao, M.; Santana-Oikawa, G. A.; Broido, B. B.; Garcia, K. C.; Svensson, K. J. AlphaFold2 enables accurate deorphanization of ligands to single-pass receptors. *Cell Syst.* **2024**, *15*, 1046−1060.
(20) Gupta, S.; Azadvari, N.; Hosseinzadeh, P. Design of Protein Segments and Peptides for Binding to Protein Targets. *BioDesign Res.* **2022**, *2022*, No. 9783197.
(21) Goles, M.; Daza, A.; Cabas-Mora, G.; Sarmiento-Varón, L.; Sepúlveda-Yañez, J.; Anvari-Kazemabad, H.; Davari, M. D.; Uribe-Paredes, R.; Olivera-Nappa, Á.; Navarrete, M. A.; Medina-Ortiz, D. Peptide-based drug discovery through artificial intelligence: towards an autonomous design of therapeutic peptides. *Briefings Bioinf.* **2025**, *25*, No. bbae275.
(22) Chen, S.; Lin, T.; Basu, R.; Ritchey, J.; Wang, S.; Luo, Y.; Li, X.; Pei, D.; Kara, L. B.; Cheng, X. Design of target specific peptide inhibitors using generative deep learning and molecular dynamics simulations. *Nat. Commun.* **2024**, *15*, No. 1611.
(23) Xie, X.; Valiente, P. A.; Kim, J.; Kim, P. M. HelixDiff, a Score-Based Diffusion Model for Generating All-Atom α-Helical Structures. *ACS Cent. Sci.* **2024**, *10*, 1001−1011.
(24) Raveh, B.; London, N.; Zimmerman, L.; Schueler-Furman, O. Rosetta FlexPepDock ab-initio: Simultaneous Folding, Docking and Refinement of Peptides onto Their Receptors. *PLoS One* **2011**, *6*, No. e18934.
(25) Strokach, A.; Kim, P. M. Deep generative modeling for protein design. *Curr. Opin. Struct. Biol.* **2022**, *72*, 226−236.
(26) Wu, K.; Bai, H.; Chang, Y.; Redler, R.; Mcnally, K. E.; Sheffler, W.; Brunette, T. J.; Hicks, D. R.; Morgan, T. E.; Stevens, T. J.; Broerman, A.; Goreshnik, I.; Dewitt, M.; Chow, C. M.; Shen, Y.; Stewart, L.; Derivery, E.; Silva, D. A.; Bhabha, G.; Ekiert, D. C.; Baker, D. De novo design of modular peptide-binding proteins by superhelical matching. *Nature* **2023**, *616*, 581−589.
(27) Groff, K.; Brown, J.; Clippinger, A. J. Modern affinity reagents: Recombinant antibodies and aptamers. *Biotechnol. Adv.* **2015**, *33*, 1787−1798.
(28) Martin, A.; Chahwan, R.; Parsa, J. Y.; Scharff, M. D. Chapter 20 - Somatic Hypermutation: The Molecular Mechanisms Underlying the Production of Effective High-Affinity Antibodies. In *Molecular Biology of B Cells*, Second ed.; Alt, F. W.; Honjo, T.; Radbruch, A.; Reth, M., Eds.; Academic Press: London, 2015; pp 363−388.
(29) Wang, B.; Král, P. Optimal Atomistic Modifications of Material Surfaces: Design of Selective Nesting Sites for Biomolecules. *Small* **2007**, *3*, 580−584.
(30) Russell, J.; Král, P. Configuration-sensitive molecular sensing on doped graphene sheets. *Nano Res.* **2010**, *3*, 472−480.
(31) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, No. 044130.
(32) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(33) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. D.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamcis Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(34) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N.log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(35) Ylilauri, M.; Pentikäinen, O. T. MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. *J. Chem. Inf. Model.* **2013**, *53*, 2626−2633.
(36) Rico, M. I.; Lebedenko, C. G.; Mitchell, S. M.; Banerjee, I. A. Molecular dynamics simulations, docking and MMGBSA studies of newly designed peptide-conjugated glucosyloxy stilbene derivatives with tumor cell receptors. *Mol. Diversity* **2022**, *26*, 2717−2743.
(37) Jupudi, S.; Rajagopal, K.; Murugesan, S.; Kumar, B. K.; Raman, K.; Byran, G.; Chennaiah, J.; Muthiah, V. P.; Dasan P, B.; Sankaran, S. Identification of Papain-Like Protease inhibitors of SARS CoV-2

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738−9746

through HTVS, Molecular docking, MMGBSA and Molecular dynamics approach. *S. Afr. J. Bot.* **2022**, *151*, 82−91.
(38) Guimarães, C. R. W.; Mathiowetz, A. M. Addressing Limitations with the MM-GB/SA Scoring Procedure using the WaterMap Method and Free Energy Perturbation Calculations. *J. Chem. Inf. Model.* **2010**, *50*, 547−559.
(39) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.

https://doi.org/10.1021/acs.jcim.5c01340
J. Chem. Inf. Model. 2025, 65, 9738−9746