JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

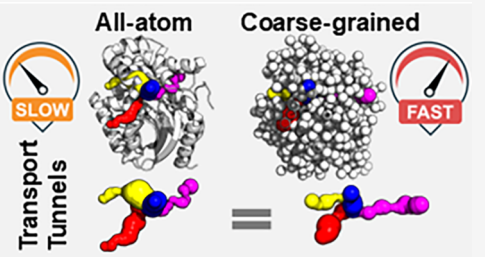
Unlocking High-Throughput Investigation of Transport Tunnels in Enzymes Using Coarse-Grained Simulation Methods

Nishita Mandal, Jan A. Stevens, Adolfo B. Poma, Bartlomiej Surpeta, Carlos Sequeiros-Borja, Aravind Selvaram Thirunavukarasu, Siewert J. Marrink, and Jan Brezovsky*

Cite This: J. Chem. Theory Comput. 2026, 22, 135−150 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Transport tunnels in enzymes with buried active sites are critical gatekeepers of enzymatic function, controlling substrate access, product release, and catalytic efficiency. Despite their importance, the transient nature of these tunnels makes them difficult to study using conventional simulation methods. In this study, we systematically evaluate three coarse-grained (CG) molecular dynamics approaches—Martini with Elastic network restraints, Martini with Gō-model restraints, and SIRAH—for their ability to characterize tunnel structure and dynamics across diverse enzyme classes. Using haloalkane dehalogenase LinB and its engineered variants as model systems, we show that CG methods accurately reproduce the geometry of tunnel ensembles observed in all-atom (AA) simulations while providing notable computational speedups. The Martini-Gō model performed particularly well, capturing subtle mutation-induced changes in tunnel dynamics, such as the closure of a main tunnel and the *de novo* opening of a transient auxiliary tunnel in LinB variants. In contrast, Martini with Elastic network restraints was limited in capturing tunnel dynamics due to the structural bias introduced by the restraints. We further validated these findings across nine enzymes from the oxidoreductase, transferase, and hydrolase classes with diverse structural folds. Although all CG methods reliably identified functionally relevant tunnels and provided fairly accurate estimates of their ensemble geometry and key bottleneck residues, they differed in their ability to replicate tunnel dynamics, with tunnel occurrences and ranking showing moderate to good correspondence with AA results. This comprehensive evaluation highlights the strengths and weaknesses of CG simulations, establishing them as powerful tools for high-throughput analysis of enzyme tunnels, which enables more efficient enzyme engineering and drug design efforts targeting these critical structural features.



INTRODUCTION

Enzymes are essential for life, enabling organisms to grow, maintain homeostasis, and reproduce. Despite decades of extensive research, the intricate mechanisms behind enzyme action remain only partially understood due to their immense complexity and diversity. Transport tunnels in enzymes with buried active sites are critical to enzymatic function by controlling substrate access, product release, and overall catalytic efficiency.[1−4] These dynamic structural features are not merely passive conduits but sophisticated regulatory elements that can significantly influence enzyme activity through conformational gating mechanisms.[2] Notably, approximately 50% of enzymes across all six Enzyme Commission (EC) classes contain such buried active sites,[5−7] making tunnels essential yet often overlooked determinants of enzymatic behavior.

Understanding the structural dynamics of these tunnels is essential for deciphering complex structure−function relationships in enzymes and has significant implications for both fundamental enzymology and applied biotechnology. Key elements such as tunnel-lining residues, bottleneck regions controlling the passage diameter, and entrance residues at the tunnel mouth form a sophisticated filtration and regulation system. Strategic engineering of these elements has proven remarkably effective for modulating catalytic properties, with even single-point mutations capable of dramatically altering substrate specificity, reaction rates, and product release kinetics.[3−8] The therapeutic relevance of enzyme tunnels is also noteworthy because numerous disease-associated enzymes contain druggable tunnels, making them attractive targets for pharmaceutical intervention.[9] This has led to the development of therapeutic agents that modulate enzyme function by binding within these tunnels rather than at the active site.[9] Furthermore, tunnels provide structural frameworks for regulatory mechanisms such as substrate inhibition and allosteric cooperativity, which are vital for maintaining cellular homeostasis.[10−16]

Received: October 14, 2025
Revised: November 28, 2025
Accepted: December 12, 2025
Published: January 2, 2026

The dynamic nature of enzyme tunnels presents a significant methodological challenge. Most tunnels are regulated by one or more molecular gates[2,9] that transition between open and closed conformations, resulting in transient tunnels that are difficult to capture using static structural approaches.[17] Consequently, molecular dynamics (MD) simulations have emerged as the gold standard for tunnel investigation,[4,9] although they come with notable limitations. Conventional all-atom (AA)-MD simulations, while providing atomic-level detail, are computationally prohibitive when sampling the microsecond-to-millisecond time scales[18] over which many biologically relevant tunnel dynamics occur.[19,20] The necessity to study slow biological processes has driven the development of enhanced sampling techniques such as metadynamics[21] and umbrella sampling,[22] which accelerate the exploration of conformational space. However, these methods typically require carefully selected collective variables (CVs), which are especially challenging to define for the complex networks of gated transient tunnels. While our previous work demonstrated that Gaussian accelerated MD simulations[20,23] can effectively sample rare tunnel conformations without requiring predefined CVs,[20] the method still demands system-specific parameterization, limiting its broad applicability to diverse enzyme systems. This persistent gap between computational feasibility and biologically relevant time scales underscores the need for alternative approaches that balance accuracy with computational efficiency.

Coarse-grained molecular dynamics (CG-MD) offers a promising solution by systematically reducing molecular complexity through grouping atoms into representative "beads."[24] This reduction in degrees of freedom provides a substantial computational advantage, with CG-MD simulations typically running 10−100 times faster than their AA counterparts.[24−26] This speed enhancement enables the exploration of conformational dynamics over longer time scales, potentially capturing rare tunnel-opening events inaccessible to conventional methods. CG approaches offer flexibility in molecular detail, ranging from highly simplified single-bead-per-residue models to near-atomistic representations where multiple beads capture distinct chemical characteristics of particular functional groups.[25] For this investigation of enzyme tunnel dynamics, we selected two well-established CG models with complementary strengths: SIRAH,[27] which maintains an unbiased approach to protein dynamics, and Martini,[28] which provides force fields with different optional structural restraints.

The Martini CG force field is one of the most rigorously validated and widely applied physics-based CG approaches in biomolecular simulation.[28] The latest Martini 3.0 implementation uses a mapping scheme in which each CG bead represents 2−4 atoms,[29] striking a balance between computational efficiency and chemical accuracy. Notably, Martini 3.0 introduces significant improvements in modeling protein cavities and protein−ligand interactions through rebalanced cross-interactions and optimized bonded parameters. These refinements better represent molecular volume, shape, and packing,[29] which are also critical for accurate tunnel geometry. Such improvements enable better predictions of binding thermodynamics and kinetics.[29−34] However, Martini 3.0 can still lack the fine structural detail needed for high chemical specificity, particularly in representing pocket−ligand interactions, differentiating enantiomers, and modeling binding directionality.[29] To maintain proper protein folding in CG simulations, Martini incorporates artificial stabilization through two distinct approaches: Elastic network models (Elastic)[35] or Gō models (Gō).[36] The Elastic model uses a network of harmonic restraints between backbone beads within a specified cutoff distance, effectively preserving the initial protein structure but limiting major conformational changes.[35] In contrast, Gō models combine two concepts to capture more efficiently the key native contacts: (i) a contact map uses atomic overlaps of heavy atoms to capture contacts beyond the standard Elastic cutoff more selectively and (ii) then applying chemically based selection of stabilizing contacts denoted as the restricted chemical structural units, often reducing the number of applied restraints by a factor of two in contrast to the Elastic model,[37] thereby allowing more flexibility in conformational sampling while maintaining overall structural integrity.[38] In this view, evaluating both approaches could help assess how different levels of conformational bias affect the accurate modeling of enzyme tunnel dynamics.

The SIRAH force field represents a distinct approach to CG that prioritizes structural fidelity through strategic bead placement derived directly from atomistic coordinates.[27] Unlike many CG models that use center-of-mass mapping, SIRAH places protein backbone beads at the actual atomic positions of N, O, and C_α atoms, preserving geometric relationships with AA dihedral angles. This design offers potential advantages for modeling intricate structural features such as enzyme tunnels by maintaining more accurate backbone geometries. SIRAH's mapping scheme also applies a context-dependent resolution, where functional groups critical for specific molecular interactions receive more detailed representation—polar and aromatic moieties involved in hydrogen bonding or π-stacking are modeled with higher granularity, while hydrophobic regions are more aggressively coarse-grained.[27] Unlike Martini, SIRAH enables the calculation of long-range electrostatics through the particle Mesh Ewald method[39,40] and typically operates without imposing artificial restraints to stabilize protein structures, potentially allowing for more natural conformational exploration. This feature makes SIRAH particularly attractive for tunnel studies because it may better capture spontaneous conformational transitions involved in tunnel gating. However, it also introduces challenges, particularly in larger proteins, where maintaining secondary structure over extended simulations can become problematic.[41] These strengths and limitations make SIRAH a valuable counterpart to Martini for evaluating CG approaches to enzyme tunnel characterization, providing complementary insights into the trade-offs between structural stability and conformational flexibility.

Despite the widespread application of Martini and SIRAH in diverse protein systems,[26] their suitability for characterizing the spatiotemporal dynamics of transport tunnels in enzymes remains largely unexplored. This gap represents a significant opportunity, which this study aims to address. We investigate three fundamental questions critical to validating CG models for tunnel analysis: (i) How effectively can CG-MD simulations detect the presence of tunnels and capture their dynamics despite inherent limitations associated with reduced molecular resolution, such as the absence of directional hydrogen bonds? (ii) Can CG-MD simulations detect subtle changes in tunnel properties resulting from targeted mutations? (iii) How do structural biasing schemes in CG models influence the sampling of transient tunnel conformations? To answer these questions, we selected haloalkane dehalogenases as model systems (Figure 1), given their well-characterized

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

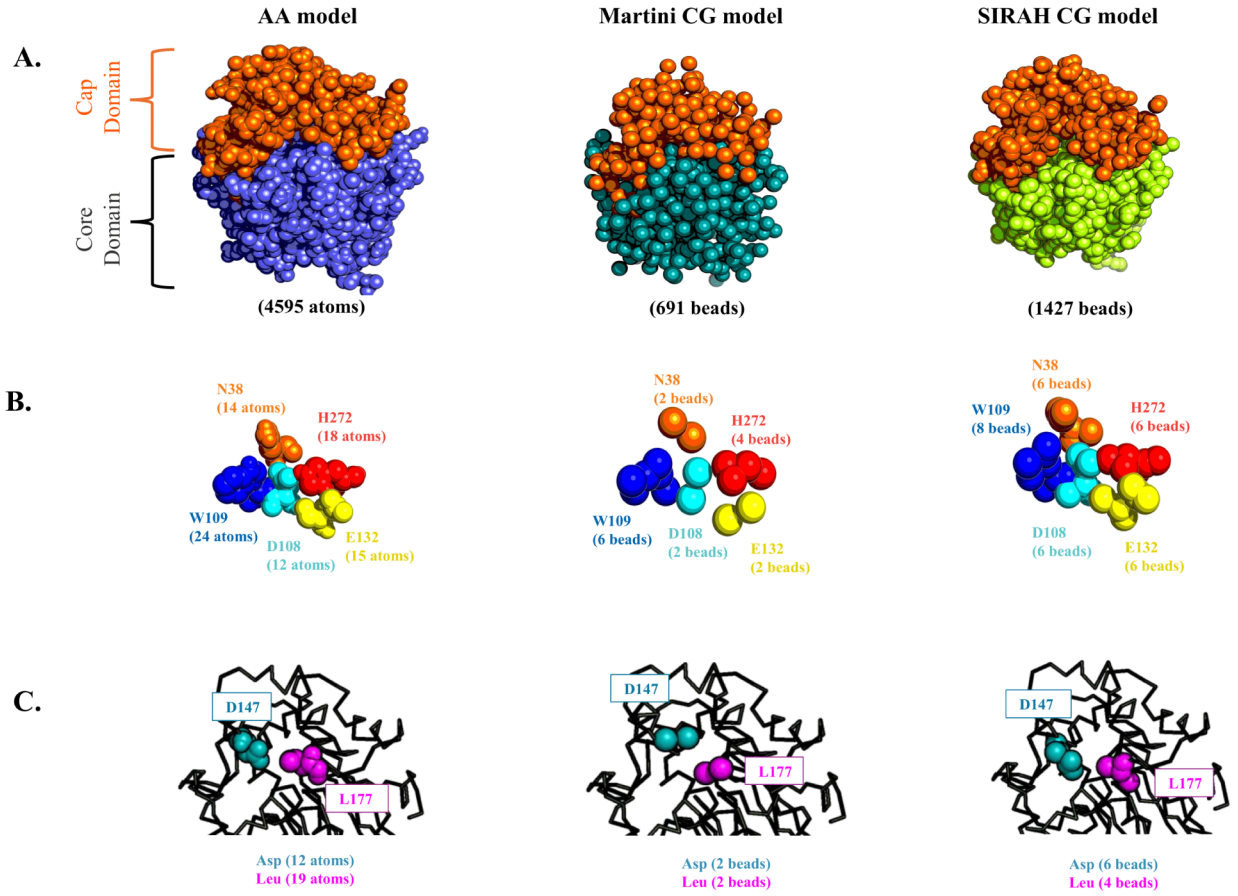


Figure 1. Representations of the level of granularity in the LinB-Wt enzyme: A) whole protein, B) catalytic pentads, and C) two key gating residues located at the mouth of the main p1 tunnel.

buried active sites connected to the solvent by multiple tunnels with varying dynamics. Their catalytic mechanisms are highly dependent on tunnel properties.[42,43] In particular, the haloalkane dehalogenase LinB and its engineered variants offer a gradual series of tunnel modifications with well-documented functional outcomes. The wild-type enzyme (LinB-Wt) features a predominant p1 tunnel complemented by less frequent p2, p3, and side tunnels (ST).[20] Two engineered variants illustrate contrasting tunnel behaviors: LinB-Closed, with reduced p1 accessibility and decreased catalytic efficiency, and LinB-Open, with a newly accessible p3 tunnel that significantly enhances catalytic performance.[44]

We conducted comprehensive comparisons between tunnel ensembles identified in the AA simulations and those obtained from three CG methods: Martini Gō, Martini Elastic, and SIRAH. By quantitatively analyzing tunnel properties—including bottleneck radius, length, curvature, and occurrence frequencies—we established metrics for evaluating how effectively each CG approach captures essential tunnel features. To ensure broader relevance beyond a single enzyme family, we extended our analysis to a diverse set of enzymes spanning three major EC classes (oxidoreductases, transferases, and hydrolases) and representing distinct structural architectures (all-α, all-β, and α/β folds). This systematic validation across functionally and structurally diverse enzymes provides a robust assessment of CG methods for high-throughput tunnel analysis. Our results demonstrate that carefully selected CG methods can accurately capture key features of enzyme tunnels while offering substantial computational efficiency gains. We identified specific strengths and limitations for each approach, with the Martini-Gō model showing particular effectiveness in capturing tunnel dynamics and subtle mutation-induced changes. Importantly, we establish that CG simulations can reliably identify functionally relevant tunnels and their ensemble properties across a wide range of enzyme classes, enabling more efficient enzyme engineering and drug design efforts to target these critical structural features. This work provides a foundation for high-throughput computational screening of tunnel properties in large enzyme data sets, opening new avenues for exploring the relationship between tunnel dynamics and enzyme function in systems previously inaccessible to AA-MD simulation approaches.

METHODS

System Setup and AA-MD Simulations

All initial LinB variant structures—LinB-Wt (PDB code: 1mj5), LinB-Closed (PDB code: 4wdq), and LinB-Open (PDB ID: 5lka)—were protonated at pH 8.5 to match the conditions of experimental assays,[44] with a salt concentration of 0.1 M using the H++ server.[45] Using the tleap module of AMBER18, proteins were positioned at the center of a truncated octahedral box with a 10 Å distance between the solute and the edge of the box and solvated with the three-charge, four-point rigid OPC water model.[46] Next, the systems were neutralized with counterions (Na^+ and Cl^-) to a final concentration of 0.1 M. Energy minimization was performed using 500 steps of steepest descent, followed by 500 steps of

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

conjugate gradient in 5 rounds with decreasing harmonic restraints, utilizing the PMEMD[47] module of AMBER18[48] with the ff14SB force field.[49] Restraints were applied as follows: 500 kcal/mol/Å^2 on all heavy atoms, then 500, 125, 25, and 0 kcal/mol/Å^2 on backbone atoms only. Next, the systems were equilibrated for 2 ns by gradually heating to 310 K under constant volume using the Langevin thermostat[50] with a collision frequency of 1.0 ps^{-1} and harmonic restraints of 5.0 kcal/mol/Å^2 on all enzyme atoms. Periodic boundary conditions (PBC) were applied, and the particle mesh Ewald (PME) method[39,40] was employed for electrostatic interactions beyond a 10 Å cutoff, using default parameters with automatic selection of charge grid density (producing ∼1.0 Å grid spacing), fourth-order B-spline interpolation, and a direct sum tolerance of 10^{-5}. A 4 fs time step was used, enabled by the SHAKE algorithm[51] and hydrogen mass repartitioning (HMR).[52] Considering the usage of HMR could result in altered time scales of studied processes,[53,54] its usage might be suboptimal if the rates of tunnel gating would be studied. However, the average properties of tunnel ensembles should not be markedly affected, as illustrated by the recent comparison of water transport via tunnels of haloalkane dehalogenase DhaA, using two data sets (with HMR and without HMR), 50 × 100 ns of adaptive sampling simulations each.[55]

Subsequently, a 200 ns unrestrained NPT simulation was performed using PMEMD.CUDA with the Monte Carlo barostat[56] and Berendsen thermostat,[57] storing frames every 20 ps. The thermostat was chosen for its ability to preserve the intrinsic dynamics and transport properties of the system, closely resembling NVE simulations while still maintaining temperature control.[58] Clustering analysis was executed with the average-linkage hierarchical agglomerative algorithm using cpptraj[59] on the 200 ns trajectory for each system. The requested number of clusters was set to five with a secondary cutoff of 4.5 Å to identify the five most diverse conformations. For each cluster, the centroid frame was selected as its representative. These representative conformations served as seed structures for five independent 5 μs unrestrained NPT simulations, with frames stored every 200 ps for all three LinB variants.

The analogous simulation protocol was applied across the EC data set composed of the following nine enzymes: EC1 (PDB IDs: 1jfb,[60] 1gp4,[61] and 3bur[62]), EC2 (PDB IDs: 1m15,[63] 1oyg,[64] and 1q20[65]), and EC3 (PDB IDs: 2oup,[66] 1dim,[67] and 1cvl[68]) with the following deviations due to the different nature of protein inputs as well as some updates in simulation methods, which become available within the project timeline.

EC1 Proteins

1jfb[60] (nitric oxide reductase, EC 1.7.1.14). The original PDB structure with HEM protoporphyrin IX containing Fe was first cleaned by stripping the nonprotein atoms. Additionally, residue CYP352 was renamed CYS, and the structure was protonated. After protonation, residue CYS352 was renamed again to CYP352, and the bond between the sulfur of CYP352 and the Fe ion of the HEME group was created. The parameters for the HEME group and CYP residue were obtained from Shahrokh et al.[69] 1gp4[61] (anthocyanidin synthase, EC 1.14.15.1). The PDB structure contains selenomethionines instead of methionines, so all residues were converted to methionine. The Fe^{2+} cofactor was adopted from the analogous 1gp5 structure with lower resolution. 3bur[62] (delta(4)-3-ketosteroid 5-beta-reductase, EC 1.3.1.3). This structure contains an NADPH cofactor that was present during protonation of the structure. The parameters for NADPH were obtained from the AMBER web repository (http://amber.manchester.ac.uk/). These three structures were protonated at the following pH to match the conditions of experimental assays: 7.2 (1jfb),[70] 7.2 (1gp4),[61] and 6.0 (3bur),[62] using the H++ web server with a salt concentration of 0.1 M, and internal and external dielectric constants of 4 and 80, respectively.

EC2 Proteins

1m15[63] (arginine kinase, EC 2.7.3.3). All ligands were removed from the structure, as they are not cofactors. 1oyg[64] (levansucrase, EC 2.4.1.10). The Ca^{2+} ion was considered a cofactor, pentabipyramidally coordinated with a water molecule and the carbonyl oxygen of Leu308 at the apices, and with Asp339 $O_{\delta1}$ and $O_{\delta2}$, Asn310 $O_{\delta1}$, Asp241 $O_{\delta1}$, and Gln272 $O_{\varepsilon1}$ at the equatorial positions. These interactions were represented as distance restraints during the simulation. The structure was protonated with the Ca^{2+} ion and a coordinating water molecule present. 1q20[65] (cholesterol sulfotransferase, EC 2.8.2.2). All ligands were removed from the structure because they are substrates, not cofactors (substrate donor product PAP and acceptor substrate pregnenolone). These three structures were protonated at the following pH to match the conditions of experimental assays: 8.0 (1m15),[71] 6.0 (1oyg),[72] and 4.0 (1q20),[73] using the H++ web server with a salt concentration of 0.1 M, and internal and external dielectric constants of 4 and 80, respectively.

EC3 Proteins

2oup[66] (PDE10A [3′,5′-cyclic-nucleotide phosphodiesterase], EC 3.1.4.17) and 1dim[67] (sialidase, EC 3.2.1.18) required no special modifications. 1cvl[68] (bacterial lipase, EC 3.1.1.3). This structure contains a Ca^{2+} ion, which is coordinated by four oxygen atoms from the protein and two water molecules. This Ca^{2+} ion was retained in the system during protonation. These three structures were protonated at the following pH levels to match the conditions of experimental assays: 7.5 (2oup),[66] 7.0 (1dim),[74] and 4.0 (1cvl),[75] using the H++ web server with a salt concentration of 0.1 M, and internal and external dielectric constants of 10 and 80, respectively.

For system building, after retrieval and protonation of the structures, water molecules were introduced around the protein using a tandem approach with the three-dimensional molecular theory of solvation (3D-RISM)[76] and the Placevent algorithm.[77] 3D-RISM-predicted and crystallographic waters were combined, retaining only water molecules located at least 2.0 Å from the protein and giving preference to crystallographic waters validated using the EDIA scorer program.[78] The systems were further processed using the Tleap module of AMBER20. Proteins were positioned at the center of a periodic truncated octahedral box with a 10 Å distance between the solute and the edge of the box and solvated with the three-charge, four-point rigid OPC water model.[46] Systems were neutralized with Na^+ and Cl^- ions to achieve a salt concentration of 0.1 M. The parameters of the systems were prepared using the ff19SB force field,[79] applying the HMR procedure to the solute atoms to enable a 4 fs time step during simulations.[52] AMBER22 software was used to perform these simulations. A single unrestrained NPT simulation was carried

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

out for 2 μs for each enzyme with frames stored every 200 ps. The cofactors (Fe^{2+} and Ca^{2+}) were restrained using distance restraints within the enzyme, but no restraint was applied to the HEME and NADP cofactors.

System Setup and CG-MD Simulations Using SIRAH

To set up the SIRAH CG models, the five most diverse structures obtained through cluster analysis in AA simulations were used as starting seed structures. These atomic seed structures were then mapped to the SIRAH CG model using SIRAH Tools[80] and the SIRAH 2.0 force field. For solvation, a pre-equilibrated truncated octahedral box of CG WT4 water molecules with a 10 Å distance between the solute and the edge of the box was used. The WT4 molecules were randomly replaced by Na^+ and Cl^- CG ions to achieve an ionic strength of 0.1 M, and counterions were added to the system following the method described by Machado et al.[81] System minimization was performed using 5,000 steps of steepest descent followed by 5,000 steps of conjugate gradient in five rounds with decreasing harmonic restraints, utilizing the PMEMD.CUDA module of AMBER18. A restraint of 500 kcal/mol/Å^2 was applied to heavy atoms, followed by 500, 125, 25, and 0 kcal/mol/Å^2 on backbone GN and GO beads. Additionally, an in-house-developed restraint protocol (Figure S1, Table S1) was applied to the dihedral angles of residues in α-helices (4 kcal/mol/Å^2) and the distances between H-bond-forming atoms of residues in β-sheets (20 kcal/mol/Å^2) to maintain structural stability. These restrained residues were selected using secondary structure analysis with the cpptraj module's secstruct function on the respective input structures. The structures were equilibrated for 2 ns, gradually heating to 310 K, followed by 5 ns of NVT and 10 ns of NPT equilibration. PBC was applied, and the PME method was employed for electrostatic interactions beyond a 10 Å cutoff, using default parameters with automatic selection of charge grid density (producing ∼1.0 Å grid spacing), fourth-order B-spline interpolation, and a direct sum tolerance of 10^{-5}. Finally, a 20 fs time step was used to perform 5 μs of unrestrained production simulation at constant pressure and temperature using a Langevin thermostat with a collision frequency of 50 ps^{-1} and a Berendsen barostat. The parameters were adopted from Machado et al.[27] The same simulation protocol was applied to seven enzymes: EC1 (PDB ID: 1gp4), EC2 (PDB IDs: 1m15, 1oyg, and 1q20), and EC3 (PDB IDs: 2oup, 1dim, and 1cvl). Updated AMBER22 software was used for performing these simulations. A single unrestrained NPT simulation was conducted for 2 μs for each enzyme with frames stored every 200 ps. Due to the unavailability of a small-molecule library in SIRAH, cofactors such as HEME and NADP were absent. Therefore, enzymes 1jfb and 3bur were not simulated using SIRAH and were excluded from the analysis. Additionally, the cofactors Fe^{2+} (replaced by Ca^{2+} due to the lack of Fe^{2+} parameters in SIRAH)[82] and Ca^{2+} were restrained using distance restraints within the enzyme.

System Setup and CG-MD Simulations Using Martini

For the CG Martini system setup, the seed structures were retained as obtained from the clustering analysis of AA simulations. These structures were converted to the CG model by the martinize2 tool.[83] Next, they were solvated with a truncated octahedral box of Martini CG water molecules with a 10 Å distance between the solute and the edge of the box. All calculations were performed using GROMACS version 2023.3. To maintain the native structure of the CG model, the Elastic network and GōMartini models were applied with default settings (Table S1), which are optimized for the description of globular proteins with the Martini force field 3.0.[35,36,38] The bonded parameters of both the Elastic network and GōMartini 3 were kept the same to compare their structural propensities. Martini force field version 3.0[29] was used. For the GōMartini 3 model, the contact map was first obtained from the web server GōContactMap.[84,85] For the Gō contact map, the distance cutoff between C_α−C_α was kept within the range of 0.3−1.1 nm. The strength of the Gō potentials was set to a default value of 9.414 kJ/mol. Secondary structures were assigned based on the AA structure using the DSSP module,[86] and the martinize2 tool was used to switch on the Elastic network with a bond force constant of 700 kJ·mol^{-1}·nm^2. The lower and upper elastic bond cutoffs were kept at their default values of 0.5 and 0.9 nm, respectively. Cutoff values ranging from 0.8 to 1.0 nm were shown to provide sufficient agreement with AA simulations.[35]

The simulation protocol was kept the same for the Elastic and Gō models. Energy minimization was performed using 400,000 steps of steepest descent. The systems were then NVT-equilibrated under position restraints for 20 ns, gradually heated to 310 K under constant volume using the V-rescale thermostat[87] with a 10 fs time step. The PBC was set to the default xyz, with full PBC applied in all three dimensions. The radius for the neighbor list (rlist) was set to 1.4 nm, and the cutoff scheme was set to Verlet. The nonbonded cutoffs, electrostatic (rcoulomb) and van der Waals (rvdw) interaction cutoffs, were both set to 1.2 nm. NVT equilibration was followed by position-restrained NPT equilibration for 20 ns with a 10 fs time step and pressure control using the C-rescale barostat.[88] Finally, a 5 μs unrestrained production simulation was run using a 20 fs time step with the Parrinello−Rahman barostat[89] and V-rescale thermostat.

The same simulation protocol was applied across nine enzymes: EC1 (PDB IDs: 1jfb, 1gp4, and 3bur), EC2 (PDB IDs: 1m15, 1oyg, and 1q20), and EC3 (PDB IDs: 2oup, 1dim, and 1cvl) to capture the diversity of enzyme classes. An unrestrained NPT simulation was performed for 2 μs for each enzyme with frames stored every 200 ps. The CG representations of HEME and NADP were adopted using Martini 2.0 beads,[90] which were updated to Martini 3.0 beads for the simulations. Due to the larger time step, positional restraints were applied to HEME (1jfb) and NADP (3bur) to keep the cofactors in place, which is necessary for enzyme function. Additionally, the cofactors Fe^{2+} (replaced by Ca^{2+}, which is represented by an S-type bead[82] due to the unavailability of Fe^{2+}) and Ca^{2+} were restrained using distance restraints within the enzyme.

Protein Stability Analysis

To confirm the stability of the enzymes explored through AA and CG-MD simulations, root mean square deviation (RMSD) and root mean square fluctuation (RMSF) were calculated with the reference set as the initial structure, considering all atoms of the protein. RMSD was calculated for all protein residues. The radius of gyration (R_g) and solvent-accessible surface area (SASA) were also calculated to confirm whether the protein remained intact throughout the simulation. These analyses were performed using cpptraj and gmx[91] tools for AMBER and GROMACS simulations, respectively. The exception was the SASA calculation for AMBER simulations

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

that were performed with MDTraj 1.9.9 software to incorporate appropriate SIRAH bead radii.[92]

Tunnel Ensemble Analysis

Tunnels were analyzed using CAVER 3.2.0 software,[17] which identifies pathways in proteins by constructing a Voronoi diagram (VD) of their atomic structure.[93] The edges and vertices of such a VD contain information about the surrounding empty space within the protein. Next, the edges smaller than the user-defined probe radius (0.9 Å here) are removed, and the simplified VD is searched for continuous tunnels from the starting point (defined here by three catalytic residues N38, D109, and H272; numbering corresponds to the crystal structure) to the protein surface using Dijkstra's algorithm.[94] Tunnels from each trajectory frame are clustered using the average-linkage agglomerative clustering method with a clustering cutoff of 3.0 and a frame reweighting coefficient of 2.0 to identify corresponding tunnel ensembles.[95] The bottleneck residues were identified as those having their surfaces closer to the surface of the tunnel bottleneck sphere than 3 Å. To calculate CG tunnels with SIRAH and Martini beads, the CAVER amino acid library was extended with the beads' names and their van der Waals radii. A divide-and-conquer approach[96] was used for efficient processing of massive tunnel data sets by splitting the tunnel calculations into 5−10 batches. For AA simulations, the CAVER filtering cutoff was set to 5 in 25,000 frames to preserve rare tunnels, while for CG methods, it was set to 250 out of 25,000 frames due to the higher frequency of tunnel detection. The reclustering algorithm[55] was later used to ensure only tunnels with consistent entrances remained in any cluster. HDBSCAN was applied to the tunnel end points, with the following parameters: cluster_selection_epsilon of 1.5, allow_single_cluster set to True, and both min_samples and min_cluster_size set to 5. The recreated tunnels were considered independent tunnel clusters, while those identified as noise were discarded.

To establish correspondence between tunnel networks found in individual AA and CG simulations of all LinB variants, we performed comparative analysis using TransportTools 0.9.3 software,[97] where the tunnel clusters from individual simulations produced by all investigated methods were clustered into uniform "superclusters" with the average-linkage method using a cutoff of 1.5, based on the average distances between the tunnel cluster surfaces measured at regular distances along their lengths. These superclusters are then ranked by their priority scores, which combine tunnel occurrences in the simulations with their average geometrical throughput. Reference structures for tunnel network alignments were obtained using the backmapping protocol for Martini using the backwards.py script.[98] For SIRAH, reference structures were obtained using the backmapping module of the SIRAH tools.[80] Only superclusters that were detected in at least 5 simulations across all identified methods were retained. In cases in which multiple branches of a given tunnel were identified, the major branch preferred across all simulations was considered for analysis. The superclusters corresponding to the already known and verified tunnels (p1, p2, p3, and ST) were identified visually based on the previously defined localization of these tunnels in the LinB structure.[20,99,100] The same procedure was followed for enzymes from the three respective EC classes, and separate TransportTools comparative analyses were performed for each protein variant. The Martini backwards.py script[98] was also used to perform backmapping of Elastic simulations for the analysis of extra tunnels present in CG results.

The occurrence frequencies of bottleneck residues for known transport pathways (p1, p2, and p3) of three LinB variants were obtained from TransportTools outputs and averaged to obtain mean occurrence frequencies. Using these data, the performance of CG methods in bottleneck residue detection was assessed by top-K overlap, i.e., the fraction of top-K bottleneck residues from AA simulations found in top-K bottleneck residues from respective CG methods for K = 5, 10, 15, and 20. The proficiency of CG methods in ranking tunnel importances was assessed in three LinB variants and 9 EC enzyme cases using 125,000 and 10,000 simulation snapshots per system, respectively. Only tunnels occurring in at least 10% of snapshots were considered. The priority scores of tunnels calculated by TransportTools were compared between AA and CG simulations using two distinct metrics: (i) Pearson correlation—linear correlation of tunnel priority scores between AA and CG. Higher correlations indicate better preservation of relative importance magnitudes; (ii) Average precision (AP)—evaluated ranking quality by calculating precision at each position where a relevant tunnel (from AA) appeared in the CG ranking. For each relevant tunnel found at position i, *precision@i = (relevant tunnels found so far)/i*. The AP was computed as the mean of these precision values across all relevant positions. This metric heavily penalizes relevant tunnels retrieved with low rankings, emphasizing early identification of important ones. The Pearson correlation measures agreement in importance scoring between methods, while AP assesses positional ranking quality throughout the tunnel list.

Statistical Analysis of Tunnel Properties

Comparisons of tunnel properties (occurrence, bottleneck radius, and length) between AA and CG methods were performed using the Python SciPy library.[101] Data normality was assessed using the Shapiro−Wilk test,[102] and variance homogeneity was tested using Levene's test.[103] Then, the nonparametric Kruskal−Wallis tests were performed,[104] followed by post hoc Mann−Whitney U test,[105] using the Bonferroni correction to control for family-wise error rates.[106]

RESULTS AND DISCUSSION

CG Models of Three LinB Variants Remain Stable during Simulations

Prior to analyzing transport tunnels in CG-MD simulations, we first assessed the structural integrity of the simulated systems. Although Martini CG methods incorporate restraints, we thoroughly evaluated protein stability to ensure reliable conformational sampling. Notably, the default SIRAH model, which lacks an established restraint protocol, exhibited significant loss of secondary structure elements during preliminary simulations (Figure S2), as previously reported for other larger proteins.[41] To address this limitation, we developed a targeted restraint strategy that selectively applies hydrogen bond constraints to β-sheets and dihedral angle restraints to α-helices (Figure S1). Implementing this protocol resulted in substantial improvements in structural stability across all simulations (Figures S2−S3), reducing the coil content by approximately 10% while maintaining the integrity of α-helical and β-sheet elements. By using this restraint protocol, we were able to produce simulations with the SIRAH

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

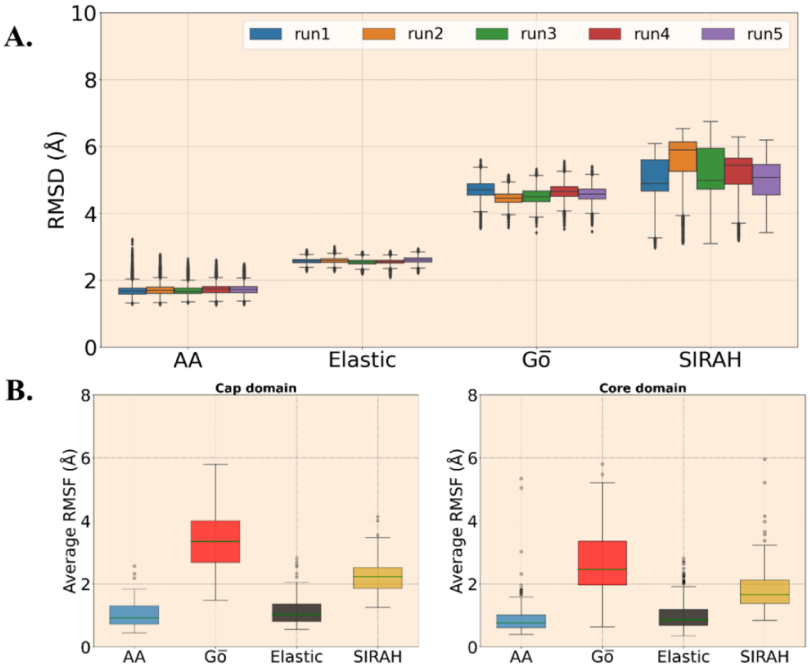


Figure 2. Stability and dynamics of LinB-Wt simulated with the investigated models. A) RMSD of the full protein in five replicates fitted to their respective seed structures clustered from the initial AA simulations. B) Average RMSF of residues in the cap and core domains. The box plots show the median (middle line in the box), with the box representing the interquartile range (from the 25th to the 75th percentile), indicating the spread of the middle 50% of the data. The whiskers extending from the box represent 1.5 times the interquartile range. Outliers are shown with open circles.

model that were stable enough to serve as the input for tunnel analyses, expanding the repertoire of CG approaches in our comparison.

To confirm the global stability of proteins in our biased CG-MD simulations, we compared the RMSD profiles of LinB-Wt with those from the AA simulations (Figure 2A). This analysis revealed a hierarchical pattern of conformational stability across the different CG models, with the Martini Elastic model exhibiting the closest alignment to the AA simulations, followed by the Gō and SIRAH models. This trend correlated directly with the extent of the restraint networks applied (Table S1). Here, it is worth noting that the stability of LinB-Wt produced with restrained SIRAH simulations (RMSD predominantly between 5−6 Å) was near to those observed for shorter (1 μs) simulations of more compact and markedly smaller proteins (RMSD $\leq$ 4.8 Å),[27] indicating that while the restraint protocol delivered baseline protein stability, it would likely still benefit from further optimization. Importantly, when considering the dynamics of whole LinB-Wt or its two domains (Figures 2B and S4), the Elastic model enforced conformational rigidity nearly analogous to that of AA simulations, while both the Gō and SIRAH models allowed for more extensive protein motion, offering complementary insights into tunnel behavior under different mobility regimes. These patterns were consistently observed in both the LinB-Closed and LinB-Open variants (Figures S5−S9).

Also, the geometry of the catalytic pentad (Figure 1B), critical for enzyme function, remained largely preserved by the CG methods (Figure S10). The highest fidelity of the pentad structure was observed in the Elastic and Gō model, exhibiting an average whole-residue RMSD $\leq$ 3 Å and backbone RMSD $\leq$ 1 Å, in line with AA simulations. In the case of the SIRAH model, the pentad structure exhibited larger deviations (average whole-residue RMSD $\leq$ 5 Å), which could be traced to the backbone mobility of residue H272. Unlike other catalytic residues, this residue is located on an unrestrained loop closer to the protein surface. In two SIRAH simulations, this loop eventually adopted a distinct conformation, perturbing the position of this histidine, which could not be entirely prevented with the current restraint protocol. Furthermore, both R_g and SASA remained stable across all LinB variants. The Elastic model yielded more compact structures, while SIRAH simulations showed an average R_g increase of about 1 Å (5%) compared to that of AA (Figures S11−S13). While all CG models showed larger SASA values in line with their larger porosity (Figures S14−S16), the most notable increase was observed with SIRAH, consistent with the highest RMSD and R_g values observed for this model, in agreement with previously reported tendencies.[27] Overall, these analyses validated the AA and CG-MD simulations, confirming their suitability for subsequent detailed tunnel analyses (Figure 2, Figures S4−S16).

Exploration of Tunnels Using CG Models in Comparison to the AA Model

After establishing the stability of all simulations, we proceeded with a quantitative comparison of tunnel networks identified in CG-MD versus AA simulations using the comparative module from the TransportTools package,[97] focusing only on tunnel ensembles with more than 10% occurrence along the simulations that were sampled in at least 3 independent simulations.

Most importantly, our results revealed that all tested CG methods successfully identified both the main and three transient tunnels known in LinB-Wt (Figure 3A, Table S2), demonstrating the fundamental ability of CG approaches to identify those experimentally validated tunnels, despite their reduced molecular resolution. However, we also noticed that

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

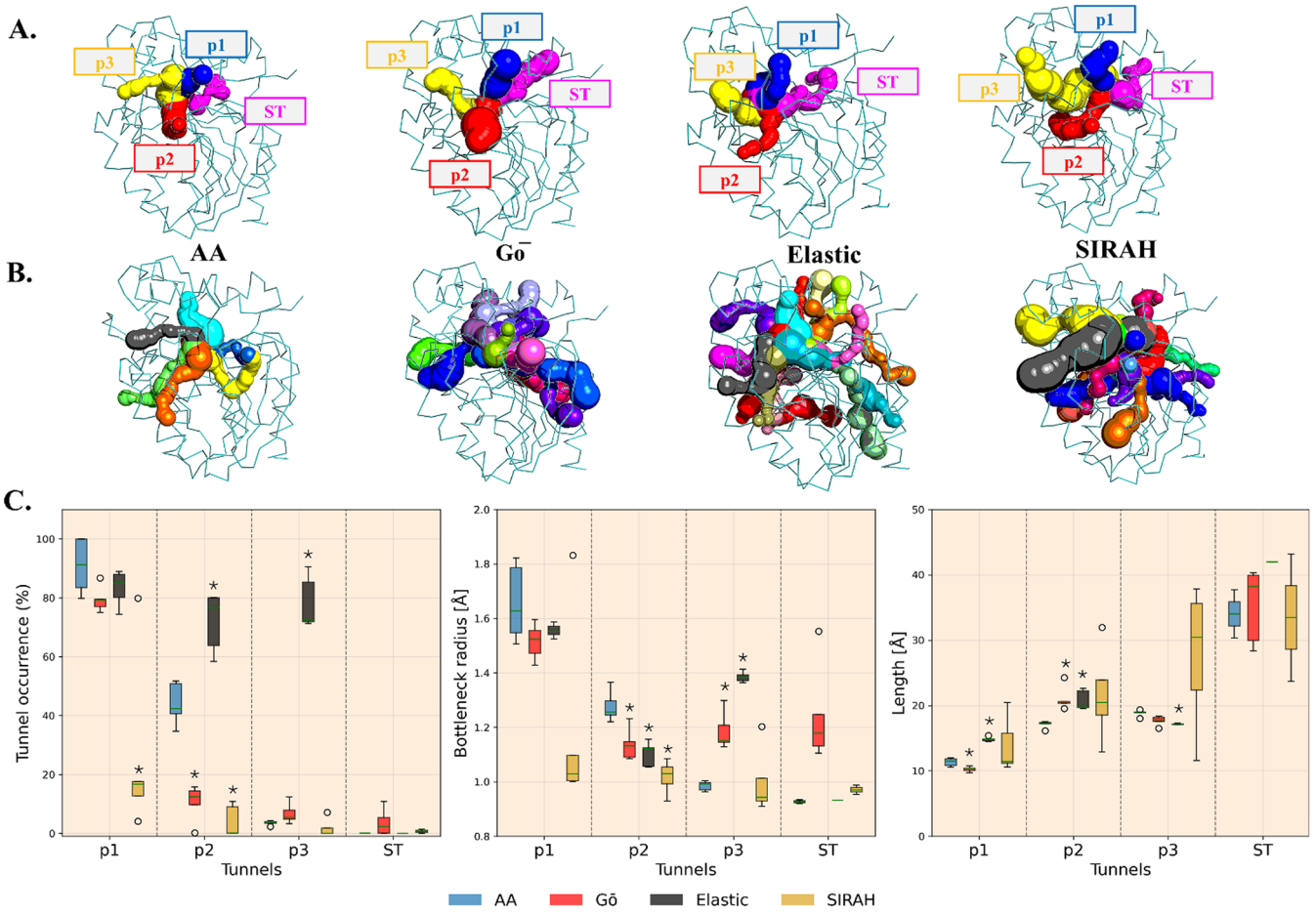


Figure 3. Comparison of tunnel structure and occurrence in the LinB-Wt. A) Representation of four known, previously validated tunnels obtained through individual methods, the major p1 tunnel, followed by transient tunnels p2 and p3, and finally, a very rare ST tunnel. B) Additional tunnels captured by each method (applied cutoff to keep only tunnels seen in at least 10% of the total simulation time and at least three simulations). C) Properties of these tunnels were captured using different methods. The box plot shows the median (middle line in the box), with the box representing the interquartile range (from the 25th to the 75th percentile), indicating the spread of the middle 50% of the data. The whiskers extending from the box represent 1.5 times the interquartile range. Outliers are shown with open circles. The asterisks denote statistically significant differences from AA simulations (corrected *P*-value <0.016667). Note that properties of the ST tunnel cannot be quantitatively contrasted due to the limited sampling of this rare tunnel in AA simulations.[20] For details of statistical analysis, see Table S3.

the rarest of these tunnels, ST, has been insufficiently explored in AA simulations, unlike with most CG methods. This is in agreement with the previous observation that efficient exploration of the ST tunnel requires the use of enhanced sampling AA simulations.[20] The lack of reference data from AA simulations disallowed the accurate quantitative analysis of ST tunnel properties. Along these lines, we noted that CG simulations consistently revealed a more expansive tunnel landscape, detecting numerous additional transient tunnels and tunnel branches not observed in AA simulations (Figure 3B, Table S2). To disentangle whether these additional tunnels arose from the inherent granularity of CG representations or enhanced conformational sampling, we reconstructed atomistic structures from CG coordinates obtained through Elastic simulations of LinB-Wt with backmapping and analyzed the resulting tunnel networks in the backmapped ensemble. We observed that the occurrences of even the most prevalent additional tunnels were markedly reduced after backmapping, from up to 40% to below 20%, but were not completely absent (Figure S17). This result indicates that the additional tunnels primarily represent geometrically legitimate rarer tunnels or their branches, accessed through the enhanced sampling efficiency of CG-MD simulations rather than artifacts due to the coarseness of CG beads. Such exploration of a broader conformational landscape of tunnel networks suggests the potential benefit of CG methods in identifying candidate tunnels for further in-depth exploration that conventional AA simulations may miss. Considering the lower resolution of CG-MD simulations, the functional relevance of these candidate tunnels could only be confirmed with extensive AA simulations, or better yet, by probing their explicit utilization by relevant small molecules.[19,20]

Next, we analyzed tunnel occurrence rates (Figure 3C), which are frequently used as indicators for assessing the functional relevance of tunnels in enzymatic systems.[107] First, from a qualitative point of inference, testing if CG methods can distinguish the primary p1 tunnel, known to be nearly permanently open in LinB-Wt,[44] from the other transient tunnels based on their occurrences and their ranking. In this respect, the Gō model performed the best in maintaining a clear separation between permanent and transient tunnels and preserving their relative prioritization analogous to that in AA simulations (Figure 3C). In contrast, the Elastic model exhibited a significant limitation by markedly overestimating

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

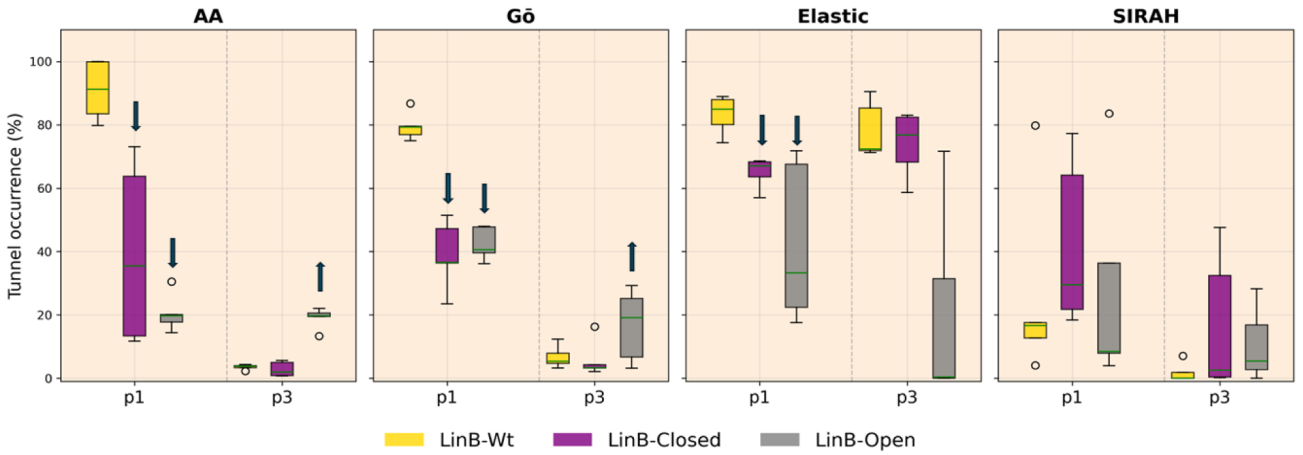


Figure 4. Comparison of the effect of mutations on tunnel occurrence delineated by different methods. Tunnel p1 and p3 from the AA, Gō model, Elastic model, and SIRAH model in LinB-Wt, LinB-Closed, and LinB-Open. The box plot shows the median (middle line in the box), with the box representing the interquartile range (from the 25th to the 75th percentile), indicating the spread of the middle 50% of the data. The whiskers extending from the box represent 1.5 times the interquartile range. Outliers are shown with open circles.

the relevance of both p2 and p3 tunnels. A detailed investigation revealed that this discrepancy arose from fundamentally different sampling patterns: AA simulations distributed tunnel events across multiple distinct branches, whereas the constrained dynamics of the Elastic model artificially concentrated occurrences within a single dominant branch (Figures S18−S19). This consolidation effect blurred the critical distinction between persistent and transient pathways, compromising the functional relevance of tunnel ranking in this model. Such behavior is consistent with the well-documented tendency of Elastic network models to overstabilize initial backbone conformations, imposing artificial rigidity that limits conformational exploration—an issue previously noted in protein−protein interaction studies.[108−110] Intriguingly, this overstabilization effect showed context-dependent behavior across the LinB variants. While the enhanced p3 tunnel occurrence persisted in LinB-Closed, it was absent in LinB-Open (Figure S20, Table S4), where mutations replacing bulky gating residues with smaller side chains fundamentally altered the tunnel's branching potential.[44] Conversely, the SIRAH model lacked discrimination between the main p1 tunnel and transient alternatives in LinB-Wt due to its uniquely underrepresented p1 occurrence (Figure 3C). Notably, this effect was absent in LinB variants with partially occluded p1 tunnels due to the L177W mutation (Figure S20). Next, we evaluated the overall quantitative ranking of tunnels based on their priority scores, which reflect their prevalence as well as their geometric throughputs (Figure S21). For all CG methods, we observed a moderate correlation of tunnel ranking with the AA results (Pearson coefficients 0.46−0.51). More clear differences were observed when AP was considered: Elastic model showed good ranking precision, the Gō model retained moderate performance, while the weak outcomes of SIRAH could be linked to its inability to properly prioritize the p1 tunnel in LinB-Wt. Such values indicate that while CG methods can provide valuable insights into tunnel ranking, their outputs need to be carefully considered and further validated. Overall, these analyses suggest that Gō restraints strike a better balance between stability and conformational flexibility of the tunnels, making this method the most effective in recognizing major tunnels from auxiliary ones and allowing further investigations on the key tunnels. However, when the quantitative ranking of all tunnels is important, the Elastic model represents a preferential choice, despite its drawback of overestimating certain transient tunnels in some LinB variants.

Finally, we analyzed to what degree the key structural properties of the detected tunnel ensembles, such as their length, bottleneck radius, and composition of bottleneck-forming residues, as determined from AA simulations, were preserved in CG-MD simulations. Considering the major p1 tunnel, our analysis showed remarkable agreement across all CG models in all LinB variants (Figures 3C and S20). The only minor discrepancies were detected for the length of this tunnel in LinB-Wt, which was about 1 Å shorter with the Gō method and 3.6 Å longer with the Elastic method (relative errors of 9% and 31%, respectively). Here, we would also like to note the narrower bottleneck of the p1 tunnel observed with SIRAH (by 0.5 Å), which, however, was not statistically significant due to its rather large variance. This high degree of structural correspondence for major tunnels establishes CG methods as useful platforms for preliminary tunnel mapping, offering geometrically accurate starting points for more detailed investigations into tunnel utilization by substrate and product molecules using more advanced computational approaches.[19,20] Considering transient p2 and p3 tunnels (Figures 3C and S20), the CG methods mostly reached relative errors in their bottleneck radii ≤20%, i.e., minor differences within 0.2 Å. The only marked exception was the considerably overestimated radii of the p3 tunnel in LinB-Wt and LinBClosed with the Elastic model (by circa 0.4 Å, with relative errors of 34−41%). The deviations in the length of these tunnels did not surpass those observed for the major p1 tunnel, except for the notably longer p3 tunnel with the SIRAH method. Given the appreciable agreement in tunnel geometries, we evaluated the performance of CG methods in detecting bottleneck residues, which represent the major targets for tunnel engineering.[8] We observed 40−79% overlap between top-5, top-10, top-15, and top-20 most frequent bottleneck residues identified from AA simulations and all CG methods (Figure S22), constituting valuable insights given that LinB proteins comprise about 300 residues. The Elastic method was the most accurate in this respect, providing 64% agreement even for a very strict task of finding the top 5

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

bottleneck residues. Considering broader screening for 10−20 bottleneck residues, all CG methods correctly found most of the key bottleneck residues in such a region.

Sensitivity of CG-MD Simulations to Mutation-Induced Alterations in Tunnel Dynamics

Beyond characterizing native tunnel dynamics, we performed a more stringent test of CG methods by evaluating their ability to detect subtle changes in tunnel dynamics upon mutations, a key capability for enzyme engineering applications. For this purpose, we conducted a comparative analysis of p1 and p3 tunnels across LinB-Wt and two engineered variants with distinct mutational effects on tunnel architecture.[20,44]

The primary p1 transport tunnel exhibits reduced accessibility in both the LinB-Open and LinB-Closed variants due to the L177W mutation at its entrance, which partially occludes the tunnel. In contrast, the LinB-Open variant features an engineered enhancement of the p3 tunnel through a triple mutation (W140A + F143L + I211L) targeting key bottleneck residues. This design successfully reconfigured the p3 tunnel to nearly match the occurrence rates of the native p2 tunnel, creating an additional transport route. Our analysis revealed striking differences in how accurately each CG approach captured these mutational effects (Figure 4). As our reference standard, AA simulations, along with the Gō model, demonstrated comparable sensitivities, accurately capturing both mutational effects. The Elastic model showed partial success, correctly recapitulating the p1 tunnel constriction but failing to capture the enhancement of the p3 accessibility. This limitation stems from the previously noted overestimation of the p3 occurrence in both the LinB-Wt and LinB-Closed variants, which obscures the relative increase in the engineered LinB-Open variant. Most notably, the SIRAH model failed to capture both mutational effects despite its adequate performance in native-state characterization. This limitation is likely due to the interplay between its CG representation and our applied secondary-structure-based restraint protocol, which may overly constrain the local conformational adaptations needed to manifest mutation-induced tunnel perturbations. This underscores the importance of optimizing restraint schemes when applying CG methods to mutational analysis and suggests that SIRAH may require alternative stabilization strategies to enable more effective analyses.

Case Study: Evaluating CG Methods across Diverse Enzymatic Systems

To evaluate the broader applicability of CG methods beyond a single enzyme family, we expanded our investigation to include a diverse set of enzymatic systems, encompassing both structural and functional variability. We selected nine representative enzymes spanning three major EC classes: oxidoreductases (EC1), transferases (EC2), and hydrolases (EC3). Within each functional class, we included enzymes with distinct structural architectures: all-α, all-β, and α/β folds. The EC1 class of enzymes catalyzes redox reactions[111] that involve the transfer of electrons from a donor to an acceptor. These enzymes frequently utilize cofactors, such as NAD^+/NADH or $FAD/FADH_2$ to facilitate electron transfer. They are vital in processes such as detoxification and respiration, where they play key roles in energy production through electron transfer during metabolic pathways. In this study, we focused on three enzymes with buried active sites from this class—nitric oxide reductase with a HEME cofactor (PDB ID: 1jfb; all-α), anthocyanidin synthase (PDB ID: 1gp4; all-β), and delta(4)-3-ketosteroid 5-beta-reductase with an NADP cofactor (PDB ID: 3bur; α/β). Notably, SIRAH simulations could only be performed on anthocyanidin synthase (1gp4) due to limitations in the force field's cofactor parameter library, which lacks representations for the HEME and NADP cofactors used in the other two enzymes. Consequently, SIRAH analysis was restricted to a single all-β protein in this EC class. The EC2 class of enzymes catalyzes the transfer of functional groups from one molecule to another.[111] These enzymes are essential for biosynthesis and metabolic regulation and are involved in processes such as phosphorylation and glycosylation. In this study, we examined three enzymes with buried active sites specific to this class—arginine kinase (PDB ID: 1m15; all-α), *Bacillus subtilis* levansucrase (PDB ID: 1oyg; all-β), and human cholesterol sulfotransferase SULT2B1b (PDB ID: 1q20; α/β). Finally, we evaluated the CG methods with the EC3 class of enzymes, which catalyze the hydrolytic breakdown of various bonds such as peptide or ester bonds. These enzymes are crucial in digestive processes, breaking down macromolecules such as proteins, fats, and nucleic acids into smaller components, aiding in their metabolism and recycling. We analyzed three enzymes with buried active sites from this class—phosphodiesterase 10 (PDB ID: 2oup; all-α), *Salmonella typhimurium* LT2 neuraminidase (PDB ID: 1dim; all-β), and bacterial lipase (PDB ID: 1cvl; α/β). This strategic selection of enzymatic systems, all featuring catalytically essential, buried active sites with diverse tunnel networks, provided a robust evaluation framework for CG methods under varying protein topologies, sizes, and dynamic characteristics. Encouragingly, the stability and dynamics of all proteins simulated with all four methods followed the same trends observed for LinB enzymes (Figures S23 and S32), supporting the generalizability of the applied CG protocols and indicating their robust performance in comparative tunnel analysis independent of protein architecture. Of note, unlike with LinB variants, AA simulations of several enzymes included in this data set revealed much larger fluctuations in particular regions, mostly formed by loops that were to some extent approximated with the Gō model but were not accessible to simulations with Elastic and SIRAH models (Figures S23−S32).

Regarding the captured tunnel landscapes, our comprehensive analysis of these nine diverse enzymes reveals several generalizable patterns in the CG method performance. Most notably, all tested CG methods successfully captured the most prevalent tunnels identified in reference AA simulations, with particularly strong correspondence for high-occurrence tunnels, which are likely to carry the greatest functional relevance (Figures S33−S36). Consistent with our findings with LinB enzymes, CG methods identified additional tunnels and tunnel branches beyond those captured in AA simulations, predominantly representing lower-occurrence pathways (Figure S36). Moreover, the CG methods failed to capture only a small subset of the low-frequency tunnels identified by AA methods (Figure S36). This asymmetric pattern suggests that CG methods have the capacity to expand the detectable tunnel landscape, a feature that could be advantageous when aiming to map a comprehensive range of potential transport tunnels, particularly the rare and transient ones. To quantify the correspondence between CG and AA tunnel geometries, we conducted correlation analyses comparing tunnel bottleneck radii and lengths across all models (Figures S37−S44). These geometric characteristics of tunnel ensembles showed sub-

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

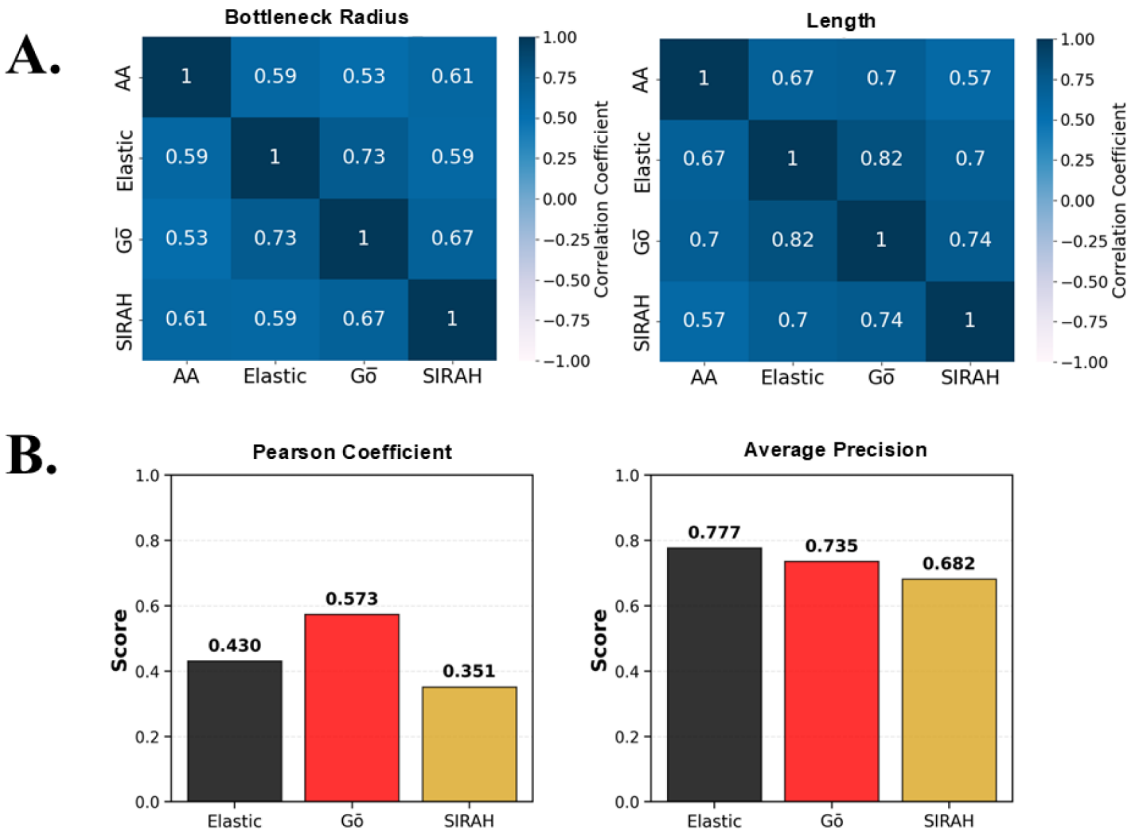


Figure 5. Correspondence of tunnel properties and their prioritization between AA and CG methods in the studied enzymes from EC1−EC3 classes. A) Heatmaps with Pearson correlation coefficients between geometrical properties of tunnels present simultaneously in AA and each CG method across the whole data set. Data on the statistical significance of the observed correlations are available in Table S5. B) Performance of CG methods in ranking tunnels according to tunnel priority scores obtained with AA simulations.

stantial agreement between AA and CG simulations (Figure 5A), despite the considerable diversity of our enzymatic data set, which includes tunnels with varying lengths and bottleneck dimensions. This pattern parallels our findings with LinB variants, suggesting that CG methods can capture the structural features of tunnel ensembles, even though they exhibit method-specific biases in representing tunnel dynamics. This reinforces the conclusion that CG approaches can reliably map the geometric properties of transport pathways, even when their predicted occurrence frequencies require careful interpretation. Intriguingly, we observed stronger correlations among CG methods (Figure 5A), suggesting that distinct structural biases might still alter the sampling of tunnel ensembles in an analogous manner. Finally, we assessed the ranking of tunnels based on their priority scores on this diverse data set (Figure 5B). Similarly to LinB variants, we observed a moderate correlation of tunnel ranking with the AA results for the Elastic and Gō models. In contrast, with the SIRAH model, the extended data set showed only weak correlation with tunnel prioritization. However, we observed a notable improvement in the AP in tunnel ranking across all CG methods (0.68−0.78).

Overall, the observed ranking performance across varied protein architectures strongly supports the applicability of CG approaches for the initial enzyme tunnel discovery and characterization. In particular, the Gō model was the most proficient, in line with its above-discussed ability to reflect larger fluctuations found in AA simulations of EC data set enzymes. This consistent performance advantage in capturing dynamic characteristics aligns with our observations in LinB variants, reinforcing the conclusion that Gō-based restraints achieve the best balance between structural stability and conformational flexibility, making them particularly suited for tunnel analysis, while avoiding more pronounced failures in separating primary and transient tunnels, which were observed with the other two methods on the LinB data set. When evaluating the geometries of tunnel ensembles, the Elastic method exhibited the best correlation with both bottleneck radii and lengths. Such performance patterns observed across nine enzymes with diverse structural architectures and catalytic functions strongly support our findings from the LinB variants, establishing a solid foundation for selecting appropriate CG approaches based on the specific characteristics of interest in tunnel investigations.

CONCLUSIONS

This study represents the first comprehensive evaluation of CG methods for characterizing enzyme transport tunnels, demonstrating that selected CG approaches can accurately capture tunnel dynamics with substantial computational efficiency compared with AA simulations. Our benchmarking across three EC classes and diverse protein architectures establishes that the Martini-Gō model provides the best trade-off in reproducing both the structural and dynamic characteristics of enzyme tunnels, including the subtle effects of mutations on tunnel networks. Moreover, all benchmarked CG methods consistently showed good performance in detecting and structurally characterizing all known functional tunnels, in

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

agreement with AA simulations. While CG simulations identified additional tunnels not observed in AA simulations, these extra tunnels typically exhibited low occurrences, which did not interfere with the identification of functionally relevant pathways. Importantly, backmapping analyses revealed that most of these additional tunnels represent legitimate rare tunnel variants rather than artifacts of the CG representation.

Except for the Elastic model, the investigated CG methods could most often distinguish highly important known permanent tunnels from auxiliary transient tunnels in three LinB systems, and all CG methods exhibited high AP on the diverse data set of EC enzymes, suggesting their utility for tunnel prioritization. Also, CG methods were able to recapitulate key structural properties of corresponding tunnel ensembles, i.e., their bottleneck radii and lengths, as well as provide guidance in the detection of most of the key bottleneck-forming residues, at which the Elastic model excelled. Very promising results were achieved with the Gō model, which successfully captured mutation-induced alterations in tunnel dynamics in the LinB variants. Importantly, CG methods provided significant speedup for simulations and their analyses compared to AA methods, notably reducing sampling time while accessing more diverse tunnel networks.

Furthermore, as we observed the different trade-offs between the applicability of stabilizing networks used in the Martini approach (Elastic or Gō), it would be interesting to consider their further fine-tuning to optimize correspondence with insights from AA data.[30] While the generation of extensive AA simulations to optimize such stabilizing networks would undermine CG models' viability as a high-throughput screening tool for tunnel networks in diverse protein systems, we assume that even shorter AA simulations could be used for partial optimization, as was recently demonstrated with the optimization of an elastic network using much smaller AA models of microtubules.[112] Alternatively, AlphaFold confidence scores were successfully evaluated on 28 different proteins as a means to develop optimized elastic network parameters, eliminating the limiting requirement for AA simulations.[113] Finally, in this context, it will also be of interest to test the performance of the OLIVES approach,[114] a recent variant of Gō Martini that more directly mimics hydrogen bonding networks that stabilize protein structure, similar to the approach used here for stabilizing SIRAH simulations.

Taken together, CG models, especially the Martini Gō model, contain a promising approach for effectively investigating tunnels in proteins, opening new possibilities for identifying tunnels in large protein systems and massive data sets.

■ ASSOCIATED CONTENT

Data Availability Statement

The underlying data for this study are available in the published article, the Supporting Information, and in the Zenodo repository at 10.5281/zenodo.15488339.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.5c01727.

Detailed restraint parameters and validation data for coarse-grained simulations, secondary structure stability analysis, RMSD/RMSF comparisons, radius of gyration and solvent-accessible surface area measurements, comprehensive tunnel identification and occurrence statistics, tunnel property comparisons (bottleneck radius, length, occurrence), statistical analysis results, performance evaluation for tunnel ranking and bottleneck residue detection, scatter plot with correlations for EC1−EC3 enzyme classes, and distribution analyses of shared versus method-specific tunnels (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jan Brezovsky − *Laboratory of Biomolecular Interactions and Transport, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznań 61-614, Poland; International Institute of Molecular and Cell Biology in Warsaw, Warsaw 02-109, Poland;* orcid.org/0000-0001-9677-5078; Email: janbre@amu.edu.pl

Authors

Nishita Mandal − *Laboratory of Biomolecular Interactions and Transport, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznań 61-614, Poland; International Institute of Molecular and Cell Biology in Warsaw, Warsaw 02-109, Poland;* orcid.org/0000-0002-9114-0281

Jan A. Stevens − *Molecular Dynamics Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen 9747 AG, The Netherlands*

Adolfo B. Poma − *Biosystems and Soft Matter Division, Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw 02-106, Poland;* orcid.org/0000-0002-8875-3220

Bartlomiej Surpeta − *International Institute of Molecular and Cell Biology in Warsaw, Warsaw 02-109, Poland;* orcid.org/0000-0002-7436-8284

Carlos Sequeiros-Borja − *Laboratory of Biomolecular Interactions and Transport, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznań 61-614, Poland;* orcid.org/0000-0002-7817-0240

Aravind Selvaram Thirunavukarasu − *Laboratory of Biomolecular Interactions and Transport, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznań 61-614, Poland*

Siewert J. Marrink − *Molecular Dynamics Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen 9747 AG, The Netherlands;* orcid.org/0000-0001-8423-5277

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.5c01727

Author Contributions

N.M. set up the LinB-Wt and mutant AA models, LinB mutant CG models, and EC class 1−3 enzyme CG models. N.M. also performed all AA and CG simulations, analyzed all tunnels, prepared the figures, and drafted the manuscript. N.M. and B.S. devised the restraint protocol for the stabilization of the SIRAH CG model. J.A.S. contributed to the setup and optimization of the CG simulation protocol with the Martini Elastic model. A.B.P. helped set up and optimize the CG simulation protocol with the Gō-Martini model for LinB-Wt

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

and supervised its application to other systems. B.S. selected diverse EC2 class enzymes with suitable tunnel networks and performed their parametrization and optimization of their AA simulation protocol; performed analyses of tunnel prioritization and bottleneck residues' detection. A.S.T. selected diverse EC3 class enzymes with suitable tunnel networks and performed their parametrization and optimization of their AA simulation protocol. C.S.-B. selected diverse EC1 class enzymes with suitable tunnel networks and performed their parametrization and optimization of their AA simulation protocol. S.J.M. supervised the setup and optimization of the CG simulation protocol with the Martini Elastic model. N.M. and J.B. conceived the project and acquired the funding. J.B. coordinated the project and co-wrote the draft of the manuscript. N.M., B.S., and J.B. analyzed the data and interpreted the results. The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Centre, Poland (grant nos. 2023/49/N/NZ2/02567 to N.M., 2022/45/B/NZ1/02519 to A.B.P., and 2017/26/E/NZ1/00548 to J.B.). The training and research stay of N.M. at the Molecular Dynamics Group at the University of Groningen was cofunded by the Initiative of Excellence—Research University Program of Adam Mickiewicz University (grant no. ID-UB 048/13/UAM/0016). Computations by N.M., B.S., A.S.T., C.S.-B., and J.B. were performed at the Poznań Supercomputing and Networking Center. A.B.P. gratefully acknowledges the Polish high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within the computational grant no. PLG/2024/017332.

■ REFERENCES

(1) Prokop, Z.; Gora, A.; Brezovsky, J.; Chaloupkova, R.; Stepankova, V.; Damborsky, J. Engineering of Protein Tunnels: Keyhole-Lock-Key Model for Catalysis by the Enzymes with Buried Active Sites. In *Protein Engineering Handbook*. Bornscheuer, L.; ed.; Wiley-VCH: Weinheimpp, 2012, pp. 421−464.

(2) Gora, A.; Brezovsky, J.; Damborsky, J. Gates of Enzymes. *Chem. Rev.* **2013**, *113* (8), 5871−5923.

(3) Kingsley, L. J.; Lill, M. A. Substrate Tunnels in Enzymes: Structure−Function Relationships and Computational Methodology. *Proteins: Struct., Funct., Bioinf.* **2015**, *83* (4), 599−611.

(4) Marques, S. M.; Brezovsky, J.; Damborsky, J. Role of Tunnels and Gates in Enzymatic Catalysis. In *Understanding Enzymes: function, Design, Engineering, and Analysis*. Svendsen, A.; ed.; Pan Stanford Publishinged., 2016, pp. 421−463.

(5) Monzon, A. M.; Zea, D. J.; Fornasari, M. S.; Saldaño, T. E.; Fernandez-Alberti, S.; Tosatto, S. C. E.; Parisi, G. Conformational Diversity Analysis Reveals Three Functional Mechanisms in Proteins. *PLoS Comput. Biol.* **2017**, *13* (2), No. e1005398.

(6) Pravda, L.; Berka, K.; Svobodová Vařeková, R.; Sehnal, D.; Banáš, P.; Laskowski, R. A.; Koča, J.; Otyepka, M. Anatomy of Enzyme Channels. *BMC Bioinf.* **2014**, *15* (1), 379.

(7) Vavra, O.; Tyzack, J.; Haddadi, F.; Stourac, J.; Damborsky, J.; Mazurenko, S.; Thornton, J. M.; Bednar, D. Large-Scale Annotation of Biochemically Relevant Pockets and Tunnels in Cognate Enzyme−Ligand Complexes. *J. Cheminf.* **2024**, *16* (1), 114.

(8) Kokkonen, P.; Bednar, D.; Pinto, G.; Prokop, Z.; Damborsky, J. Engineering Enzyme Access Tunnels. *Biotechnol. Adv.* **2019**, *37* (6), 107386.

(9) Marques, S. M.; Daniel, L.; Buryska, T.; Prokop, Z.; Brezovsky, J.; Damborsky, J. Enzyme Tunnels and Gates As Relevant Targets in Drug Design. *Med. Res. Rev.* **2017**, *37* (5), 1095−1139.

(10) Porter, C. M.; Miller, B. G. Cooperativity in Monomeric Enzymes with Single Ligand-Binding Sites. *Bioorg. Chem.* **2012**, *43*, 44−50.

(11) Reed, M. C.; Lieb, A.; Nijhout, H. F. The Biological Significance of Substrate Inhibition: A Mechanism with Diverse Functions. *BioEssays* **2010**, *32* (5), 422−429.

(12) Biedermannová, L.; Prokop, Z.; Gora, A.; Chovancová, E.; Kovács, M.; Damborský, J.; Wade, R. C. A Single Mutation in a Tunnel to the Active Site Changes the Mechanism and Kinetics of Product Release in Haloalkane Dehalogenase LinB. *J. Biol. Chem.* **2012**, *287* (34), 29062−29074.

(13) Denisov, I. G.; Sligar, S. G. A Novel Type of Allosteric Regulation: Functional Cooperativity in Monomeric Proteins. *Arch. Biochem. Biophys.* **2012**, *519* (2), 91−102.

(14) Denisov, I. G.; Frank, D. J.; Sligar, S. G. Cooperative Properties of Cytochromes P450. *Pharmacol. Ther.* **2009**, *124* (2), 151−167.

(15) Wu, B. Substrate Inhibition Kinetics in Drug Metabolism Reactions. *Drug Metab. Rev.* **2011**, *43* (4), 440−456.

(16) Colletier, J.-P.; Fournier, D.; Greenblatt, H. M.; Stojan, J.; Sussman, J. L.; Zaccai, G.; Silman, I.; Weik, M. Structural Insights into Substrate Traffic and Inhibition in Acetylcholinesterase. *EMBO J.* **2006**, *25* (12), 2746−2756.

(17) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; Biedermannova, L.; Sochor, J.; Damborsky, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **2012**, *8* (10), No. e1002708.

(18) Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* **2009**, *5* (6), 1632−1639.

(19) Sarkar, D. K.; Surpeta, B.; Brezovsky, J. Incorporating Prior Knowledge in the Seeds of Adaptive Sampling Molecular Dynamics Simulations of Ligand Transport in Enzymes with Buried Active Sites. *J. Chem. Theory Comput.* **2024**, *20* (14), 5807−5819.

(20) Mandal, N.; Surpeta, B.; Brezovsky, J. Reinforcing Tunnel Network Exploration in Proteins Using Gaussian Accelerated Molecular Dynamics. *J. Chem. Inf. Model.* **2024**, *64* (16), 6623−6635.

(21) Laio, A.; Gervasio, F. L. Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science. *Rep. Prog. Phys.* **2008**, *71* (12), 126601.

(22) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187−199.

(23) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11* (8), 3584−3595.

(24) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; De Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The Power of Coarse Graining in Biomolecular Simulations. *WIREs Comput. Mol. Sci.* **2014**, *4* (3), 225−248.

(25) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.* **2016**, *116* (14), 7898−7936.

(26) Borges-Araújo, L.; Patmanidis, I.; Singh, A. P.; Santos, L. H. S.; Sieradzan, A. K.; Vanni, S.; Czaplewski, C.; Pantano, S.; Shinoda, W.; Monticelli, L.; Liwo, A.; Marrink, S. J.; Souza, P. C. T. Pragmatic Coarse-Graining of Proteins: Models and Applications. *J. Chem. Theory Comput.* **2023**, *19* (20), 7112−7135.

(27) Machado, M. R.; Barrera, E. E.; Klein, F.; Sóñora, M.; Silva, S.; Pantano, S. The SIRAH 2.0 Force Field: Altius, Fortius, Citius. *J. Chem. Theory Comput.* **2019**, *15* (4), 2719−2733.

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

(28) Periole, X.; Marrink, S.-J. The Martini Coarse-Grained Force Field. In *Biomolecular Simulations, Methods in Molecular Biology*, Monticelli, L.; Salonen, E.; Eds.; Humana Press: Totowa, NJ, 2013; Vol. *924*; pp. 533−565. DOI: .
(29) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18* (4), 382−388.
(30) Souza, P. C. T.; Thallmair, S.; Conflitti, P.; Ramírez-Palacios, C.; Alessandri, R.; Raniolo, S.; Limongelli, V.; Marrink, S. J. Protein−Ligand Binding with the Coarse-Grained Martini Model. *Nat. Commun.* **2020**, *11* (1), 3714.
(31) Nemchinova, M.; Schuurman-Wolters, G. K.; Whittaker, J. J.; Arkhipova, V.; Marrink, S. J.; Poolman, B.; Guskov, A. Exploring the Ligand Binding and Conformational Dynamics of the Substrate-Binding Domain 1 of the ABC Transporter GlnPQ. *J. Phys. Chem. B* **2024**, *128* (32), 7822−7832.
(32) Diamanti, E.; Souza, P. C. T.; Setyawati, I.; Bousis, S.; Monjas, L.; Swier, L. J. Y. M.; Shams, A.; Tsarenko, A.; Stanek, W. K.; Jäger, M.; Marrink, S. J.; Slotboom, D. J.; Hirsch, A. K. H. Identification of Inhibitors Targeting the Energy-Coupling Factor (ECF) Transporters. *Commun. Biol.* **2023**, *6* (1), 1182.
(33) Waclawiková, B.; Cesar Telles De Souza, P.; Schwalbe, M.; Neochoritis, C. G.; Hoornenborg, W.; Nelemans, S. A.; Marrink, S. J.; El Aidy, S. Potential Binding Modes of the Gut Bacterial Metabolite, 5-Hydroxyindole, to the Intestinal L-Type Calcium Channels and Its Impact on the Microbiota in Rats. *Gut Microbes* **2023**, *15* (1), 2154544.
(34) Bartocci, A.; Grazzi, A.; Awad, N.; Corringer, P.-J.; Souza, P. C. T.; Cecchini, M. A Millisecond Coarse-Grained Simulation Approach to Decipher Allosteric Cannabinoid Binding at the Glycine Receptor A1. *Nat. Commun.* **2024**, *15* (1), 9040.
(35) Periole, X.; Cavalli, M.; Marrink, S.-J.; Ceruso, M. A. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* **2009**, *5* (9), 2531−2543.
(36) Souza, P. C. T.; Borges-Araújo, L.; Brasnett, C.; Moreira, R. A.; Grünewald, F.; Park, P.; Wang, L.; Razmazma, H.; Borges-Araújo, A. C.; Cofas-Vargas, L. F.; Monticelli, L.; Mera-Adasme, R.; Melo, M. N.; Wu, S.; Marrink, S. J.; Poma, A. B.; Thallmair, S. GōMartini 3: From Large Conformational Changes in Proteins to Environmental Bias Corrections. *Nat. Commun.* **2025**, *16* (1), 4051.
(37) Cofas-Vargas, L. F.; Moreira, R. A.; Poblete, S.; Chwastyk, M.; Poma, A. B. The GōMartini Approach: Revisiting the Concept of Contact Maps and the Modelling of Protein Complexes. *Acta Phys. Polym. A* **2024**, *145* (3), S9−S20.
(38) Poma, A. B.; Cieplak, M.; Theodorakis, P. E. Combining the MARTINI and Structure-Based Coarse-Grained Approaches for the Molecular Dynamics Studies of Conformational Transitions in Proteins. *J. Chem. Theory Comput.* **2017**, *13* (3), 1366−1374.
(39) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \cdot \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(40) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577−8593.
(41) Klein, F.; Soñora, M.; Helene Santos, L.; Nazareno Frigini, E.; Ballesteros-Casallas, A.; Rodrigo Machado, M.; Pantano, S. The SIRAH Force Field: A Suite for Simulations of Complex Biological Systems at the Coarse-Grained and Multiscale Levels. *J. Struct. Biol.* **2023**, *215* (3), 107985.
(42) Koudelakova, T.; Chaloupkova, R.; Brezovsky, J.; Prokop, Z.; Sebestova, E.; Hesseler, M.; Khabiri, M.; Plevaka, M.; Kulik, D.; Kuta Smatanova, I.; Rezacova, P.; Ettrich, R.; Bornscheuer, U. T.; Damborsky, J. Engineering Enzyme Stability and Resistance to an Organic Cosolvent by Modification of Residues in the Access Tunnel. *Angew. Chem. Int. Ed.* **2013**, *52* (7), 1959−1963.
(43) Koudelakova, T.; Chovancova, E.; Brezovsky, J.; Monincova, M.; Fortova, A.; Jarkovsky, J.; Damborsky, J. Substrate Specificity of Haloalkane Dehalogenases. *Biochem. J.* **2011**, *435* (2), 345−354.
(44) Brezovsky, J.; Babkova, P.; Degtjarik, O.; Fortova, A.; Gora, A.; Iermak, I.; Rezacova, P.; Dvorak, P.; Smatanova, I. K.; Prokop, Z.; Chaloupkova, R.; Damborsky, J. Engineering a de Novo Transport Tunnel. *ACS Catal.* **2016**, *6* (11), 7597−7610.
(45) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating pK_a s and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368−W371.
(46) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5* (21), 3863−3871.
(47) Mermelstein, D. J.; Lin, C.; Nelson, G.; Kretsch, R.; McCammon, J. A.; Walker, R. C. Fast and Flexible Gpu Accelerated Binding Free Energy Calculations within the Amber Molecular Dynamics Package. *J. Comput. Chem.* **2018**, *39* (19), 1354−1358.
(48) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E. I.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K., et al. *AMBER 2018*; University of California: San Francisco, 2018.
(49) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.
(50) Zwanzig, R. Nonlinear Generalized Langevin Equations. *J. Stat. Phys.* **1973**, *9* (3), 215−220.
(51) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8* (5), 1542−1555.
(52) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864−1874.
(53) Sahil, M.; Sarkar, S.; Mondal, J. Long-Time-Step Molecular Dynamics Can Retard Simulation of Protein-Ligand Recognition Process. *Biophys. J.* **2023**, *122* (5), 802−816.
(54) Sahil, M.; Singh, T.; Ghosh, S.; Mondal, J. 3site Multisubstrate-Bound State of Cytochrome P450cam. *J. Am. Chem. Soc.* **2023**, *145* (43), 23488−23502.
(55) Sequeiros-Borja, C.; Surpeta, B.; Thirunavukarasu, A. S.; Dongmo Foumthuim, C. J.; Marchlewski, I.; Brezovsky, J. Water Will Find Its Way: Transport through Narrow Tunnels in Hydrolases. *J. Chem. Inf. Model.* **2024**, *64* (15), 6014−6025.
(56) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular Dynamics Simulations of Water and Biomolecules with a Monte Carlo Constant Pressure Algorithm. *Chem. Phys. Lett.* **2004**, *384* (4−6), 288−294.
(57) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684−3690.
(58) Basconi, J. E.; Shirts, M. R. Effects of Temperature Control Algorithms on Transport Properties and Kinetics in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9* (7), 2887−2899.
(59) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084−3095.
(60) Shimizu, H.; Park, S.-Y.; Shiro, Y.; Adachi, S. X-Ray Structure of Nitric Oxide Reductase (Cytochrome P450nor) at Atomic Resolution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58* (1), 81−89.
(61) Wilmouth, R. C.; Turnbull, J. J.; Welford, R. W. D.; Clifton, I. J.; Prescott, A. G.; Schofield, C. J. Structure and Mechanism of Anthocyanidin Synthase from Arabidopsis Thaliana. *Structure* **2002**, *10* (1), 93−103.

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

(62) Di Costanzo, L.; Drury, J. E.; Penning, T. M.; Christianson, D. W. Crystal Structure of Human Liver Δ4-3-Ketosteroid 5β-Reductase (AKR1D1) and Implications for Substrate Binding and Catalysis. *J. Biol. Chem.* **2008**, *283* (24), 16830−16839.
(63) Yousef, M. S.; Fabiola, F.; Gattis, J. L.; Somasundaram, T.; Chapman, M. S. Refinement of the Arginine Kinase Transition-State Analogue Complex at 1.2 Å Resolution: Mechanistic Insights. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58* (12), 2009−2017.
(64) Meng, G.; Fütterer, K. Structural Framework of Fructosyl Transfer in Bacillus Subtilis Levansucrase. *Nat. Struct. Mol. Biol.* **2003**, *10* (11), 935−941.
(65) Lee, K. A.; Fuda, H.; Lee, Y. C.; Negishi, M.; Strott, C. A.; Pedersen, L. C. Crystal Structure of Human Cholesterol Sulfotransferase (SULT2B1b) in the Presence of Pregnenolone and 3′-Phosphoadenosine 5′-Phosphate. *J. Biol. Chem.* **2003**, *278* (45), 44593−44599.
(66) Wang, H.; Liu, Y.; Hou, J.; Zheng, M.; Robinson, H.; Ke, H. Structural Insight into Substrate Specificity of Phosphodiesterase 10. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (14), 5782−5787.
(67) Crennell, S. J.; Garman, E. F.; Philippon, C.; Vasella, A.; Laver, G. W.; Vimr, E. R.; Taylor, G. L. The Structures of *Salmonella* typhimurium LT2 Neuraminidase and Its Complexes with Three Inhibitors at High Resolution. *J. Mol. Biol.* **1996**, *259* (2), 264−280.
(68) Lang, D.; Hofmann, B.; Haalck, L.; Hecht, H.-J.; Spener, F.; Schmid, R. D.; Schomburg, D. Crystal Structure of a Bacterial Lipase from *Chromobacterium viscosum* ATCC 6918 Refined at 1.6 Å Resolution. *J. Mol. Biol.* **1996**, *259* (4), 704−717.
(69) Shahrokh, K.; Orendt, A.; Yost, G. S.; Cheatham, T. E. Quantum Mechanically Derived AMBER-compatible Heme Parameters for Various States of the Cytochrome P450 Catalytic Cycle. *J. Comput. Chem.* **2012**, *33* (2), 119−133.
(70) Kaya, M.; Matsumura, K.; Higashida, K.; Hata, Y.; Kawato, A.; Abe, Y.; Akita, O.; Takaya, N.; Shoun, H. Cloning and Enhanced Expression of the Cytochrome P450nor Gene (*nicA*; *CYP55A5*) Encoding Nitric Oxide Reductase from *Aspergillus Oryzae*. *Biosci., Biotechnol., Biochem.* **2004**, *68* (10), 2040−2049.
(71) Davulcu, O.; Skalicky, J. J.; Chapman, M. S. Rate-Limiting Domain and Loop Motions in Arginine Kinase. *Biochemistry* **2011**, *50* (19), 4011−4018.
(72) Porras-Domínguez, J. R.; Ávila-Fernández, Á.; Miranda-Molina, A.; Rodríguez-Alegría, M. E.; Munguía, A. L. Bacillus Subtilis 168 Levansucrase (SacB) Activity Affects Average Levan Molecular Weight. *Carbohydr. Polym.* **2015**, *132*, 338−344.
(73) He, D.; Falany, C. N. Characterization of Proline-Serine-Rich Carboxyl Terminus in Human Sulfotransferase 2B1b: Immunogenicity, Subcellular Localization, Kinetic Properties, and Phosphorylation. *Drug Metab. Dispos.* **2006**, *34* (10), 1749−1755.
(74) Hoyer, L. L.; Roggentin, P.; Schauer, R.; Vimr, E. R. Purification and Properties of Cloned *Salmonella* Typhimurium LT2 Sialidase with Virus-Typical Kinetic Preference for Sialyl A2→3 Linkages1. *J. Biochem.* **1991**, *110* (3), 462−467.
(75) Maruyama, T.; Nakajima, M.; Kondo, H.; Kawasaki, K.; Seki, M.; Goto, M. Can Lipases Hydrolyze a Peptide Bond? *Enzyme Microb. Technol.* **2003**, *32* (6), 655−657.
(76) Luchko, T.; Gusarov, S.; Roe, D. R.; Simmerling, C.; Case, D. A.; Tuszynski, J.; Kovalenko, A. Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber. *J. Chem. Theory Comput.* **2010**, *6* (3), 607−624.
(77) Sindhikara, D. J.; Yoshida, N.; Hirata, F. *Placevent*: An Algorithm for Prediction of Explicit Solvent Atom Distribution-Application to HIV-1 Protease and F-ATP Synthase. *J. Comput. Chem.* **2012**, *33* (18), 1536−1543.
(78) Meyder, A.; Nittinger, E.; Lange, G.; Klein, R.; Rarey, M. Estimating Electron Density Support for Individual Atoms and Molecular Fragments in X-Ray Structures. *J. Chem. Inf. Model.* **2017**, *57* (10), 2437−2447.
(79) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16* (1), 528−552.
(80) Machado, M. R.; Pantano, S. SIRAH Tools: Mapping, Backmapping and Visualization of Coarse-Grained Models. *Bioinformatics* **2016**, *32* (10), 1568−1570.
(81) Machado, M. R.; Pantano, S. Split the Charge Difference in Two! A Rule of Thumb for Adding Proper Amounts of Ions in MD Simulations. *J. Chem. Theory Comput.* **2020**, *16* (3), 1367−1372.
(82) De Jong, D. H.; Liguori, N.; Van Den Berg, T.; Arnarez, C.; Periole, X.; Marrink, S. J. Atomistic and Coarse Grain Topologies for the Cofactors Associated with the Photosystem II Core Complex. *J. Phys. Chem. B* **2015**, *119* (25), 7791−7803.
(83) Kroon, P. C.; Grunewald, F.; Barnoud, J.; van Tilburg, M.; Brasnett, C.; Souza, P. C.; Wassenaar, T. A.; Marrink, S. J. Martinize2 and Vermouth provide a unified framework for molecular topology generation. *eLife* **2025**, *12*, RP90627.
(84) Mahmood, M. I.; Poma, A. B.; Okazaki, K.-I. Optimizing Gō-MARTINI Coarse-Grained Model for F-BAR Protein on Lipid Membrane. *Front. Mol. Biosci.* **2021**, *8*, 619381.
(85) Liu, Z.; Moreira, R. A.; Dujmović, A.; Liu, H.; Yang, B.; Poma, A. B.; Nash, M. A. Mapping Mechanostable Pulling Geometries of a Therapeutic Anticalin/CTLA-4 Protein Complex. *Nano Lett.* **2022**, *22* (1), 179−187.
(86) Gorelov, S.; Titov, A.; Tolicheva, O.; Konevega, A.; Shvetsov, A. DSSP in GROMACS: Tool for Defining Secondary Structures of Proteins in Trajectories. *J. Chem. Inf. Model.* **2024**, *64* (9), 3593−3598.
(87) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.
(88) Bernetti, M.; Bussi, G. Pressure Control Using Stochastic Cell Rescaling. *J. Chem. Phys.* **2020**, *153* (11), 114107.
(89) Ke, Q.; Gong, X.; Liao, S.; Duan, C.; Li, L. Effects of Thermostats/Barostats on Physical Properties of Liquids by Molecular Dynamics Simulations. *J. Mol. Liq.* **2022**, *365*, 120116.
(90) Sousa, F. M.; Lima, L. M. P.; Arnarez, C.; Pereira, M. M.; Melo, M. N. Coarse-Grained Parameterization of Nucleotide Cofactors and Metabolites: Protonation Constants, Partition Coefficients, and Model Topologies. *J. Chem. Inf. Model.* **2021**, *61* (1), 335−346.
(91) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(92) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109* (8), 1528−1532.
(93) Aurenhammer, F. Voronoi Diagrams-a Survey of a Fundamental Geometric Data Structure. *ACM Comput. Surv.* **1991**, *23* (3), 345−405.
(94) Dijkstra, E. W. A Note on Two Problems in Connexion with Graphs. *Numerische Math.* **1959**, *1*, 269−271.
(95) Pavelka, A.; Sebestova, E.; Kozlikova, B.; Brezovsky, J.; Sochor, J.; Damborsky, J. CAVER Algorithms for Analyzing Dynamics of Tunnels in Macromolecules. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2016**, *13* (3), 505−517.
(96) Sequeiros-Borja, C.; Surpeta, B.; Marchlewski, I.; Brezovsky, J. Divide-and-Conquer Approach to Study Protein Tunnels in Long Molecular Dynamics Simulations. *MethodsX* **2023**, *10*, 101968.
(97) Brezovsky, J.; Thirunavukarasu, A. S.; Surpeta, B.; Sequeiros-Borja, C. E.; Mandal, N.; Sarkar, D. K.; Dongmo Foumthuim, C. J.; Agrawal, N. TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport through Them. *Bioinformatics* **2022**, *38* (6), 1752−1753.
(98) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **2014**, *10* (2), 676−690.

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150

(99) Klvana, M.; Pavlova, M.; Koudelakova, T.; Chaloupkova, R.; Dvorak, P.; Prokop, Z.; Stsiapanava, A.; Kuty, M.; Kuta-Smatanova, I.; Dohnalek, J.; Kulhanek, P.; Wade, R. C.; Damborsky, J. Pathways and Mechanisms for Product Release in the Engineered Haloalkane Dehalogenases Explored Using Classical and Random Acceleration Molecular Dynamics Simulations. *J. Mol. Biol.* **2009**, *392* (5), 1339−1356.
(100) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J. Redesigning Dehalogenase Access Tunnels as a Strategy for Degrading an Anthropogenic Substrate. *Nat. Chem. Biol.* **2009**, *5* (10), 727−733.
(101) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261−272.
(102) Shapiro, S. S.; Wilk, M. B. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* **1965**, *52* (3−4), 591−611.
(103) Glaser, R. E. Levene's Robust Test of Homogeneity of Variances. In *Encyclopedia of Statistical Sciences*, Kotz, S.; Read, C. B.; Balakrishnan, N.; Vidakovic, B., Eds.; Wiley, 2006. DOI: .
(104) Kruskal, W. H.; Wallis, W. A. Use of Ranks in One-Criterion Variance Analysis. *J. Am. Stat. Assoc.* **1952**, *47* (260), 583−621.
(105) Mann, H. B.; Whitney, D. R. On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other. *Ann. Math. Statist.* **1947**, *18* (1), 50−60.
(106) Armstrong, R. A. When to Use the Bonferroni Correction. *Ophthalmic Physiol. Opt.* **2014**, *34* (5), 502−508.
(107) Brezovsky, J.; Chovancova, E.; Gora, A.; Pavelka, A.; Biedermannova, L.; Damborsky, J. Software Tools for Identification, Visualization and Analysis of Protein Tunnels and Channels. *Biotechnol. Adv.* **2013**, *31* (1), 38−49.
(108) Javanainen, M.; Martinez-Seara, H.; Vattulainen, I. Excessive Aggregation of Membrane Proteins in the Martini Model. *PLoS One* **2017**, *12* (11), No. e0187936.
(109) Stark, A. C.; Andrews, C. T.; Elcock, A. H. Toward Optimized Potential Functions for Protein−Protein Interactions in Aqueous Solutions: Osmotic Second Virial Coefficient Calculations Using the MARTINI Coarse-Grained Force Field. *J. Chem. Theory Comput.* **2013**, *9* (9), 4176−4185.
(110) Majumder, A.; Straub, J. E. Addressing the Excessive Aggregation of Membrane Proteins in the MARTINI Model. *J. Chem. Theory Comput.* **2021**, *17* (4), 2513−2521.
(111) McDonald, A. G.; Tipton, K. F. Enzyme Nomenclature and Classification: The State of the Art. *FEBS J.* **2023**, *290* (9), 2214−2231.
(112) Sahoo, A.; Hanson, S. M. Microtubules in Martini: Parameterizing a Heterogeneous Elastic-Network towards a Mechanically Accurate Microtubule. *PNAS Nexus* **2025**, *4* (7), pgaf202.
(113) Jussupow, A.; Kaila, V. R. I. Effective Molecular Dynamics from Neural Network-Based Structure Prediction Models. *J. Chem. Theory Comput.* **2023**, *19* (7), 1965−1975.
(114) Pedersen, K. B.; Borges-Araújo, L.; Stange, A. D.; Souza, P. C. T.; Marrink, S. J.; Schiøtt, B. OLIVES: A Gō-like Model for Stabilizing Protein Structure via Hydrogen Bonding Native Contacts in the Martini 3 Coarse-Grained Force Field. *J. Chem. Theory Comput.* **2024**, *20* (18), 8049−8070.

https://doi.org/10.1021/acs.jctc.5c01727
J. Chem. Theory Comput. 2026, 22, 135−150