



Review

Dynamic prediction of protein binding sites: Integrating structural biology, molecular simulation, solvent probe mapping, and artificial intelligence

Qiu Yue*

Jinhua Key laboratory of Quality Evaluation and Standard Research of Traditional Chinese Medicine, Jinhua Institute For Food And Drug Control, Jinhua Advanced Research Institute of Technology, Jinhua 321019, China

* **Correspondence:** Email: sean28299@gmail.com; Tel: +8615631635591.

Abstract: Identifying protein binding sites is essential for mechanistic biology and structure-based drug discovery. However, single static structures often fail to capture functionally relevant sites that are shallow, transient, allosteric, or cryptic. This review focuses primarily on protein binding-site prediction and presents a unified biophysical framework emphasizing conformational dynamics, thermodynamics, and solvent-mediated recognition. We evaluated current experimental and computational approaches, with particular attention to molecular dynamics, enhanced sampling, Markov state models, and mixed-solvent molecular dynamics for identifying transient cavities and chemically favorable hotspots across conformational ensembles. Rather than relying on single-score predictions, we propose a multi-factor evidence framework that integrates pocket emergence, probe enrichment, local chemistry, reproducibility, and functional or ligand-based support. RNA binding regions and macromolecular interfaces are discussed as important extensions, where related principles of conformational heterogeneity, electrostatics, hydration, and molecular recognition also apply. Finally, we examine the emerging role of ensemble-aware artificial intelligence in integrating structural, dynamic, thermodynamic, and experimental information, and highlight the need for interpretable, uncertainty-aware, and prospectively validated prediction strategies.

Keywords: macromolecular binding sites; cryptic pockets; allosteric sites; molecular dynamics; druggability; geometric deep learning; structure-based drug design

1. Introduction

1.1. *Biological significance of macromolecular binding sites*

Macromolecular function is organized around sites at which molecules recognize one another. Enzyme active sites position substrates and catalytic residues; orthosteric sites transmit endogenous signals; allosteric sites reshape conformational populations; protein–protein and protein–nucleic acid interfaces assemble molecular machines; and structured RNA pockets bind metabolites, ions, proteins, and regulatory small molecules. Accurate site identification affects nearly every stage of structure-guided discovery, from target nomination and functional annotation to fragment screening, virtual screening, lead optimization, resistance analysis, and the design of chemical probes.

The central difficulty is that a biologically meaningful site is not simply a concavity on a molecular surface. It is a spatially and temporally organized microenvironment that can stabilize a binding partner under relevant solution conditions and, in many cases, transmit a functional effect. Geometry is necessary but rarely sufficient. A cavity may be deep but chemically unfavorable; an exposed patch may lack enclosure but contain a compact energetic hotspot; and a region absent from an apo crystal structure may become ligandable after a modest loop or side-chain rearrangement. Consequently, binding-site prediction has expanded from detecting static surface depressions to estimating which conformational states can support binding, how frequently those states occur, and whether occupation will alter macromolecular function. The dynamic and state-dependent nature of binding pockets has been extensively documented, particularly for transient and cryptic sites whose accessibility and druggability can vary substantially across conformational ensembles [1,2].

1.2. *Why single-structure prediction is insufficient*

Experimental coordinates are indispensable, but every structure is conditional. Crystal packing can stabilize one arrangement and suppress another. Cryogenic data collection may narrow one conformational distribution. Flexible loops can be unresolved, engineered constructs can shift equilibria, and apo structures may omit cavities stabilized only by ligands, partner proteins, membrane lipids, ions, or particular protonation states. Even high-resolution holo structures preferentially reveal states that crystallize and may not report the population or lifetime of the pocket in solution.

Cryptic sites provide the clearest counterexample to a single-structure view. They are absent or insufficiently formed in an available unbound structure but appear in a ligand-bound or otherwise perturbed conformation [2,3]. Transient pockets are related but need not have an experimentally resolved ligand-bound state; they may open and close spontaneously on a simulation-accessible timescale. Allosteric sites add another layer because functional relevance depends not only on local ligandability but also on coupling to distant catalytic, recognition, or assembly coordinates. Protein binding pockets therefore fluctuate in volume, enclosure, hydration, polarity, and accessibility rather than switching between two idealized structures [1].

1.3. *Site detection, ligandability, and druggability*

Three questions should be kept distinct throughout a prediction workflow.

1. Site detection: Where is a geometrically or energetically plausible interaction region?

2. Ligandability: Can the region bind a defined class of molecules with measurable and specific affinity?
3. Druggability: Can it bind a molecule with the affinity, selectivity, physicochemical properties, and functional consequences required for a viable therapeutic program?

Ligandability is ligand-class dependent, whereas druggability is also development-context dependent. A charged RNA groove may bind aminoglycosides or other cationic ligands yet still be challenging to target selectively because similar electrostatic and structural features can occur across multiple RNA regions. Such a site may therefore be ligandable while remaining difficult to develop into a selective drug target with favorable pharmacological properties. A shallow protein–protein interface can contain a high-value hotspot despite scoring poorly as a conventional small-molecule pocket. Conversely, a high SiteScore or probe density does not guarantee favorable cellular exposure, selectivity, or functional modulation. The concept of the druggable genome [4], fragment-based estimates of target druggability [5], and pocket descriptors used in tools such as SiteMap [6] all reinforce the need to state which level of inference a score supports.

1.4. Scope and organizing principle

This review primarily focuses on protein binding sites while considering RNA binding regions and macromolecular interfaces as essential extensions of the same biophysical principle: molecular recognition is governed by the interplay between structural organization, conformational dynamics, physicochemical complementarity, and functional regulation. Although proteins represent the most extensively studied macromolecular systems for binding site prediction, many fundamental concepts, including transient pocket formation, electrostatic recognition, solvent-mediated interactions, and conformational selection, are also applicable to RNA molecules and protein–protein or protein–nucleic acid interfaces. Therefore, we discuss binding site prediction from a unified perspective that integrates structural, dynamic, and functional information. This article was prepared as a narrative review rather than a systematic review or meta-analysis. Relevant literature was identified primarily through PubMed. The literature surveyed extended through 2026, with earlier seminal studies included where necessary to establish the historical and methodological foundations of the field. Representative studies were selected based on methodological relevance, influence on field development, experimental or independent validation (where available), and their relevance to the structural, dynamic, thermodynamic, and data-driven aspects of binding-site prediction.

The organizing principle of this review follows a conceptual transition from static identification toward dynamic, functional, and multimodal prediction (Figure 1). Static structural approaches initially identify candidate binding regions based on geometric properties, physicochemical features, evolutionary conservation, and structure-based learning models. Because macromolecules populate conformational ensembles, molecular dynamics simulations, enhanced sampling, and ensemble analyses can subsequently identify transient or low-population binding-competent states through dynamic pocket discovery or emergence, including spontaneous pocket opening, conformational selection, and ligand- or probe-stabilized conformations. Solvent-probe mapping and mixed-solvent molecular dynamics can further provide chemically resolved information on favorable interaction regions, while kinetic analyses such as Markov state models characterize the populations and transitions of relevant conformational states. Multimodal artificial intelligence may integrate structural, dynamic, solvent-derived, evolutionary, and experimental information. These components do not

constitute a universally sequential pipeline and may instead be applied iteratively, in parallel, or in different orders depending on the biological question and available data.

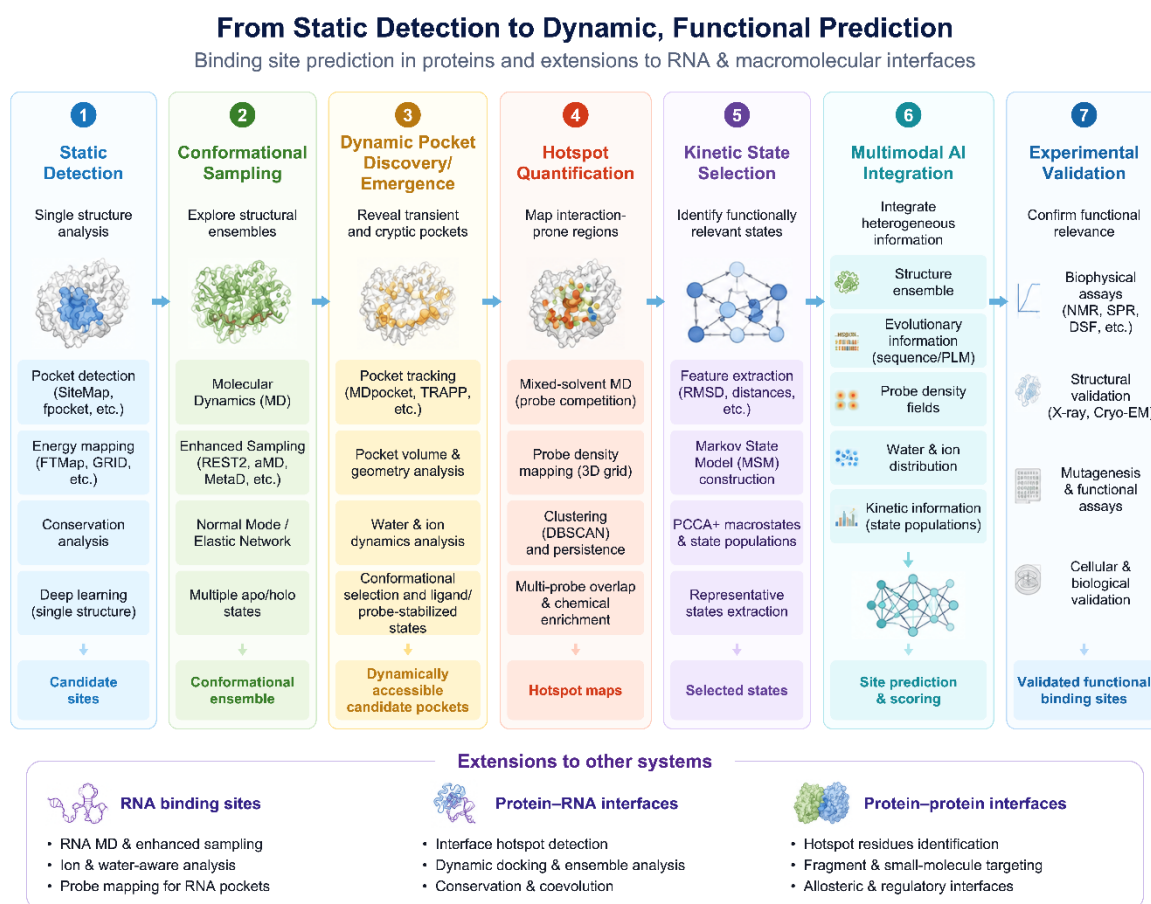


Figure 1. Conceptual framework for the evolution of binding-site prediction from static structural analysis toward dynamic pocket discovery or emergence, solvent- and hotspot-based characterization, kinetic-state analysis, multimodal artificial intelligence, and experimental validation. The illustrated stages represent complementary components of a possible workflow and may be applied iteratively or in different orders depending on the biological system and available data.

2. Biophysical basis of macromolecular binding sites

2.1. Binding-site classes are overlapping functional descriptions

Orthosteric sites bind endogenous ligands or substrates at the primary recognition region. Catalytic sites are defined by chemical transformation and may overlap an orthosteric pocket. Allosteric sites bind away from the primary site and alter function through conformational or dynamic coupling. Cryptic sites are defined operationally by their absence from an available unbound structure and emergence in another state. Transient pockets open only intermittently, while protein–protein and protein–nucleic acid hotspots describe localized energetic contributions within interfaces that are often

larger and flatter than small-molecule pockets. RNA–small molecule sites combine grooves, junctions, bulges, base triples, metal-ion coordination, and structured hydration [7,8].

These labels are not mutually exclusive. A cryptic pocket can be allosteric, an orthosteric site can be transient, and a protein–protein hotspot can be targeted by a fragment or macrocycle. Classification is most useful when it records the evidence supporting each descriptor (Table 1): structural location, state dependence, binding partner, functional consequence, and mode of validation.

Table 1. Classification of macromolecular binding sites and characteristic evidence.

Site class	Defining feature	Typical biophysical signature	Main prediction challenge	Strong validation
Orthosteric	Binds the primary endogenous ligand or substrate	Geometric complementarity and conserved recognition chemistry	Distinguishing functional from adventitious cavities	Holo structure, competition, or catalytic kinetics
Catalytic	Supports bond making or breaking	Catalytic residues, electrostatic preorganization, water or metal control	Catalysis may occur in a shallow or highly polar site	Mutagenesis plus kinetic mechanism
Allosteric	Occupation changes a distant function	Dynamic coupling, population shift, or network communication	Local ligandability does not prove functional coupling	Ligand binding plus distant functional response
Cryptic	Pocket is absent or poorly formed in an unbound reference	Ligand- or solvent-stabilized opening	Rare-state sampling and false positives	Apo-holo comparison or experimentally confirmed ligand
Transient	Pocket opens intermittently	Time-dependent volume, enclosure, and hydration	Dependence on sampling length and state definition	Reproducibility across trajectories and experimental support
Protein–protein hotspot	Small subset of interface contributes disproportionately to affinity	Aromatic, hydrophobic, or electrostatic clusters on broad surfaces	Weak enclosure and partner dependence	Alanine scanning, fragment binding, or complex structure
Protein–nucleic acid interface	Recognizes DNA or RNA	Electrostatics, base or backbone readout, water and ion mediation	Large flexible interfaces and sequence dependence	Complex structure, crosslinking, or mutagenesis
RNA–small molecule pocket	Structured RNA region binds a small molecule	Base stacking, electrostatics, metal ions, and conformational stabilization	Data scarcity, high flexibility, and force-field sensitivity	NMR, crystallography, cryo-EM, chemical probing, or binding assay

2.2. Geometric determinants

Pocket volume, depth, enclosure, mouth area, curvature, compactness, surface exposure, and the continuity of channels or tunnels establish whether a molecular fragment can enter and remain in a site. Deep enclosed pockets generally reduce translational freedom and protect contacts from bulk solvent, but no universal geometric cutoff defines binding. Large cavities may contain several subpockets with distinct chemistry, while narrow tunnels can be functionally important despite low volume. At macromolecular interfaces, local concavity and protrusion complementarity can be more informative than whole-interface area.

Geometry is state dependent. Backbone shifts of less than an angstrom, side-chain rotamer changes, or displacement of one ordered water can connect two subcavities or transform an exposed groove into an enclosed pocket. These changes help explain why coarse global measures such as whole-protein RMSD may appear stable while site-level druggability changes substantially.

2.3. Physicochemical, hydration, and energetic determinants

Binding sites concentrate chemically complementary features: hydrophobic patches, hydrogen-bond donors and acceptors, aromatic surfaces, cationic or anionic regions, metal-coordination motifs, and polarizable atoms. Productive binding requires a balance. A highly polar pocket may offer many possible hydrogen bonds but impose a prohibitive desolvation cost if those interactions are geometrically unsatisfied. A hydrophobic cavity may bind fragments readily yet suffer from poor selectivity if it lacks directional interactions. Water is an active component of the site. Conserved waters can bridge protein–ligand contacts, stabilize pocket shape, and complete coordination networks, whereas thermodynamically unfavorable waters can provide a driving force for ligand binding when displaced [9]. Ions similarly alter local electrostatics and can organize RNA folds, catalytic sites, and membrane-associated pockets. These solvent and ion effects are particularly important because hydration structure and metal-ion distributions can reshape local conformational stability, electrostatic organization, and the energetic cost of ligand recognition [8,10]. A binding-site model that treats all solvent as featureless misses both opportunities and penalties. Solvent-probe mapping is powerful partly because it measures competition against water in an explicit dynamic environment.

From a thermodynamic perspective, favorable local interactions must be considered together with the energetic cost of reorganizing both the ligand and its environment. Ligand binding generally involves competition among direct macromolecule–ligand interactions, ligand and site desolvation, solvent reorganization, and changes in conformational and configurational entropy. Hydration can therefore either favor or oppose binding. Displacement of thermodynamically unfavorable or weakly coordinated water may contribute favorably to ligand recognition, whereas removal of strongly stabilized water can impose a substantial desolvation penalty, unless compensating interactions are formed. Consequently, geometric complementarity or locally favorable interaction energies alone cannot establish favorable binding thermodynamics.

2.4. Conformational, thermodynamic, and kinetic determinants

Induced fit and conformational selection describe complementary limits rather than mutually exclusive mechanisms. In induced fit, binding drives structural adaptation after initial encounter [11].

In conformational selection, a ligand preferentially binds a pre-existing, perhaps low-population state and shifts the ensemble [12,13]. Many real systems combine both processes: a ligand selects a permissive state and then induces local refinement. Protein dynamics span side-chain motions, loop breathing, hinge displacement, domain rearrangement, partial unfolding, and changes in oligomeric state [14]. The relevant state must be not only energetically accessible but kinetically reachable on the timescale of binding.

Conformational populations also contribute directly to binding thermodynamics. When a ligand recognizes a low-population open or binding-competent state, formation of that state carries a conformational free-energy cost relative to the dominant unbound ensemble. Thus, even a conformation exhibiting highly favorable local ligand interactions may contribute weakly to overall binding if it is rarely populated. Conversely, stabilization of a pre-existing accessible state can shift the conformational equilibrium and increase its population. The free-energy landscape of pocket opening should therefore be distinguished from the interaction free energy of a ligand within an already formed pocket, although both contribute to the observed binding process.

3. Experimental structural and biophysical approaches

3.1. Structural determination and conformational characterization

X-ray crystallography remains a primary source of atomic binding-site evidence and underlies curated structural databases [15]. Holo structures define direct contacts, pocket shape, ordered solvent, and alternative ligand poses. Apo-holo series can reveal induced fit, cryptic opening, or side-chain reorganization. Fragment crystallography is particularly valuable because weak fragments sample compact interaction motifs and can expose secondary or allosteric sites that are missed by high-affinity ligands. Interpretation must account for crystal packing, engineered mutations, truncations, cryoprotection, occupancy refinement, and construct-dependent conformations. A ligand observed at high concentration may occupy an adventitious site, while a real site can remain invisible if the corresponding state is incompatible with the lattice. Structural evidence is therefore strongest when combined with affinity, competition, or functional data.

3.2. Cryogenic electron microscopy

Cryo-EM is increasingly important for large assemblies, membrane proteins, and systems that occupy multiple conformational states [16]. Focused classification can separate ligand-bound and ligand-free populations, while local resolution and map quality determine whether small molecules or side-chain rearrangements can be assigned confidently. Cryo-EM can reveal interdomain and intersubunit pockets that do not exist in isolated domains, but conformational heterogeneity and model bias must be considered when interpreting weak density as a ligand or water network.

3.3. Nuclear magnetic resonance

NMR provides solution-state information across a broad range of timescales. Chemical shift perturbation maps ligand-sensitive residues; relaxation and dispersion identify exchange between low- and high-population states; paramagnetic probes report solvent accessibility and long-range distances;

and ligand-observed methods detect weak fragment binding [17]. NMR is especially valuable for shallow, transient, or RNA sites because it can detect binding without requiring a rigid, crystallizable complex. Its limitations include sample size, spectral overlap, molecular-weight constraints, and the fact that a perturbation may reflect an allosteric response rather than direct contact.

3.4. Hydrogen-deuterium exchange mass spectrometry and related methods

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) reports changes in solvent protection and backbone dynamics upon ligand binding, mutation, or complex formation [18]. It can identify protected regions, reveal distal allosteric effects, and compare conformational states under near-solution conditions. Spatial resolution is generally peptide-level unless advanced fragmentation is used, and changes can reflect either direct occlusion or altered dynamics. Crosslinking MS, native MS, ion mobility, and limited proteolysis provide complementary evidence for complex topology, stoichiometry, conformational distributions, and stabilization [19].

3.5. Biophysical screening and chemical capture

Fragment screening by crystallography, NMR, surface plasmon resonance (SPR), thermal shift assays, microscale thermophoresis, or native MS provides orthogonal tests of ligandability [19–21]. Photoaffinity labeling and covalent fragment approaches can capture transient interactions and localize sites by MS, although reactive-group selectivity and nonspecific labeling require controls [22]. No single assay can establish a binding site unambiguously. A reliable experimental strategy therefore integrates complementary evidence, including direct binding measurements, spatial localization, functional responses, and perturbation approaches such as mutagenesis. Table 2 summarizes representative experimental approaches for macromolecular binding site characterization, highlighting their primary inputs, capacity to capture dynamic information, ability to detect cryptic sites, throughput limitations, and complementary strengths and weaknesses.

Table 2. Comparison of experimental and computational site-prediction approaches.

Approach	Primary input	Dynamic information	Cryptic-site capability	Cost or throughput	Main strength	Main limitation
X-ray crystallography	Crystals, diffraction	Low to moderate through structural series	Moderate	Low throughput	Atomic contacts and waters	Crystal and construct bias
Cryo-EM	Vitrified particles	Moderate through classification	Moderate	Low to moderate	Large complexes and membrane systems	Local resolution and heterogeneity

Continued on next page

Approach	Primary input	Dynamic information	Cryptic-site capability	Cost or throughput	Main strength	Main limitation
NMR	Solution sample	High across selected timescales	High	Low throughput	Weak and transient binding	Size and spectral complexity
HDX-MS	Solution exchange and MS	Moderate	Indirect	Moderate	Conformational and allosteric changes	Limited spatial resolution
Fragment screening	Protein plus fragment library	Assay dependent	Moderate to high	Moderate	Direct ligandability evidence	Weak affinity and assay artifacts
Static pocket detection	One or more structures	None unless repeated over an ensemble	Low	Very high	Fast, interpretable localization	Misses state-dependent sites
MD and trajectory pockets	Force field and starting structure	High	Moderate to high	Moderate to high	Site opening, hydration, and state dependence	Sampling and force-field dependence
MSMD or cosolvent mapping	Structure, water, organic probes	High	High	High	Hotspot chemistry plus induced opening	Probe and concentration dependence
Deep or geometric learning	Sequence, structure, surface, or graph	Usually implicit	Variable	Very high after training	Scalable feature learning	Dataset bias and calibration

4. Static computational prediction of binding sites

4.1. Geometry-based pocket detection

Geometry-based methods identify cavities without requiring a known ligand. SURFNET fills gaps with spheres [23]; CASTp characterizes pockets and cavities using computational geometry [24]; fpocket uses Voronoi tessellation and alpha spheres [25]; and DoGSiteScorer applies difference-of-Gaussian filters and druggability descriptors [26]. SiteMap combines geometric and physicochemical fields to identify and characterize candidate sites [6,27]. These approaches are fast enough for proteome-scale or ensemble screening and provide interpretable descriptors such as volume, enclosure, exposure, and hydrophobicity. The main limitation is representational: a method can score only the conformation supplied. Shallow interfaces and closed cryptic sites may be absent, whereas packing defects, clefts between domains, or exposed grooves can be overpredicted. Consensus across

independent algorithms reduces idiosyncratic failures but does not convert a static model into a dynamic one.

4.2. Static energetic and probe mapping

GRID evaluates interaction energies of chemical probes at spatial points around a target [28]. FTMap docks a panel of small organic probes and clusters favorable positions to identify consensus hotspots [29]. Fragment hotspot maps and SiteMap-like energetic descriptors similarly estimate where hydrophobic, donor, acceptor, or charged interactions are favored. These methods add chemical information to geometry and can identify hotspots on comparatively flat surfaces. Static mapping nevertheless depends on the input conformation and simplified solvent treatment. It can rank a preformed pocket well but may miss the energetic consequences of pocket opening, water rearrangement, or loop motion. For this reason, static maps are well suited to hypothesis generation and cross-validation of dynamic hotspots.

4.3. Evolutionary and template information

Conserved residues, coevolution, catalytic motifs, and family-specific sequence patterns can indicate functional regions even when geometric evidence is weak. Tools such as ConSurf map evolutionary conservation onto structures [30]. Template methods transfer sites from homologous or structurally similar proteins, while apo-holo comparisons identify recurring conformational responses. These approaches are powerful when reliable relatives exist but can fail for novel folds, divergent function within a family, or allosteric sites that are poorly conserved. Evolutionary information also requires interpretation. Catalytic residues are often conserved, whereas druggable allosteric pockets may be lineage-specific. A lack of conservation does not invalidate a ligandable site, and strong conservation can reflect structural stability rather than binding. The practical value of evolutionary and template-based information has been demonstrated in several settings. ConSurf, for example, used phylogenetic conservation mapped onto three-dimensional structures to identify functionally important surface regions in SH2 and PTB signaling domains, illustrating how conserved surface patches can help prioritize protein–protein and ligand-recognition sites [31]. Template-based approaches can extend this concept when close sequence identity is unavailable. FINDSITE identifies ligand-binding sites by comparing binding-site similarity among weakly homologous template structures obtained through threading, retaining substantial predictive accuracy even when approximate protein models and low-sequence-identity templates were used [32]. These applications illustrate how evolutionary conservation and structural templates can provide experimentally actionable hypotheses when direct ligand-bound structures are unavailable.

4.4. Machine learning and deep learning on static structures

P2Rank predicts the ligandability of points on the solvent-accessible surface using local geometric and physicochemical features and clusters high-scoring points into sites [33]. DeepSite reframed pocket detection as three-dimensional image analysis, using voxelized atom-property channels and a 3D convolutional neural network [34]. Surface and graph representations reduce dependence on rigid Cartesian grids. MaSIF learns molecular-surface interaction fingerprints [35],

ScanNet learns interpretable atom- and residue-scale geometric patterns [36], and PeSTo predicts protein interaction interfaces directly from atomic coordinates [37]. PocketMiner differs conceptually because it was trained to predict where cryptic pockets are likely to open in MD simulations even when only one structure is provided [38]. Ligand-aware approaches such as LABind explicitly encode both protein and ligand information, allowing the predicted site to depend on the intended ligand class [39]. These models expand representational capacity but inherit the structural and labeling biases of their training data. Holo structures leak ligand-shaped geometry into inputs; homologous proteins can appear across train and test sets; and residue imbalance can inflate receiver-operating-characteristic metrics. Predictions on modeled or apo structures are usually harder than predictions on cognate holo structures. A static neural network can learn the likelihood that a region could open, but unless dynamics are supplied explicitly, it cannot directly estimate open-state population, lifetime, or solvent reorganization.

4.5. Practical role of static consensus

Static prediction is most useful as a rapid first layer. Multiple structures should be analyzed when available, including apo, holo, mutant, homologous, cryo-EM, and predicted models. Candidate sites can then be ranked by spatial consensus, conservation, chemical complementarity, and known functional data. In the PLpro work, SiteMap, the LigBuilder V3 cavity module, and DeepSite recovered known regions while also proposing thumb, finger, and back-surface candidates [40]. Differences among conformations were themselves informative because apparently small structural changes altered which sites were detected.

Independent systems further demonstrate why static consensus should generally be followed by dynamic analysis. Bowman and Geissler used extensive molecular simulations and Markov state models of β -lactamase, interleukin-2, and RNase H to identify transient cryptic allosteric pockets arising from equilibrium conformational fluctuations, including bound-like conformations that were not evident from a single static structure [41]. These findings illustrate that static pocket predictors can provide candidate locations, whereas conformational ensembles are required to determine whether apparently absent or poorly formed sites become accessible through intrinsic dynamics. Static consensus should therefore be regarded as a hypothesis-generation and prioritization layer rather than a final assignment of binding-site accessibility. Representative software and computational strategies for binding-site prediction, together with their representations, dynamic inputs, major outputs, applications, and key limitations, are summarized in Table 3.

Table 3. Representative software and computational strategies.

Method	Representation	Dynamic input	Main output	Best use	Important caveat
CASTp	Surface geometry	No	Pockets, cavities, area, volume	Structural annotation	No direct ligandability estimate
fpocket	Voronoi and alpha spheres	No, but can be applied framewise	Ranked pockets and descriptors	Fast screening	Sensitive to conformation and thresholds

Continued on next page

Method	Representation	Dynamic input	Main output	Best use	Important caveat
SiteMap	Geometry plus interaction fields	No	SiteScore, Dscore, site points	Pocket detection and druggability	Proprietary and static
FTMap	Docked organic probes	No	Consensus energetic hotspots	Fragment hotspot mapping	Limited conformational adaptation
P2Rank	Surface points plus machine learning	No	Ranked pockets and residues	High-throughput prediction	Training-set dependence
DeepSite	3D voxel CNN	No	Binding-site probability map	Static site detection	Grid cost and holo bias
MaSIF	Molecular surface patches	No	Interaction fingerprints	Ligand and interface recognition	Surface construction choices
PocketMiner	Graph neural network trained on MD labels	Implicit in labels	Cryptic-pocket propensity	Rapid cryptic-site triage	Does not replace state sampling
MDpocket	Framewise pocket geometry	Yes	Pocket frequency and volume	Dynamic pocket monitoring	Depends on trajectory coverage
TRAPP	Pocket grids and ML descriptors	Yes	Transient-pocket properties and druggability	Comparing site states	Site definition and model calibration
MixMD / MSMD	Explicit water and organic probes	Yes	Probe densities and induced pockets	Hotspots and cryptic sites	Probe chemistry, concentration, aggregation
SILCS	Competitive fragment mapping	Yes	FragMaps and grid free energies	Functional group mapping	Protocol and force-field dependence
Probe GridMap Builder	CPPTRAJ grid occupancy plus DBSCAN	Yes	Ranked density clusters and maps	Reproducible hotspot prioritization	Density is not binding free energy

5. Dynamic binding site prediction through molecular dynamics and solvent probe mapping

5.1. Conventional MD as a pocket-discovery instrument

Atomistic MD connects a structural model to a time-ordered ensemble governed by a force field and thermodynamic conditions [42]. For binding-site analysis, the most useful observables are local: pocket volume, mouth opening, enclosure, hydration, side-chain rotamers, loop distances, residue contacts, probe or ligand residence, and communication with distal regions. Global RMSD is valuable

for quality control but is not a site metric. A stable global fold can contain strongly varying local pockets, and a flexible domain can move without changing the chemistry of a particular site. Conventional MD can reveal spontaneous opening and closing, channel gating, water exchange, and alternative hydrogen-bond networks. The probability of observing a rare pocket, however, depends on the free-energy barrier, total aggregate sampling, number of independent replicas, initial conditions, and the sensitivity of the analysis. A single long trajectory and several shorter replicas answer different questions. Replicas provide evidence of reproducibility and reduce dependence on one stochastic path, whereas a longer trajectory may cross slower barriers. Ideally, both are combined. Trajectory-based tools such as MDpocket apply geometric analysis across frames and report pocket frequency, volume, or density [43]. TRAPP traces the spatial and physicochemical properties of a user-defined pocket and can apply logistic-regression or convolutional models to estimate druggability [44,45]. Dynamic pharmacophore and water-network analyses add chemical interpretation. These approaches should be aligned to a stable reference and accompanied by sensitivity analyses for grid spacing, pocket residues, clustering cutoffs, and frame stride.

Many cryptic openings occur beyond the routine timescale of unbiased MD. Enhanced-sampling methods alter the exploration strategy but make different assumptions. Metadynamics deposits history-dependent bias along chosen collective variables [46]. Its success, therefore, depends on variables that distinguish between closed and open states without hiding slow orthogonal coordinates. Accelerated MD lowers energetic barriers more globally [47], while Gaussian accelerated MD uses a smooth boost potential that can support energetic reweighting under suitable conditions [48]. Replica-exchange methods couple replicas across temperature or Hamiltonian space to promote barrier crossing [49]. Adaptive sampling launches new simulations from underexplored or kinetically informative states, and FAST-like strategies can prioritize a structural objective while preserving exploration [50]. Weighted-ensemble methods redistribute trajectory weights among progress-coordinate bins to estimate rare-event kinetics [51]. Steered MD applies external work to explore defined transitions, but a forced path should not be interpreted as an equilibrium mechanism without additional analysis [52].

No enhanced-sampling method is universally best. A known opening coordinate favors metadynamics; an unknown and heterogeneous cryptic-site landscape may favor adaptive sampling or cosolvent-assisted exploration; rare-event rates may motivate weighted ensembles; and large domain transitions may require coarse-grained or multiscale approaches. Validation should test whether the inferred states are reproducible under unbiased dynamics and whether reweighted populations are stable.

5.2. Kinetic state modeling and ensemble-based conformational selection

Molecular dynamics simulations generate extensive conformational ensembles, but identifying functionally relevant states from heterogeneous trajectories remains a major challenge. Markov state models (MSMs) provide a quantitative framework to transform high-dimensional trajectories into kinetic networks by describing transitions among conformational states [53,54]. After feature selection, time-lagged independent component analysis (TICA) or related variational approaches identify slowly decorrelating collective variables associated with long-timescale motions [55]. The resulting conformational space is discretized into microstates through clustering, and transition probability matrices are constructed at a defined lag time to estimate state-to-state kinetics. When the Markovian assumption is satisfied, stationary populations describe equilibrium state distributions, while implied

timescales characterize slow conformational transitions. PCCA or PCCA+ methods can further coarse-grain microstates into metastable macrostates representing kinetically distinct functional states [56]. Software platforms such as PyEMMA facilitate MSM estimation, validation, and interpretation [57].

Feature selection is critical for interpreting MSMs, as global coordinates may capture motions unrelated to ligand recognition, whereas site-focused descriptors such as inter-residue distances, torsions, and local pocket features better describe binding-site transitions. Variational approaches such as VAMP-2 can assist feature evaluation, but model performance should be further validated through implied timescales, Chapman–Kolmogorov tests, state connectivity, and robustness across lag times and clustering schemes. The kinetically resolved states generated by MSMs provide a rational basis for ensemble docking and conformational selection [58]. Compared with single-structure docking, ensemble docking considers multiple receptor conformations derived from experimental structures, MD simulations, enhanced sampling, or MSM macrostates. The key challenge is selecting biologically representative states rather than simply increasing ensemble size. A reliable workflow combines site-focused clustering, kinetic state populations, structural quality control, and post-docking evaluation. When experimental ligand information is unavailable, additional evidence from fragment mapping, solvent probe enrichment, and chemical interaction consistency can improve confidence in predicted binding modes.

Practical applications illustrate how kinetic-state modeling can alter structural interpretation and ligand prioritization. In the dopamine D3 receptor, conventional docking to the available crystal structure produced ligand poses that were inconsistent with site-directed mutagenesis. Large-scale MD simulations combined with MSM analysis identified an alternative binding mode involving displacement of the conserved F6.52 residue and opening of a cryptic pocket, yielding a structural interpretation consistent with the experimental mutational data [59]. This example demonstrates how kinetic ensembles can reveal binding-competent states that are poorly represented by a single experimental structure. A related strategy has been applied to PPM1D/Wip1 phosphatase, for which no experimental structure was initially available. MD simulations initiated from an AlphaFold-predicted structure were organized into an MSM that revealed a cryptic pocket at the flap–hinge interface. Docking and deep learning–based pose evaluation favored this dynamically exposed site, and affinity predictions aggregated across the MSM reproduced the relative potency of known allosteric inhibitors more accurately than predictions based on the single static AlphaFold structure [60]. This application illustrates how MSM-derived conformational ensembles can improve virtual screening and binding-site interpretation when structural information is limited.

5.3. Solvent-probe mapping, mixed-solvent molecular dynamics, and probe selection

MSMD simulates a macromolecule in water containing one or more organic probe species. The probes compete with water and sample favorable interaction regions, while the macromolecule remains flexible. Persistent enrichment can mark local chemistry that stabilizes fragment-like groups, and probe occupancy can help stabilize or reveal a pocket. MixMD refers specifically to workflows developed by the Carlson group [61–63], whereas SILCS, MDmix-like protocols, inverse MSMD, and other cosolvent approaches use related principles with different probe sets, concentrations, scoring formalisms, and restraints [64,65]. These methods should not be treated as interchangeable implementations of one score. Compared with static docking, MSMD includes protein motion and explicit solvent competition. Compared with pure-water MD, probes provide both a perturbation and

a chemical reporter. This dual role is a strength and a source of ambiguity. A probe can reveal a pre-existing low-population state, induce a state that would rarely form in pure water, or accumulate nonspecifically on an exposed surface. The analysis must determine which interpretation is supported.

Probe chemistry determines which microenvironments are emphasized. Acetonitrile (ACN) is small and polar; isopropanol (IPA) combines a hydrophobic group with donor and acceptor capability; N-methylacetamide (NMA) mimics peptide-like amide interactions; phenol (PHN) and pyrimidine (PYR) report aromatic and mixed polar contacts. The mixed-solvent MD suite expanded the built-in set to 13 probes—ACN, IPA, NMA, PHN, PYR, FMA, DMS, MEH, GUA, ACE, BNZ, TOL, and IND—with prepared compatibility for CHARMM27 and AMBER99SB-ILDN in the reported implementation [66]. User-defined probes can broaden coverage when appropriate parameters and coordinates are supplied. Representative probe classes, the environments they preferentially report, and their common risks are summarized in Table 4.

Table 4. Probe classes and the environments they preferentially report.

Probe class	Representative examples	Dominant information	Common risk
Small polar aprotic	ACN, acetone	Polar accessibility and compact acceptor environments	Diffuse accumulation on exposed polar surfaces
Alcohol	IPA, methanol, ethanol	Mixed hydrophobic and hydrogen-bonding regions	Self-association and concentration effects
Amide	NMA, formamide	Peptide-like donor and acceptor patterns	Strong water competition and broad polar binding
Aromatic phenol or heterocycle	PHN, PYR, indole	Aromatic stacking plus directional polar contacts	Low solubility or aggregation
Hydrophobic aromatic	Benzene, toluene	Nonpolar and aromatic subpockets	Aggregation and artificial pocket overexpansion
Charged or strongly basic	Guanidinium-like probes	Acidic sites and phosphate-rich environments	Strong nonspecific electrostatic attraction
Custom fragments	Target- or modality-specific	Tailored chemical hypothesis	Parameter quality and reduced transferability

5.4. Chemical hotspot mapping and cryptic site prioritization

Probe concentration controls the compromise between sampling and perturbation. Higher concentration increases encounter frequency but also promotes probe aggregation, changes solvent properties, and can overstabilize open conformations. Reported MixMD studies commonly use percent-level aqueous probe solutions, but the effective molarity differs across probe molecular weights and density definitions. Concentration should be specified unambiguously as % w/v, % v/v, mole fraction, or mol/L. System preparation should document the macromolecular construct, protonation states, missing residues, cofactors, metals, membrane environment, water model, ion concentration, probe parameters, force field, box geometry, equilibration, production length, and number of replicas. Probe–probe and probe–water interactions require validation because parameters

optimized for neat liquid or dilute solution may behave differently at simulation concentration [67]. Repulsive dummy potentials or occupancy restraints used in some protocols alter the physical model and must be reported. The mixed-solvent MD suite addresses reproducibility at the operational level by automating structure preparation utilities, water counting, concentration-based probe-number calculation, box setup, ion addition, and generation of GROMACS- and AMBER-compatible inputs [66]. The system builder is modular rather than a new sampling formalism. This distinction matters: automation reduces setup errors but does not eliminate force-field, sampling, or interpretation uncertainty. GROMACS, AMBER, OpenMM, and CPPTRAJ provide mature simulation or analysis infrastructure [68–71].

Trajectory analysis commonly bins selected probe atoms into a three-dimensional grid. High-density grid points can be clustered into contiguous hotspots. The Probe GridMap Builder uses CPPTRAJ for grid occupancy and DBSCAN for density-connected clustering and ranking [72]. DBSCAN-based clustering is widely used for identifying probe-enriched hotspots because it can remove isolated noise and detect density-connected regions without predefined cluster numbers. However, hotspot interpretation depends on parameters such as ϵ and minimum points, which influence cluster size and ranking. Therefore, clustering settings should be evaluated considering grid resolution, probe properties, and sensitivity analysis. Absolute probe densities are not directly comparable among different probes due to variations in concentration, molecular size, diffusion behavior, and reference states. Normalization, within-probe ranking, replica consensus, and comparison with local protein chemistry improve interpretation reliability. As illustrated in Figure 2, the six evidence layers should be interpreted within a graded confidence framework rather than as an all-or-none requirement. A candidate site may be proposed when preliminary structural, dynamic, or probe-based evidence indicates a potentially relevant interaction region. Confidence increases when multiple computational evidence layers converge, including reproducible pocket emergence, probe enrichment or persistent occupancy, favorable local physicochemical complementarity, and consistency across independent trajectories, trajectory segments, probe types, or analysis settings. Sites supported by several mutually consistent computational criteria can therefore be classified as high-confidence predicted sites even when direct experimental confirmation is not yet available. Experimental support represents a higher level of evidence and may include ligand-bound structures, mutagenesis, fragment binding, biophysical measurements, activity assays, or other functional observations. Accordingly, experimental validation strengthens and confirms a prediction but is not treated as a prerequisite for identifying a computationally meaningful candidate site.

Thermodynamic interpretation of probe enrichment requires additional caution. Preferential probe accumulation indicates that a local environment is sampled more favorably than an appropriate reference environment, but raw probe occupancy or density is not equivalent to ligand-binding free energy. Under a defined equilibrium reference state, normalized spatial probability distributions may be transformed into free-energy-like quantities or potentials of mean force. However, such interpretations depend on adequate sampling, probe concentration, force-field accuracy, probe–probe interactions, and the definition of the bulk reference state. In addition, probe enrichment does not explicitly account for the full conformational and desolvation costs associated with binding a larger ligand. Density-based hotspot ranking should therefore be distinguished from rigorous binding free-energy estimation.

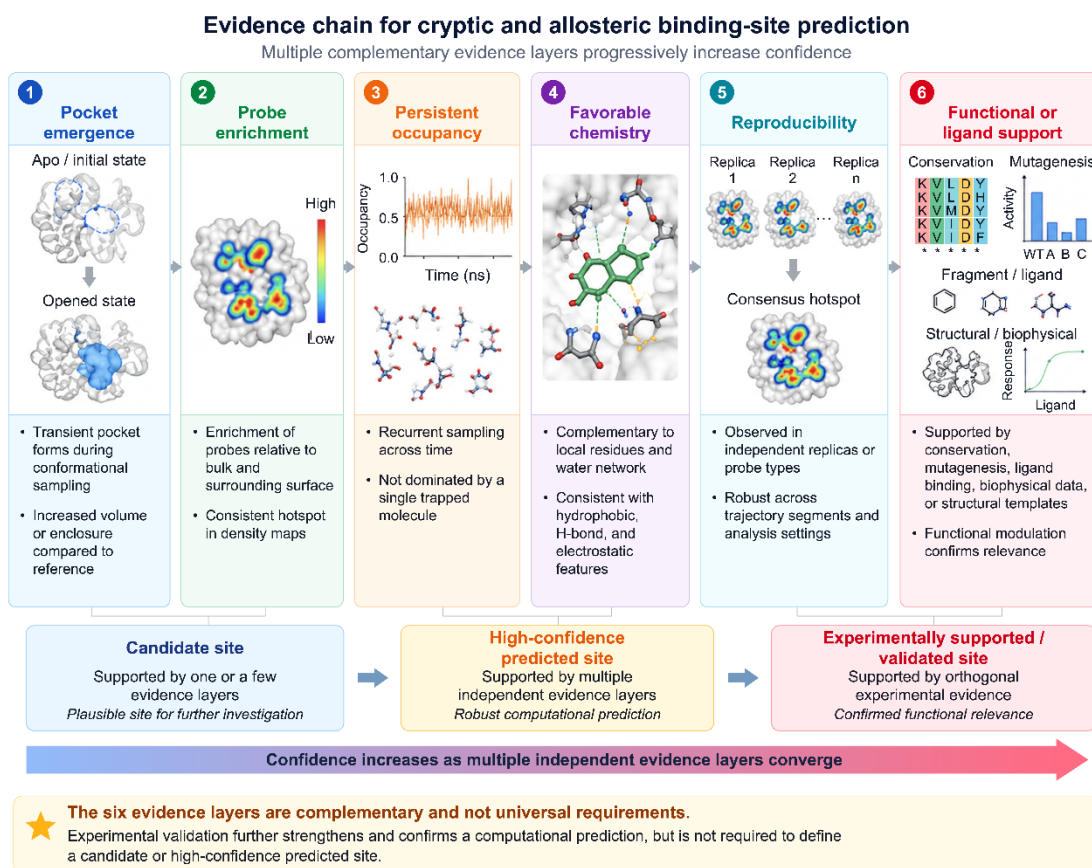


Figure 2. Hierarchical evidence framework for cryptic and allosteric binding-site prediction. The six evidence dimensions are complementary rather than universal mandatory criteria. Candidate sites may be supported by one or a few structural, dynamic, or probe-based evidence layers; high-confidence predicted sites are supported by multiple independent and mutually consistent computational evidence layers; and experimentally supported or validated sites additionally incorporate orthogonal structural, ligand-binding, mutational, biophysical, or functional evidence. Confidence increases as independent evidence layers converge.

The applicability of this framework extends beyond conventional protein–small molecule pockets. For example, mixed-solvent MD simulations of the polo-box domain of polo-like kinase 1 (PLK1-PBD), a protein–protein interaction module, used amino-acid and tripeptide probes to map interaction hotspots and identified a previously uncharacterized cryptic allosteric pocket. The predicted pocket exhibited favorable druggability characteristics and dynamic coupling to the canonical protein–protein interaction interface, illustrating how probe enrichment, pocket opening, and communication with a functional site can be combined when evaluating comparatively shallow macromolecular interfaces [73]. The PLpro study illustrates why these layers are needed [74,75]. Back site 1 attracted several probe types strongly, yet TRAPP suggested poor enclosure and limited druggability improvement. The finger region accumulated aromatic probes but did not show a convincing gain in druggability. Thumb site 2 combined multi-probe support with ACN-associated changes in volume, hydrophobicity, hydrogen-bond features, and TRAPP scores, making it a more coherent candidate. The result was a prioritization, not proof of binding. RNA binding sites introduce additional environmental variables that are less prominent in many globular protein pockets. Atomistic simulations of the guanidine-I

riboswitch showed that ligand recognition and pocket stability were strongly dependent on the ionic environment, including a localized arrangement of Mg^{2+} and K^{+} ions near the binding site. In particular, the local K^{+} ion contributed to stabilization of interactions important for ligand binding and maintenance of the folded RNA pocket [76]. This example emphasizes that evidence frameworks for RNA binding sites must explicitly consider ion occupancy, electrostatic organization, structured hydration, and force-field-dependent stabilization in addition to geometric pocket formation.

The six components of the proposed evidence framework should be interpreted as complementary evidence layers rather than universal binary requirements. Their relative importance depends on the stage of discovery, the available experimental information, and the computational protocol. Pocket emergence can be evaluated using changes in pocket volume, enclosure, solvent accessibility, or frame wise occurrence relative to an appropriate reference ensemble, such as a matched aqueous simulation. Probe enrichment can be assessed by comparing local probe density with bulk density or with neighboring surface regions, preferably after normalization for probe concentration and sampling differences. Persistent occupancy can be characterized by occupancy fractions, repeated visits, residence-related measures, or recurrence across independent trajectory segments rather than by a single long-lived probe event. Favorable chemistry refers to agreement between the physicochemical properties of enriched probes and the local residue, hydration, ion, and electrostatic environment. Reproducibility should be assessed across independent simulations, trajectory segments, probe chemistries, or reasonable variations in clustering and grid-analysis parameters. Functional or ligand support may include structural evidence, mutagenesis, fragment binding, biophysical assays, activity measurements, or known regulatory effects.

No universal numerical threshold is expected to define these criteria across all systems. Probe concentration, simulation length, force field, grid resolution, pocket definition, clustering parameters, and local molecular environment can all influence their apparent magnitude. Accordingly, evidence should be interpreted through normalized comparisons, sensitivity analysis, replica agreement, and consistency across independent evidence types. Importantly, probe density or occupancy should not be directly equated with binding free energy unless an appropriate thermodynamic reference state and sampling framework are defined.

6. Druggability assessment and artificial intelligence for functional binding site prediction

6.1. Ligandability, druggability, and functional relevance

Pocket volume, depth, enclosure, compactness, and solvent exposure describe whether a site can physically host a ligand. Hydrophobic fraction, hydrogen-bond capacity, charge balance, aromaticity, metal coordination, and hydration describe whether binding chemistry is plausible. Druggability models combine these descriptors, often using statistical or machine-learning models trained on known pockets. SiteMap provides SiteScore and Dscore-like assessments [6], LigBuilder V3 supplies cavity and knowledge-based DrugScore functions [77], and TRAPP evaluates transient pockets with global descriptors and 3D grid representations [44,45]. JEDI-like collective variables can encourage exploration of druggable pocket states [78]. Scores from different tools are not commensurate. They use different site definitions, training labels, and intended outcomes. A score should be interpreted relative to the method's calibration set and, preferably, relative to known sites in the same target or

family. Time series are often more informative than averages: a site with a moderate mean but a reproducible high-scoring metastable state may be valuable for conformational selection.

Druggability does not establish functional relevance. A distal pocket can bind a compound without altering activity, and a strongly coupled allosteric region may be difficult to bind with a small molecule. Functional assessment can integrate conservation, disease or resistance mutations, dynamic cross-correlation, residue-interaction networks, perturbation-response analysis, and known regulatory partners. Experimental tests should separate binding from modulation. For enzymes, activity assays and substrate competition are informative; for interfaces, complex formation and downstream signaling are relevant; for RNA, structural probing and functional translation or splicing assays may be needed. An allosteric-site claim is strongest when three conditions are met: direct occupation is demonstrated, a distant functional coordinate changes, and a plausible coupling pathway or state redistribution is observed. Mutations at the predicted site should alter binding or modulation without globally destabilizing the macromolecule.

6.2. *From voxels to surfaces, graphs, and equivariant models*

Early deep-learning site predictors used 3D voxel grids [34]. Modern approaches represent atoms, residues, molecular surfaces, or spatial graphs and can build invariance or equivariance to rotation and translation into the architecture. Recent surveys of protein binding-site prediction similarly show a transition from handcrafted geometric descriptors and voxel-based convolutional models toward surface, point-cloud, graph-based, and geometric deep-learning representations, although generalization across structurally distinct targets remains an important challenge [79]. MaSIF learns local surface fingerprints [35]; ScanNet and PeSTo learn interaction-site patterns from atomic geometry [36]; PocketMiner learns cryptic-opening propensity from simulation-derived labels [38]; and LABind uses protein and ligand encoders with graph and cross-attention mechanisms [39]. GeoNet and related coordinate-aware models seek interpretable local geometric representations for multiple ligand classes [80]. The choice of target label matters as much as architecture. Center prediction, residue classification, pocket segmentation, cryptic-opening propensity, ligand-specific binding, and druggability are distinct tasks. Training one model to answer all of them without task-specific labels can blur interpretation. Multitask learning is promising when outputs remain explicit, and uncertainty is retained. The methods discussed above represent implemented AI-based approaches with defined training objectives and prediction targets, although their tasks differ substantially. DeepSite predicts spatial binding-site probabilities from voxelized protein structures, whereas MaSIF, ScanNet, and PeSTo use surface- or geometry-based representations to characterize molecular interaction regions. PocketMiner specifically targets cryptic-pocket opening propensity using simulation-derived labels, while LABind incorporates ligand information into site prediction, and GeoNet emphasizes interpretable geometric representations. These established methods should therefore be distinguished according to the specific prediction object for which they were trained rather than treated as interchangeable binding-site predictors.

6.3. *Protein language models and predicted structures*

Large protein language models learn sequence representations from evolutionary-scale sequence corpora [81]. Structure-prediction systems such as AlphaFold2 and ESMFold greatly expand structural

coverage [82,83], making structure-based site prediction possible for proteins without experimental models. Predicted structures should be treated as hypotheses with spatially varying confidence. Local backbone error, side-chain uncertainty, domain orientation, missing ligands, and collapse of alternative states can directly distort pockets. Language-model embeddings can add functional and evolutionary context to geometric features, while predicted-aligned-error or local confidence can inform uncertainty. A robust model should not interpret a high-confidence fold as evidence that a closed predicted pocket is the only accessible state. Where feasible, predicted structures should seed ensembles rather than serve as immutable receptors. Protein language models and predicted structures therefore provide complementary sequence and structural representations for binding-site prediction, but they do not by themselves resolve conformational populations, pocket-opening kinetics, or solvent-mediated recognition. Their value in dynamic binding-site prediction depends on how these representations are integrated with experimentally determined structures, conformational ensembles, simulations, or other state-dependent information

6.4. Ensemble-aware AI and prospective probe-guided learning

Most established binding-site predictors analyze individual structures independently, whereas ensemble-aware approaches aim to incorporate multiple conformations, state populations, or trajectory-derived features. Relevant inputs may include pocket-volume changes, hydration patterns, residue-contact frequencies, dynamic correlations, solvent accessibility, probe-density fields, and MSM populations. In principle, such information can support prediction of functional-site probability, conformational accessibility, kinetic persistence, ligand compatibility, druggability, and uncertainty associated with sampling and model limitations. However, broadly validated ensemble-aware binding-site models remain limited, and simple averaging across conformations may obscure rare but functionally important states. Attention mechanisms, multiple-instance learning, and MSM-based weighting provide potential strategies for identifying informative conformations, but their systematic validation for binding-site prediction is still developing. A major challenge remains the limited availability of datasets containing matched structural ensembles, sufficiently sampled trajectories, kinetic information, and experimental validation.

Probe density provides a chemically annotated spatial observation that could serve as an additional input for future multimodal models. It can be represented as voxel channels, a point cloud, a surface field, or a graph connecting probe clusters to nearby residues. Equivariant graph networks can combine local topology with probe type and density [84], while atomistic message-passing models such as SchNet illustrate how chemical environments can be learned from coordinates [85]. These architectures demonstrate relevant representational capabilities, but their direct integration with probe-density fields for binding-site prediction should currently be regarded as a prospective research direction rather than an established methodology. Similarly, diffusion models could potentially treat sparse density as partial evidence and generate a complete site surface or candidate fragment occupancy [86]. An implemented example of this integration is CrypTothML, which combines mixed-solvent molecular dynamics with machine learning for cryptic-site prediction [87]. In this framework, MSMD simulations using six chemically diverse probes are first used to identify probe-enriched hotspots, after which a machine-learning model ranks these regions using hotspot-derived and protein-specific features. Evaluation on a curated cryptic-site dataset yielded an AUC-ROC of approximately 0.88 and identified cryptic sites that were missed by several comparison methods. Case analyses

further illustrated the use of the framework for prioritizing cryptic hotspots in individual protein targets. CrypTothML, therefore, provides a practical example of how solvent-derived dynamic information can be incorporated into a data-driven model, while more complex equivariant and generative integration of full probe-density fields remains a prospective direction that requires dedicated training data and prospective validation.

Future probe-guided AI models should avoid interpreting probe enrichment alone as a binding-site indicator. Instead, they should integrate bulk normalization, replica variability, cluster consistency, pocket geometry, local physicochemical properties, and negative examples from nonspecific exposed surfaces. Multi-task frameworks may simultaneously predict site boundaries, interaction preferences, conformational accessibility, and uncertainty, but the practical value of these approaches will depend on rigorous benchmarking and validation on genuinely independent targets. Ultimately, prospective validation on blinded systems will be essential for establishing the reliability and generalizability of multimodal AI-based binding-site prediction.

Site predictions can guide costly downstream experiments and should therefore be auditable. Explanations can report residue contributions, geometric features, probe types, conformations, water or ion features, and training-set neighbors. Uncertainty should distinguish at least three sources: model uncertainty, structural uncertainty, and sampling uncertainty. Deep ensembles, Bayesian approximations, or calibrated conformal methods can support confidence estimates [88], but their coverage must be tested on genuinely independent targets. The applicability domain should state whether a model was trained on proteins, RNA, metal sites, interfaces, holo structures, or predicted structures. A confident score outside that domain should not be treated as reliable evidence. The most valuable future AI systems will therefore combine prediction with a structured evidence report and identify which additional simulation or experiment would most effectively reduce uncertainty.

6.5. Benchmarking, validation, and generalizability

Benchmarking binding-site prediction methods is complicated by the fact that different approaches predict different objects and are often optimized for different applications. Some models predict pocket centers or spatial regions, whereas others classify binding-site residues, segment molecular surfaces, estimate cryptic-pocket opening propensity, or assess ligandability and druggability. Consequently, performance metrics should be matched to the specific prediction target. Distance-based measures may be appropriate for pocket-center prediction; residue-level precision, recall, F1 score, or area-under-curve metrics are more suitable for residue classification; and overlap- or surface-based measures are required for pocket segmentation. Druggability and ligandability models additionally require evaluation against experimentally defined binding or functional outcomes rather than geometric overlap alone.

Benchmark construction is equally important. Evaluation on holo structures can overestimate performance because ligand-bound conformations already contain pocket geometry shaped by binding, whereas apo structures provide a more stringent test of whether a method can identify sites before ligand-induced stabilization. For cryptic-site prediction, apo-to-holo evaluation is particularly informative because the relevant pocket may be absent or poorly formed in the starting structure. Training and test sets should also be separated according to sequence and structural similarity to reduce homologous leakage. Closely related proteins across training and test sets can inflate apparent generalizability, especially when conserved binding-site geometries or residue patterns are shared.

Class imbalance and pocket-definition ambiguity represent additional challenges. Binding-site residues or pocket voxels usually constitute only a small fraction of the molecular structure, making accuracy alone an unreliable metric. Moreover, benchmark labels depend on how binding sites are defined from ligand distance, residue contacts, cavity boundaries, or experimental annotations. Calibration should therefore accompany discrimination performance where possible, particularly when prediction scores are intended to guide expensive experiments. External validation on independent datasets, testing on unseen target families, and prospective evaluation on targets not used during model development provide stronger evidence of generalizability than random train/test splits. For ensemble-aware methods, benchmarking should additionally evaluate whether incorporating multiple conformations improves prediction relative to single-structure baselines and whether predicted confidence remains stable across alternative conformational ensembles, simulation protocols, and sampling depths.

7. Conclusions

In this review, we summarize the evolving landscape of binding-site prediction from static structural identification toward dynamic, functional, and multimodal inference. Experimental structures and static computational approaches remain essential for defining molecular recognition regions, but they provide incomplete descriptions of flexible, transient, allosteric, and cryptic sites. A biologically meaningful binding site should therefore be considered as a dynamic entity shaped by conformational populations, solvent organization, physicochemical complementarity, thermodynamic accessibility, and functional coupling. Molecular dynamics simulations, enhanced sampling, Markov state models, and mixed-solvent molecular dynamics have substantially expanded the ability to characterize hidden conformational states and chemically favorable hotspots beyond experimentally observed structures. However, computational predictions require careful assessment of sampling quality, reproducibility, force-field dependence, physicochemical consistency, and experimental support. Artificial intelligence offers increasing opportunities to integrate structural ensembles, molecular simulations, solvent-derived features, evolutionary information, and experimental measurements. Several major challenges nevertheless remain unresolved. Standardized benchmark datasets containing conformational ensembles and consistently defined prediction targets are still lacking, and experimentally validated cryptic and transient binding sites remain comparatively scarce. High-quality datasets for RNA binding regions and macromolecular interfaces are even more limited. Force-field accuracy, conformational sampling, pocket definitions, and analysis parameters continue to introduce substantial uncertainty into dynamic predictions. In parallel, ensemble-aware and multimodal AI approaches have received limited prospective validation on genuinely unseen targets, making generalizability and uncertainty calibration important open problems. Future progress will therefore depend on benchmark standardization, improved experimental validation, explicit treatment of physical and sampling uncertainty, and prospective evaluation of integrative models. Ultimately, the most reliable binding-site prediction strategies will combine physics-based simulation, biophysical evidence, and interpretable data-driven learning to generate mechanistically meaningful and experimentally actionable hypotheses.

Data availability statement

No new datasets were generated or analyzed in this review.

Use of generative-AI tools declaration

The author declares that ChatGPT (OpenAI), a Generative AI tool, was used in an assistive capacity during the preparation and revision of this article for English language editing, sentence restructuring, improving clarity and readability, editorial refinement, and reference formatting. All scientific content, interpretations, references, figures, tables, and conclusions were independently reviewed and verified by the author, who takes full responsibility for the final manuscript.

Acknowledgments

We acknowledge financial support from the Scientific Research Projects of the Jinhua Advanced Research Institute of Technology (G202509).

Conflict of interest

The author declare no conflict of interest.

Author contributions

Qiu Yue: Conceptualization, Methodology, Visualization, Data curation, Writing-Original draft preparation.

References

1. Stank A, Kokh DB, Fuller JC, et al. (2016) Protein binding pocket dynamics. *Acc Chem Res* 49: 809–815. <https://doi.org/10.1021/acs.accounts.5b00516>
2. Vajda S, Beglov D, Wakefield AE, et al. (2018) Cryptic binding sites on proteins: definition, detection, and druggability. *Curr Opin Chem Biol* 44: 1–8. <https://doi.org/10.1016/j.cbpa.2018.05.003>
3. Kuzmanic A, Bowman GR, Juarez-Jimenez J, et al. (2020) Investigating cryptic binding sites by molecular dynamics simulations. *Acc Chem Res* 53: 654–661. <https://doi.org/10.1021/acs.accounts.9b00613>
4. Hopkins AL, Groom CR (2002) The druggable genome. *Nat Rev Drug Discov* 1: 727–730. <https://doi.org/10.1038/nrd892>
5. Hajduk PJ, Huth JR, Fesik SW (2005) Druggability indices for protein targets derived from NMR-based screening data. *J Med Chem* 48: 2518–2525. <https://doi.org/10.1021/jm049131r>
6. Halgren TA (2009) Identifying and characterizing binding sites and assessing druggability. *J Chem Inf Model* 49: 377–389. <https://doi.org/10.1021/ci800324m>
7. Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* 267: 383–386. <https://doi.org/10.1126/science.7529940>

8. Sun LZ, Zhang D, Chen SJ (2017) Theory and modeling of RNA structure and interactions with metal ions and small molecules. *Annu Rev Biophys* 46: 227–246. <https://doi.org/10.1146/annurev-biophys-070816-033920>
9. Baron R, McCammon JA (2013) Molecular recognition and ligand association. *Annu Rev Phys Chem* 64: 151–175. <https://doi.org/10.1146/annurev-physchem-040412-110047>
10. Bellissent-Funel MC, Hassanali A, Havenith M, et al. (2016) Water determines the structure and dynamics of proteins. *Chem Rev* 116: 7673–7697. <https://doi.org/10.1021/acs.chemrev.5b00664>
11. Koshland DE (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA* 44: 98–104. <https://doi.org/10.1073/pnas.44.2.98>
12. Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* 12: 88–118. [https://doi.org/10.1016/S0022-2836\(65\)80285-6](https://doi.org/10.1016/S0022-2836(65)80285-6)
13. Boehr DD, Nussinov R, Wright PE (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* 5: 789–796. <https://doi.org/10.1038/nchembio.232>
14. Henzler-Wildman K, Kern D (2007) Dynamic personalities of proteins. *Nature* 450: 964–972. <https://doi.org/10.1038/nature06522>
15. Burley SK, Berman HM, Bhikadiya C, et al. (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* 47: D464–D474. <https://doi.org/10.1093/nar/gky1004>
16. Bai XC, McMullan G, Scheres SH (2015) How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci* 40: 49–57. <https://doi.org/10.1016/j.tibs.2014.10.005>
17. Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* 73: 1–16. <https://doi.org/10.1016/j.pnmrs.2013.02.001>
18. Masson GR, Burke JE, Ahn NG, et al. (2019) Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods* 16: 595–602. <https://doi.org/10.1038/s41592-019-0459-y>
19. Leney AC, Heck AJ (2017) Native mass spectrometry: what is in the name? *J Am Soc Mass Spectrom* 28: 5–13. <https://doi.org/10.1007/s13361-016-1545-3>
20. Patching SG (2014) Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. *Biochim Biophys Acta* 1838: 43–55. <https://doi.org/10.1016/j.bbamem.2013.04.028>
21. Erlanson DA, Fesik SW, Hubbard RE, et al. (2016) Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov* 15: 605–619. <https://doi.org/10.1038/nrd.2016.109>
22. Smith E, Collins I (2015) Photoaffinity labeling in target- and binding-site identification. *Future Med Chem* 7: 159–183. <https://doi.org/10.4155/fmc.14.152>
23. Laskowski RA (1995) SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph* 13: 323–330. [https://doi.org/10.1016/0263-7855\(95\)00073-9](https://doi.org/10.1016/0263-7855(95)00073-9)
24. Dundas J, Ouyang Z, Tseng J, et al. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res* 34: W116–W118. <https://doi.org/10.1093/nar/gkl282>
25. Le Guilloux V, Schmidtke P, Tuffery P (2009) Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* 10: 168. <https://doi.org/10.1186/1471-2105-10-168>

26. Volkamer A, Kuhn D, Grombacher T, et al. (2012) Combining global and local measures for structure-based druggability predictions. *J Chem Inf Model* 52: 360–372. <https://doi.org/10.1021/ci200454v>
27. Halgren T (2007) New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69: 146–148. <https://doi.org/10.1111/j.1747-0285.2007.00483.x>
28. Goodford PJ (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* 28: 849–857. <https://doi.org/10.1021/jm00145a002>
29. Kozakov D, Grove LE, Hall DR, et al. (2015) The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat Protoc* 10: 733–755. <https://doi.org/10.1038/nprot.2015.043>
30. Ashkenazy H, Abadi S, Martz E, et al. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44: W344–W350. <https://doi.org/10.1093/nar/gkw408>
31. Armon A, Graur D, Ben-Tal N (2001) ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J Mol Biol* 307: 447–463. <https://doi.org/10.1006/jmbi.2000.4474>
32. Brylinski M, Skolnick J (2008) A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc Natl Acad Sci USA* 105: 129–134. <https://doi.org/10.1073/pnas.0707684105>
33. Krivak R, Hoksza D (2018) P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. *J Cheminform* 10: 39. <https://doi.org/10.1186/s13321-018-0285-8>
34. Jimenez J, Doerr S, Martinez-Rosell G, et al. (2017) DeepSite: protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics* 33: 3036–3042. <https://doi.org/10.1093/bioinformatics/btx350>
35. Gainza P, Sverrisson F, Monti F, et al. (2020) Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nat Methods* 17: 184–192. <https://doi.org/10.1038/s41592-019-0666-6>
36. Tubiana J, Schneidman-Duhovny D, Wolfson HJ (2022) ScanNet: an interpretable geometric deep learning model for structure-based protein binding site prediction. *Nat Methods* 19: 730–739. <https://doi.org/10.1038/s41592-022-01490-7>
37. Krapp LF, Abriata LA, Cortes Rodriguez F, et al. (2023) PeSTo: parameter-free geometric deep learning for accurate prediction of protein binding interfaces. *Nat Commun* 14: 2175. <https://doi.org/10.1038/s41467-023-37701-8>
38. Meller A, Ward M, Borowsky J, et al. (2023) Predicting locations of cryptic pockets from single protein structures using the PocketMiner graph neural network. *Nat Commun* 14: 1177. <https://doi.org/10.1038/s41467-023-36699-3>
39. Zhang Z, Quan L, Wang J, et al. (2025) LABind: identifying protein binding ligand-aware sites via learning interactions between ligand and protein. *Nat Commun* 16: 7712. <https://doi.org/10.1038/s41467-025-62899-0>
40. Qiu Y, Hong WL, Wang X, et al. (2025) Discovery of inhibitors of SARS-CoV-2 papain-like protease using molecular simulation and Markov state model. *J Phys Chem A* 129: 9110–9118. <https://doi.org/10.1021/acs.jpca.5c04728>

41. Bowman GR, Geissler PL (2012) Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites. *Proc Natl Acad Sci USA* 109: 11681–11686. <https://doi.org/10.1073/pnas.1209309109>
42. Karplus M (2002) Molecular dynamics simulations of biomolecules. *Acc Chem Res* 35: 321–323. <https://doi.org/10.1021/ar020082r>
43. Schmidtke P, Bidon-Chanal A, Luque FJ, et al. (2011) MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* 27: 3276–3285. <https://doi.org/10.1093/bioinformatics/btr550>
44. Kokh DB, Richter S, Henrich S, et al. (2013) TRAPP: a tool for analysis of transient binding pockets in proteins. *J Chem Inf Model* 53: 1235–1252. <https://doi.org/10.1021/ci4000294>
45. Yuan JH, Han SB, Richter S, et al. (2020) Druggability assessment in TRAPP using machine learning approaches. *J Chem Inf Model* 60: 1685–1699. <https://doi.org/10.1021/acs.jcim.9b01185>
46. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99: 12562–12566. <https://doi.org/10.1073/pnas.202427399>
47. Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J Chem Phys* 120: 11919–11929. <https://doi.org/10.1063/1.1755656>
48. Miao Y, Feher VA, McCammon JA (2015) Gaussian accelerated molecular dynamics: unconstrained enhanced sampling and free energy calculation. *J Chem Theory Comput* 11: 3584–3595. <https://doi.org/10.1021/acs.jctc.5b00436>
49. Zhou R (2007) Replica exchange molecular dynamics method for protein folding simulation. *Methods Mol Biol* 350: 205–223. <https://doi.org/10.1385/1-59745-189-4:205>
50. Zimmerman MI, Bowman GR (2015) FAST conformational searches by balancing exploration/exploitation trade-offs. *J Chem Theory Comput* 11: 5747–5757. <https://doi.org/10.1021/acs.jctc.5b00737>
51. Huber GA, Kim S (1996) Weighted-ensemble Brownian dynamics simulations for protein association reactions. *Biophys J* 70: 97–110. [https://doi.org/10.1016/S0006-3495\(96\)79552-8](https://doi.org/10.1016/S0006-3495(96)79552-8)
52. Abou-Hatab S, Abrams CF (2025) Minimal collective variables for conformational transitions in steered and temperature-accelerated MD simulations: a T4 lysozyme case study. *J Phys Chem B* 129: 5176–5188. <https://doi.org/10.1021/acs.jpcc.5c01129>
53. Prinz JH, Wu H, Sarich M, et al. (2011) Markov models of molecular kinetics: generation and validation. *J Chem Phys* 134: 174105. <https://doi.org/10.1063/1.3565032>
54. Bowman GR, Beauchamp KA, Boxer G, et al. (2009) Progress and challenges in the automated construction of Markov state models for full protein systems. *J Chem Phys* 131: 124101. <https://doi.org/10.1063/1.3216567>
55. Perez-Hernandez G, Paul F, Giorgino T, et al. (2013) Identification of slow molecular order parameters for Markov model construction. *J Chem Phys* 139: 015102. <https://doi.org/10.1063/1.4811489>
56. Röblitz S, Weber M (2013) Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification. *Adv Data Anal Classif* 7: 147–179. <https://doi.org/10.1007/s11634-013-0134-6>
57. Scherer MK, Trendelkamp-Schroer B, Paul F, et al. (2015) PyEMMA 2: a software package for estimation, validation, and analysis of Markov models. *J Chem Theory Comput* 11: 5525–5542. <https://doi.org/10.1021/acs.jctc.5b00743>

58. Amaro RE, Baudry J, Chodera J, et al. (2018) Ensemble docking in drug discovery. *Biophys J* 114: 2271–2278. <https://doi.org/10.1016/j.bpj.2018.02.038>
59. Ferruz N, Doerr S, Vanase-Frawley MA, et al. (2018) Dopamine D3 receptor antagonist reveals a cryptic pocket in aminergic GPCRs. *Sci Rep* 8: 897. <https://doi.org/10.1038/s41598-018-19345-7>
60. Meller A, De Oliveira S, Davtyan A, et al. (2023) Discovery of a cryptic pocket in the AI-predicted structure of PPM1D phosphatase explains the binding site and potency of its allosteric inhibitors. *Front Mol Biosci* 10: 1171143. <https://doi.org/10.3389/fmolb.2023.1171143>
61. Ghanakota P, Carlson HA (2016) Moving beyond active-site detection: MixMD applied to allosteric systems. *J Phys Chem B* 120: 8685–8695. <https://doi.org/10.1021/acs.jpcb.6b03515>
62. Smith RD, Carlson HA (2021) Identification of cryptic binding sites using MixMD with standard and accelerated molecular dynamics. *J Chem Inf Model* 61: 1287–1299. <https://doi.org/10.1021/acs.jcim.0c01002>
63. Ung PM, Ghanakota P, Graham SE, et al. (2016) Identifying binding hot spots on protein surfaces by mixed-solvent molecular dynamics: HIV-1 protease as a test case. *Biopolymers* 105: 21–34. <https://doi.org/10.1002/bip.22742>
64. Ghanakota P, DasGupta D, Carlson HA (2019) Free energies and entropies of binding sites identified by MixMD cosolvent simulations. *J Chem Inf Model* 59: 2035–2045. <https://doi.org/10.1021/acs.jcim.8b00925>
65. MacKerell AD Jr, Jo S, Lakkaraju SK, et al. (2020) Identification and characterization of fragment binding sites for allosteric ligand design using the site identification by ligand competitive saturation hotspots approach (SILCS-Hotspots). *Biochim Biophys Acta Gen Subj* 1864: 129519. <https://doi.org/10.1016/j.bbagen.2020.129519>
66. Qiu Y, Liu Q, Xie JW (2026) Mixed-Solvent MD Suite: modular tools for building and analyzing mixed-solvent molecular dynamics systems. *J Comput Aided Mol Des* 40: 128. <https://doi.org/10.1007/s10822-026-00829-y>
67. Lexa KW, Goh GB, Carlson HA (2014) Parameter choice matters: validating probe parameters for use in mixed-solvent simulations. *J Chem Inf Model* 54: 2190–2199. <https://doi.org/10.1021/ci400741u>
68. Van Der Spoel D, Lindahl E, Hess B, et al. (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26: 1701–1718. <https://doi.org/10.1002/jcc.20291>
69. Case DA, Aktulga HM, Belfon K, et al. (2021) Amber 2021 reference manual. University of California, San Francisco.
70. Roe DR, Cheatham TE III (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* 9: 3084–3095. <https://doi.org/10.1021/ct400341p>
71. Eastman P, Swails J, Chodera JD, et al. (2017) OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* 13: e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>
72. Ester M, Kriegel HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press, 226–231.
73. Mukhopadhyay S, Chakrabarty S (2026) Mapping protein-protein interaction hotspots and unveiling a cryptic allosteric pocket in PLK1 PBD via mixed-solvent molecular dynamics. *ChemPhysChem* 27: e202500907. <https://doi.org/10.1002/cphc.202500907>

74. Qiu Y, Liu Q, Tu G, et al. (2022) Discovery of the cryptic sites of SARS-CoV-2 papain-like protease and analysis of its druggability. *Int J Mol Sci* 23: 11265. <https://doi.org/10.3390/ijms231911265>
75. Qiu Y, Jiang HJ, Yang YS, et al. (2025) Review of the crystallized structures of the SARS-CoV-2 papain-like protease. *J Mol Struct* 1333: 141730. <https://doi.org/10.1016/j.molstruc.2025.141730>
76. Franke L, Globisch C, Karakurt MC, et al. (2024) Atomistic simulations reveal crucial role of metal ions for ligand binding in guanidine-I riboswitch. *Macromol Rapid Commun* 45: e2400606. <https://doi.org/10.1002/marc.202400606>
77. Yuan Y, Pei J, Lai L (2020) LigBuilder V3: a multi-target de novo drug design approach. *Front Chem* 8: 142. <https://doi.org/10.3389/fchem.2020.00142>
78. Cuchillo R, Pinto-Gil K, Michel J (2015) A collective variable for the rapid exploration of protein druggability. *J Chem Theory Comput* 11: 1292–1307. <https://doi.org/10.1021/ct501072t>
79. Xia Y, Pan X, Shen HB (2024) A comprehensive survey on protein-ligand binding site prediction. *Curr Opin Struct Biol* 86: 102793. <https://doi.org/10.1016/j.sbi.2024.102793>
80. Han J, Zhang S, Guan M, et al. (2024) GeoNet enables the accurate prediction of protein-ligand binding sites through interpretable geometric deep learning. *Structure* 32: 2435–2448.e5. <https://doi.org/10.1016/j.str.2024.10.011>
81. Rives A, Meier J, Sercu T, et al. (2021) Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc Natl Acad Sci USA* 118: e2016239118. <https://doi.org/10.1073/pnas.2016239118>
82. Jumper J, Evans R, Pritzel A, et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
83. Lin Z, Akin H, Rao R, et al. (2023) Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379: 1123–1130. <https://doi.org/10.1126/science.ade2574>
84. Batzner S, Musaelian A, Sun L, et al. (2022) E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat Commun* 13: 2453. <https://doi.org/10.1038/s41467-022-29939-5>
85. Schütt KT, Sauceda HE, Kindermans PJ, et al. (2018) SchNet: a deep learning architecture for molecules and materials. *J Chem Phys* 148: 241722. <https://doi.org/10.1063/1.5019779>
86. Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. *Adv Neural Inf Process Syst* 33: 6840–6851.
87. Motono C, Yanagisawa K, Koseki J, et al. (2025) CrypTothML: an integrated mixed-solvent molecular dynamics simulation and machine learning approach for cryptic site prediction. *Int J Mol Sci* 26: 4710. <https://doi.org/10.3390/ijms26104710>
88. Gal Y, Ghahramani Z (2016) Dropout as a Bayesian approximation: representing model uncertainty in deep learning. Proceedings of the 33rd International Conference on Machine Learning. *Proc Mach Learn Res* 48: 1050–1059.



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)