

# Rational Drug Design - A Review of Modern Computational and Structure-Based Approaches

Vikas Rao, Prateek Yadav

M. Pharm, B. Pharm

Department of Pharmacology

Institute of Pharmacy, Bundelkhand University, Jhansi, India

**Abstract :** Drug discovery has undergone a transformative evolution from serendipitous observations and empirical screening to sophisticated, mechanism-driven approaches. Traditional high-throughput screening methods, while successful, are constrained by extensive resource requirements, limited chemical space exploration, and high attrition rates in clinical development. The emergence of rational drug design has revolutionized the pharmaceutical landscape by integrating computational intelligence with structural biology, enabling targeted therapeutic development with unprecedented precision. Modern structure-based drug design (SBDD) leverages advanced techniques including cryo-electron microscopy, X-ray crystallography, and nuclear magnetic resonance spectroscopy to elucidate biomolecular architectures at atomic resolution. Recent advances (2020-2026) in artificial intelligence and machine learning have exponentially enhanced molecular docking algorithms, binding affinity predictions, and de novo drug design capabilities. Ligand-based approaches, particularly pharmacophore mapping and 3D-QSAR modeling, have evolved to incorporate deep learning frameworks that identify critical molecular features governing biological activity. Virtual screening technologies now integrate multi-target profiling, ADMET prediction, and quantum mechanical calculations to prioritize lead candidates with optimal drug-likeness properties. Fragment-based drug discovery, enhanced by computational workflows, has emerged as a powerful strategy for identifying novel chemical scaffolds against challenging targets. The synergistic integration of physics-based simulations, knowledge-driven databases, and predictive algorithms is accelerating hit-to-lead optimization while reducing development timelines and costs. As we advance toward precision medicine, the convergence of structural genomics, polypharmacology networks, and cloud-based computational platforms promises to unlock previously "undruggable" targets and deliver next-generation therapeutics with improved efficacy and safety profiles.

**IndexTerms - Rational Drug Design, Structure-Based Drug Design, Pharmacophore Mapping, Virtual Screening, Computer-Aided Drug Design, Drug-Likeness.**

## 1. INTRODUCTION

### 1.1 Historical Context of Drug Discovery

Drug discovery has evolved from serendipitous observations to sophisticated computational approaches over the past century. Early pharmaceutical breakthroughs, including penicillin (1928) and aspirin, emerged through fortunate accidents rather than systematic design. The paradigm shift toward rational drug design began in the 1980s with advances in computational chemistry and structural biology, enabling researchers to visualize molecular interactions at atomic resolution[1]. High-throughput screening (HTS) technologies in the 1990s allowed screening of millions of compounds, though hit rates remained dismally low at 0-0.01%[2,3].

The past decade (2014-2024) has witnessed unprecedented transformation driven by artificial intelligence and machine learning. The Protein Data Bank now contains over 200,000 experimentally determined structures, while AlphaFold's revolutionary predictions cover over 200 million proteins[4,5]. Deep learning algorithms now integrate with traditional molecular dynamics and docking methods across the entire drug discovery pipeline[6,7]. In January 2025, the FDA issued comprehensive draft guidance on AI use in regulatory decision-making, based on over 500 AI-containing submissions received between 2016-2023[8,9]. The European Medicines Agency similarly published AI guidance in October 2024[10].

The first AI-designed drugs have entered clinical trials, with Insilico Medicine's TNiK inhibitor INS018\_055 progressing from target discovery to Phase II trials in approximately 18 months[11]. Despite these advances, the pharmaceutical industry invests over \$83 billion annually in R&D, representing 18-25% of total revenues, with only 12% of drugs entering clinical trials ultimately receiving FDA approval[12,13].

### 1.2 The Need for Rational Approaches

The economic burden of drug development has escalated dramatically, with current estimates ranging from \$879 million to \$2.87 billion per approved drug depending on therapeutic area and methodology[14,15,16]. The Tufts Center for the Study of Drug Development estimates average capitalized costs reached \$2.6 billion as of 2021, nearly triple the inflation-adjusted \$802 million from 2003[17]. More recent analyses using clinical trial contract data from 2000-2018 suggest median capitalized

R&D investment of \$985.3 million, with mean investment at \$1.34 billion[18]. Costs vary significantly by therapeutic domain, with oncology agents averaging \$2.77 billion per approved drug compared to \$766 million for nervous system agents[19].

The temporal burden compounds these financial pressures. Drug development spans 10-15 years from target identification to approval, with clinical development alone consuming approximately 95 months compared to 31 months for non-clinical phases[20]. Data from 2024 indicates this timeline has lengthened, with the period from IND filing to FDA submission averaging 89.8 months for drugs approved 2014-2018, representing an 8.1% increase from the previous period[21]. Phase III trial costs averaged \$36.58 million in 2024, a 30% increase from \$28.15 million in 2018, driven by protocol complexity and data collection requirements that increased 283.2% over the past decade[22].

Attrition rates remain alarmingly high, with approximately 90% of drug candidates failing during clinical trials or at regulatory approval[23]. Analysis of 2010-2017 clinical trial data identifies four primary failure reasons: lack of clinical efficacy (40-50%), unmanageable toxicity (30%), poor drug-like properties (10-15%), and commercial/strategic issues (10%)[24]. The Likelihood of Approval (LoA) for all developmental candidates declined from 9.6% in 2001-2010 to 7.9% for candidates entering development 2011-2020[25]. Only 5 of every 5,000 compounds entering preclinical phase advance to clinical trials, with typically one receiving approval—an overall 0.02% success rate[26]. The probability of progressing from Phase I to approval is only 10-12%, with highest attrition occurring during Phase II[27].

Approximately 80% of clinical trials fail to meet enrollment targets on time, requiring study extensions that inflate costs and delay timelines[28]. Trial delays increased from 4.5% in 2003 to 21.8% in 2024[29]. Global company-sponsored clinical trials increased 23% from 10,417 in 2019 to 12,837 in 2024, intensifying competition for investigators, sites, and patient populations[30].

The regulatory environment has evolved substantially, with FDA implementing adaptive trial designs, breakthrough therapy designations, and accelerated approval pathways[31]. Analysis suggests that optimization strategies including streamlined FDA review, adaptive designs, and simplified protocols could reduce development costs by 22-27%[32]. The COVID-19 pandemic demonstrated rapid development potential, with multiple vaccines developed within 12 months through unprecedented resource mobilization[33,34].

Rational drug design offers compelling solutions to these challenges. The AI in pharmaceuticals market, valued at \$1.8 billion in 2023, is projected to reach \$13.1 billion by 2030 with 18.8% compound annual growth rate, reflecting industry confidence in computational approaches[35].

### 1.3 Scope and Organization of This Review

This comprehensive review examines rational drug design methodologies with emphasis on computational and structure-based approaches developed 2020-2025. We synthesize key principles across: (1) structure-based drug design leveraging crystallography, cryo-EM, and NMR data; (2) ligand-based methodologies including pharmacophore mapping and 3D-QSAR; (3) virtual screening technologies; (4) fragment-based drug discovery; (5) AI/ML applications in molecular generation and property prediction; and (6) multi-target and polypharmacology approaches.

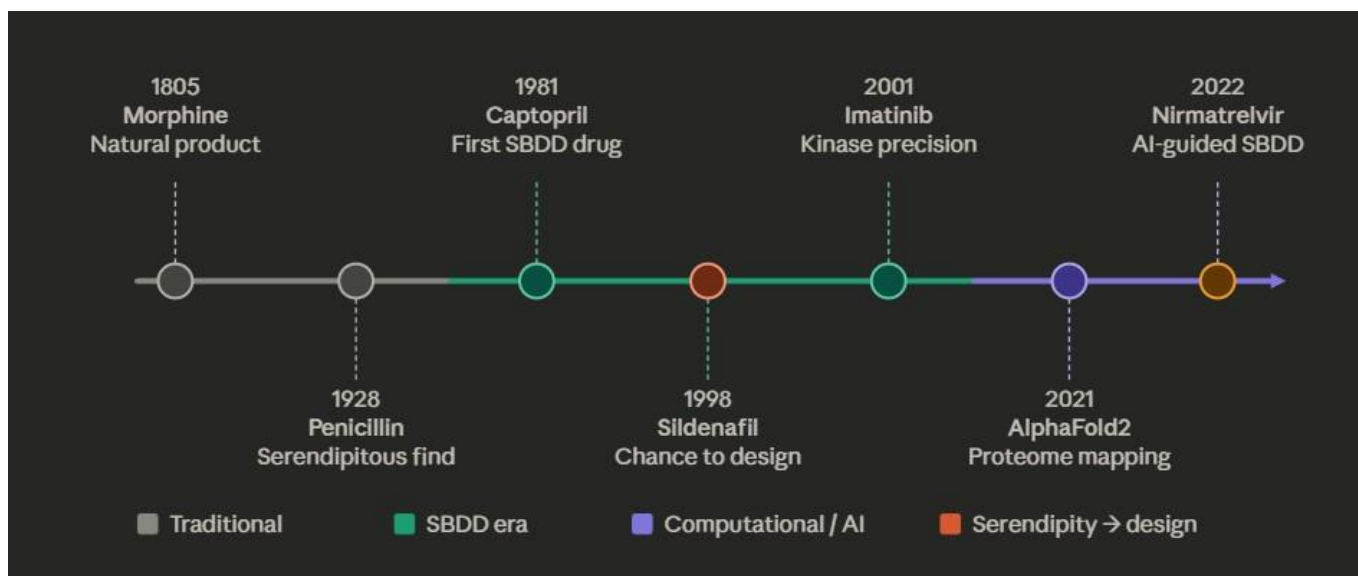
The review is organized as follows: Section 2 examines structure-based drug design principles; Section 3 addresses ligand-based approaches; Section 4 explores virtual screening methodologies; Section 5 discusses fragment-based strategies; Section 6 delves into AI/ML applications; Section 7 addresses multi-target design; Section 8 reviews ADMET prediction; and Section 9 provides future perspectives.

The target audience encompasses academic researchers, medicinal chemists, pharmaceutical scientists, graduate students, and regulatory professionals. We emphasize 2020-2025 developments, critically evaluating both successes and limitations while recognizing computational approaches as complementary tools requiring experimental validation[36].

## 2. TRADITIONAL DRUG DESIGN: FOUNDATIONS AND LIMITATIONS

### 2.1 Classical Approaches to Drug Discovery

Traditional drug discovery evolved through three principal routes: serendipitous observation, natural product screening, and empirical chemical modification. Landmark accidental discoveries — including penicillin (Fleming, 1928), aspirin's cardioprotective effects, and sildenafil's unexpected utility in erectile dysfunction — demonstrated that chance could yield transformative therapeutics, yet exposed the fundamental unreliability of unguided discovery [37, 38]. Natural products have historically contributed over 60% of FDA-approved drugs, with plants, marine organisms, and microorganisms serving as prolific sources of bioactive scaffolds; ethnopharmacological knowledge systems guided early identification of lead compounds such as morphine, quinine, and taxol [39, 40]. Structure–Activity Relationship (SAR) studies subsequently systematized lead optimization through iterative analog synthesis, bioisosterism, and QSAR modelling, enabling the "me-too" drug paradigm despite criticism for limited therapeutic novelty [41].



**Figure 1.** Timeline of drug discovery evolution — from serendipitous natural product isolation to AI-guided structure-based drug design. Milestone drugs are annotated by era (gray = traditional; teal = SBDD; purple/amber = computational/AI). Sildenafil (coral) represents the serendipity-to-rational design transition [37, 40, 90, 99].

## 2.2 High-Throughput Screening (HTS)

The advent of HTS integrated automated robotics, miniaturized 384/1536-well plate formats, and diverse detection modalities — fluorescence, luminescence, and cell-based assays — enabling simultaneous evaluation of millions of compounds against defined targets [42, 43]. While HTS offered unbiased, rapid exploration of vast chemical space and facilitated novel target identification, persistent challenges including high operational costs, false positives arising from colloidal aggregation and autofluorescence, poor hit-to-lead conversion rates (<0.1%), and chemical library quality issues substantially diminished productivity [44, 45].

## 2.3 The Transition to Rational Design

The cumulative inefficiencies of empirical methods — compounded by the escalating cost and attrition of HTS campaigns — created an urgent need for a paradigm shift. Advances in X-ray crystallography, cryo-electron microscopy, and computational modelling enabled structure-based drug design (SBDD), validated by early successes including captopril and dorzolamide, establishing mechanistic target understanding as the new foundation of drug development [41, 46].

# 3. FUNDAMENTALS OF RATIONAL DRUG DESIGN

## 3.1 Core Concepts and Principles

Rational drug design (RDD) represents a hypothesis-driven, knowledge-based paradigm wherein therapeutic candidates are conceived from a mechanistic understanding of disease biology and target architecture, contrasting sharply with empirical trial-and-error approaches [47]. Its foundational prerequisites are rigorous target identification and validation — confirming that selective modulation of a defined molecular entity yields therapeutic benefit — processes now substantially informed by CRISPR/Cas9 screens, multi-omics platforms, and AI-driven target prediction [48, 49]. Structural determination of the target — historically achieved through X-ray crystallography and NMR spectroscopy — is the cornerstone of the rational design workflow; cryo-electron microscopy (cryo-EM) has since transformed this landscape by enabling near-atomic-resolution visualization of previously intractable targets, including GPCRs, ion channels, and large membrane protein complexes, under near-physiological conditions [50, 51].

## 3.2 The Drug Design Workflow

The RDD workflow proceeds iteratively through target selection and validation, hit identification via virtual or fragment-based screening, structure-guided lead optimization, and rigorous preclinical-to-clinical progression [47, 52]. Each cycle integrates structural data with medicinal chemistry and computational modelling in a Design–Make–Test–Analyze loop that systematically narrows chemical space toward optimized candidates.

## 3.3 Advantages Over Traditional Methods

Relative to empirical approaches, RDD substantially compresses optimization timelines, reduces late-stage attrition by preemptively addressing off-target liability, and confers significant intellectual property advantages through novel, mechanistically justified scaffolds [48, 49]. Genome-scale target data, structural databases such as the PDB, and AI-enhanced structure determination are collectively accelerating the translation of structural insights into clinical candidates at unprecedented pace [51, 52].

## 4. STRUCTURE-BASED DRUG DESIGN (SBDD)

### 4.1 Theoretical Framework and Structural Biology Foundations

Structure-based drug design (SBDD) exploits high-resolution three-dimensional information of biological targets to rationally engineer molecules with optimal binding complementarity — transitioning drug discovery from empirical observation to mechanistic precision [53]. Molecular recognition in SBDD is governed by induced-fit and lock-and-key principles, wherein non-covalent interactions — hydrogen bonding, hydrophobic contacts, van der Waals forces, and electrostatics — collectively determine binding free energy and selectivity [54]. Structural data are obtained via X-ray crystallography, NMR spectroscopy, and increasingly cryo-electron microscopy (cryo-EM), which has dramatically expanded the druggable target space to include GPCRs, ion channels, and large protein complexes previously intractable to crystallisation, with the Protein Data Bank now exceeding 190,000 deposited macromolecular structures [55].

### 4.2 Molecular Docking

Molecular docking computationally predicts the preferred binding conformation and orientation of a ligand within a target's active site, evaluating interaction quality through empirical, knowledge-based, or force-field scoring functions [53, 54]. Widely used platforms include AutoDock Vina, Glide (Schrödinger), GOLD, DOCK, and MOE, each offering distinct algorithms for flexible ligand and receptor sampling. Best-practice protocols require stringent protein preparation — hydrogen assignment, protonation state correction, and grid definition — followed by re-docking of co-crystallised ligands to validate pose prediction accuracy prior to prospective screening campaigns [54].

### 4.3 Fragment-Based Drug Discovery (FBDD)

FBDD screens low-molecular-weight fragments (typically MW <300 Da) against targets using highly sensitive biophysical techniques — surface plasmon resonance (SPR), NMR spectroscopy, and X-ray crystallography — then elaborates initial weak-binding hits into potent leads through fragment growing, linking, or merging strategies [56, 57]. The approach has yielded eight FDA-approved drugs; landmark successes include vemurafenib (BRAF V600E inhibitor for melanoma) and venetoclax (BCL-2 inhibitor for CLL/AML), with 52 fragment-derived candidates advancing across clinical phases by 2024 [57, 58].

### 4.4 De Novo Drug Design

De novo approaches computationally build novel ligands atom-by-atom or fragment-by-fragment directly within the binding site, guided by geometric and energetic constraints, using tools such as LigBuilder V3, LUDI, and scaffold-hopping algorithms to generate chemically diverse scaffolds unexplored by existing libraries [53].

### 4.5 Case Studies in SBDD

SBDD's clinical impact is exemplified by several landmark molecules: HIV protease inhibitors (ritonavir, indinavir) designed against crystallographic active-site structures transformed AIDS therapy; oseltamivir (Tamiflu) was rationally designed against the neuraminidase crystal structure; imatinib and subsequent BCR-ABL kinase inhibitors defined precision oncology; and nirmatrelvir (Paxlovid), developed via rapid iterative SBDD against the SARS-CoV-2 Mpro protease structure, received FDA Emergency Use Authorization within two years of the pandemic onset — a paradigm-defining demonstration of SBDD's speed and translational power [55, 59].

## 5. PHARMACOPHORE-BASED DRUG DESIGN

### 5.1 Pharmacophore Concepts and Generation Methods

The IUPAC defines a pharmacophore as "the ensemble of steric and electronic features necessary to ensure optimal supramolecular interactions with a specific biological target structure and to trigger or block its biological response" [60]. Translated into practical chemical space, this abstract concept is operationalised through discrete three-dimensional pharmacophoric features — hydrogen bond donors and acceptors, hydrophobic regions, aromatic rings, positive/negative ionizable groups, and exclusion volumes — whose precise spatial arrangement governs molecular recognition [60, 61]. Pharmacophore models are generated via two complementary strategies. Ligand-based pharmacophore modelling derives common features by aligning and superimposing a training set of structurally diverse, biologically active compounds and is particularly valuable when no high-resolution target structure is available; structure-based pharmacophore modelling, conversely, extracts interaction fingerprints directly from the protein–ligand binding site geometry, integrating information from co-crystallised complexes or receptor active-site maps [61, 62]. Widely used software platforms include LigandScout, Phase (Schrödinger), Discovery Studio, MOE Pharmacophore module, and open-source tools such as Pharaoh and Pharmit — the latter leveraging fragment docking to construct spatial pharmacophore landmarks matched against enumerated ligand conformers [63].

### 5.2 Pharmacophore Validation and Virtual Screening

Rigorous model validation is an indispensable prerequisite before deployment in virtual screening campaigns. Standard practice employs a test set of known actives and decoys to calculate enrichment factors and receiver operating characteristic (ROC) curves, providing quantitative measures of model selectivity and sensitivity [61]. Once validated, pharmacophore models serve as rapid, computationally inexpensive three-dimensional filters in virtual screening workflows: large chemical libraries undergo conformer generation, pharmacophore mapping, and iterative post-screening refinement — commonly including docking or molecular dynamics — prior to hit selection [62].

### 5.3 Applications

Pharmacophore modelling has demonstrated broad utility across therapeutic areas. Three-dimensional pharmacophore models have been extensively deployed for GPCR ligand discovery — a target class representing approximately 40% of all prescribed medicines — enabling identification of biased agonists, allosteric ligands, and scaffold-hopping opportunities across adrenergic, opioid, serotonin, and muscarinic receptor families [64, 65]. In kinase drug discovery, integrated pharmacophore-QSAR workflows have accelerated inhibitor identification for targets including BRAF, CDK, and EGFR. Multi-target pharmacophore design — simultaneously optimising complementarity against two or more targets — is gaining prominence for complex diseases requiring polypharmacological intervention [62, 63]. Additionally, pharmacophore-guided repurposing has emerged as a cost-efficient strategy for repositioning approved drugs; computational GPCR profiling approaches identified thousands of new potential drug–receptor interactions across cannabinoid, histamine, and dopamine receptor families from existing DrugBank compounds [65].

## 6. VIRTUAL SCREENING METHODOLOGIES

### 6.1 Overview, Classification, and Workflow

Virtual screening (VS) has emerged as an indispensable computational tool in the modern drug discovery pipeline, enabling the rapid *in silico* evaluation of millions of compounds against defined biological targets to prioritize candidates for experimental testing [66]. VS strategies are broadly classified into structure-based virtual screening (SBVS) — primarily executed via molecular docking, which models ligand–receptor complementarity using three-dimensional target geometry — and ligand-based virtual screening (LBVS), which exploits physicochemical similarity, 2D fingerprints (ECFP, MACCS keys), and 3D shape metrics (Tanimoto coefficient, ROCS) to identify active analogs of known ligands [66, 67]. A typical VS workflow proceeds through library preparation, physicochemical filtering (drug-likeness), pharmacophore or similarity screening, molecular docking, and post-screening ADMET profiling before experimental hit confirmation [68].

### 6.2 Drug-Likeness and ADMET Screening

Compound triage in virtual screening pipelines is anchored by Lipinski's Rule of Five — stipulating MW  $\leq 500$  Da, LogP  $\leq 5$ , H-bond donors  $\leq 5$ , and H-bond acceptors  $\leq 10$  — complemented by Veber's rules (rotatable bonds  $\leq 10$ , PSA  $\leq 140$  Å<sup>2</sup>) and CNS-targeted permeability criteria [68, 69]. Predictive ADMET profiling has become an integral early-stage filter: tools including SwissADME, pkCSM, admetSAR, QikProp, and ADMETlab 2.0 model absorption (Caco-2 permeability, human intestinal absorption), distribution (BBB penetration, plasma protein binding), metabolism (CYP450 interactions), excretion, and toxicity endpoints including hERG cardiac liability and hepatotoxicity flags — collectively reducing costly late-stage attrition by eliminating pharmacokinetically unviable candidates before synthesis [69, 70].

### 6.3 Machine Learning and Integrated Strategies

The integration of machine learning into VS has profoundly amplified its predictive power. QSAR models employing random forests, support vector machines, and graph neural networks — collectively termed "deep QSAR" — now correlate molecular descriptors or SMILES-encoded representations with bioactivity across millions of training instances, achieving ROC-AUC values exceeding 0.92 on benchmark datasets [71, 72]. Hybrid LBVS + SBVS consensus strategies, which fuse pharmacophore, QSAR, docking, and shape-similarity scores via data fusion algorithms, consistently outperform individual methods across diverse protein targets — PPAR $\gamma$  and DPP4 consensus screens achieving AUC values of 0.90 and 0.84 respectively [73]. AI-accelerated platforms such as RosettaVS have further enabled ultra-large library docking across multi-billion compound spaces, reporting a 14% hit rate against the KLHDC2 ubiquitin ligase in prospective campaigns, validating the clinical utility of integrated computational cascades [72].

### 6.4 Performance Metrics and Validation

VS model performance is standardized through enrichment factors (EF1%, EF5%), the area under the ROC curve (AUC), Boltzmann-Enhanced Discrimination of ROC (BEDROC), and prospective experimental validation — with AUC values above 0.8 generally regarded as predictive of meaningful early enrichment [73].

## 7. INTEGRATION OF COMPUTATIONAL METHODS

### 7.1 Hybrid Approaches

Modern rational drug design rarely relies on a single computational modality; instead, it employs strategically layered hybrid workflows that exploit the complementary strengths of SBDD, pharmacophore modelling, and ligand-based methods. Structure-based pharmacophore refinement — deriving pharmacophoric constraints directly from protein–ligand interaction geometries and subsequently deploying them as three-dimensional filters in large-scale virtual screening — represents a widely validated integration strategy that combines the mechanistic fidelity of receptor-guided design with the computational efficiency of ligand-based filtering [74, 75]. Multi-target virtual screening architectures extend this further, simultaneously optimising complementarity across two or more binding sites to address polypharmacological disease contexts wherein single-target inhibition is therapeutically insufficient [74].

### 7.2 Molecular Dynamics Simulations

Molecular dynamics (MD) simulations constitute a pivotal post-docking validation layer, capturing the time-dependent conformational behaviour of protein–ligand complexes under physiologically relevant conditions and assessing thermodynamic and kinetic stability beyond the static snapshot provided by docking [76]. Induced-fit effects — conformational rearrangements of the binding pocket upon ligand accommodation — are rigorously modelled through MD

trajectories, substantially improving pose prediction accuracy for flexible targets including kinases, GPCRs, and allosteric receptors [76]. Binding free energy quantification via MM/PBSA (Molecular Mechanics/Poisson–Boltzmann Surface Area) and MM/GBSA (Generalized Born Surface Area) post-processes MD snapshots to yield  $\Delta G_{\text{bind}}$  values, providing rankings of inhibitor series that correlate more closely with experimental affinities than empirical docking scores alone; MM/GBSA is generally favoured for relative ranking tasks owing to its computational efficiency, while MM/PBSA offers superior accuracy for absolute binding free energy estimation [77, 78].

### 7.3 Artificial Intelligence and Deep Learning

The convergence of generative artificial intelligence with structural biology has ushered in a transformative new phase of computational drug design. Generative models — including variational autoencoders, recurrent neural networks, and diffusion-based architectures — are now routinely deployed for de novo molecular generation, directly producing novel chemical structures with optimised predicted affinity, drug-likeness, and synthetic accessibility within defined target binding sites [79, 80]. Retrosynthesis prediction tools, including AiZynthFinder and ASKCOS, address the historically critical gap between computationally generated structures and practical chemical synthesis by recursively decomposing candidate molecules into accessible precursors via neural network-guided reaction templates [75]. AlphaFold2 — awarded the 2024 Nobel Prize in Chemistry — and its successor AlphaFold3 (Google DeepMind, 2024) have fundamentally reshaped structural drug discovery: AF2 enabled near-atomic-accuracy structure prediction for over 200 million proteins across virtually the entire known proteome, while AF3 extended this capability to joint modelling of protein–ligand, protein–DNA, protein–RNA, and antibody–antigen complexes, reporting 50% greater accuracy over the best physics-based methods on the PoseBusters benchmark — collectively unlocking previously intractable targets for structure-based hit identification and lead optimisation [81, 82].

## 8. PRACTICAL CONSIDERATIONS AND CHALLENGES

### 8.1 Data Quality and Availability

The reliability of rational drug design pipelines is fundamentally contingent on the quality and completeness of underlying structural and bioactivity data. The Protein Data Bank (PDB) — while indispensable, exceeding 215,000 deposited structures — harbors inherent limitations: individual entries represent single conformational snapshots, resolution varies considerably across structures, and only ~24,000 structures carry associated bioactivity annotations in ChEMBL [83]. ChEMBL (v33: 24.2 million bioactivity records; 2.8 million compounds) and BindingDB serve as the principal sources of experimentally derived binding data, yet both carry significant data standardization challenges: heterogeneous assay conditions, unit inconsistencies, conflated IC<sub>50</sub>/K<sub>i</sub>/K<sub>d</sub> measurements across assay formats, and interoperability gaps between databases demand careful data curation — applying Confidence Score filters, focusing on K<sub>i</sub> values, and applying custom quality filters — before any computational modelling [83, 84, 85].

### 8.2 Computational Resources and Open-Source Alternatives

High-performance GPU computing and cloud platforms (AWS, Google Cloud, Microsoft Azure) have substantially democratized access to computationally intensive workflows including molecular dynamics simulations and large-scale virtual screening, reducing reliance on institutional supercomputing infrastructure [86]. Open-source alternatives — including AutoDock Vina, GROMACS, OpenMM, RDKit, and Pharmit — provide cost-effective routes for academic researchers, contrasting with the substantial licensing costs of commercial platforms such as Schrödinger Suite and MOE.

### 8.3 Common Pitfalls and Limitations

Despite decades of refinement, molecular docking scoring functions remain the principal source of failure in computational campaigns: whilst ligand pose sampling accuracy reaches ~80%, the scoring step correctly ranks native poses in only ~46% of benchmark cases — a discordance driven by inadequate treatment of solvent effects, entropic contributions, receptor flexibility, protonation/tautomer states, and water-mediated hydrogen bond networks [86, 87]. The largest single challenge is protein flexibility: most standard docking programs maintain receptor rigidity, failing to capture the induced-fit rearrangements that govern binding in kinases, GPCRs, and allosteric enzymes; MD simulation integration addresses this but at substantial computational cost [87]. Failure to account for these factors inflates false positive rates and undermines the correlation between predicted and experimental binding affinities.

### 8.4 Experimental Validation Requirements

Computational predictions require rigorous experimental confirmation through a hierarchy of orthogonal biophysical and biochemical assays before any hit can progress to lead status. Surface plasmon resonance (SPR) — the only technique accepted by FDA/EMA guidelines for affinity quantification in regulatory submissions — provides real-time, label-free measurement of on/off-rate kinetics ( $k_a$ ,  $k_d$ ) and equilibrium dissociation constants (K<sub>D</sub>); isothermal titration calorimetry (ITC), the thermodynamic gold standard, simultaneously delivers affinity, stoichiometry, enthalpy ( $\Delta H$ ), and entropy ( $\Delta S$ ) from a single label-free experiment; and microscale thermophoresis (MST) enables solution-phase affinity measurement without target immobilisation — uniquely suited to membrane proteins, large complexes, and targets with short residence times [88, 89]. Cell-based secondary assays and dose-response profiling then complete the hit-to-lead progression cascade, translating in vitro binding confirmation into functional pharmacological evidence.

## 9. SUCCESSFUL APPLICATIONS AND CASE STUDIES

### 9.1 FDA-Approved Drugs from Rational Design

The translational impact of rational drug design is most vividly illustrated through a lineage of landmark FDA-approved therapeutics, each traceable to deliberate, structure-informed molecular engineering. Saquinavir, the first FDA-approved HIV-1 protease inhibitor (1995), was developed via structure-based docking against crystallographic active-site models and dramatically extended survival in HIV-infected patients [90]. Dorzolamide — a carbonic anhydrase II inhibitor for glaucoma — was similarly designed through SBDD, with the binding mode of the sulfonamide warhead confirmed in the protein active site [90]. Oseltamivir (Tamiflu) was rationally engineered against the X-ray crystal structure of influenza neuraminidase; its development established that structure-guided lead optimization could yield antivirals targeting conformationally conserved active sites across strain variants [91]. Imatinib (Gleevec, FDA 2001) — the first rationally designed protein kinase inhibitor — was derived by screening chemical libraries against the BCR-ABL oncogene product and systematically optimizing a 2-phenylaminopyrimidine scaffold via SAR to achieve exquisite selectivity; its introduction elevated the five-year survival rate in chronic myeloid leukaemia from ~31% (1993) to ~90% by 2023, redefining the precision oncology paradigm [92, 93]. Aliskiren, the first direct renin inhibitor, was designed through iterative SBDD against renin crystal structures to optimize transition-state mimicry [90].

### 9.2 Recent Success Stories

The COVID-19 pandemic provided the most compressed and compelling modern demonstration of SBDD's translational power. Nirmatrelvir (Paxlovid®) was developed by Pfizer through iterative SBDD and SAR cycles — computationally fitting peptidomimetic warheads atom-by-atom into the SARS-CoV-2 main protease (M<sub>pro</sub>) active site — from structure availability to FDA Emergency Use Authorization in under two years [94, 95]. In the EPIC-HR Phase II/III trial, Paxlovid reduced the combined risk of hospitalization and death from COVID-19 by 89% compared to placebo in high-risk, unvaccinated patients [95]. Real-world meta-analyses of 32 studies subsequently confirmed pooled risk reductions of 46–64% for mortality and hospitalization during the Omicron era [96]. In oncology, the emergence of >70 FDA-approved kinase inhibitors since imatinib — across BCR-ABL, EGFR, BRAF, CDK, ALK, and VEGFR target classes — exemplifies SBDD's continued productivity in cancer therapeutics [93].

### 9.3 Comparative Analysis

Traditional drug development requires 10–15 years and an average R&D investment of ~\$879 million per approved drug inclusive of failure costs and capital outlays — an inefficiency driven primarily by high clinical attrition (overall Phase I–approval likelihood: ~7.6–14.3% across leading companies) [97, 98]. The integration of structure-based and computational design methods compresses lead optimization timelines, reduces empirical synthesis iterations, and improves mechanistic confidence entering clinical stages, as evidenced by Paxlovid's <2-year discovery-to-EUA trajectory and Insilico Medicine's ISM001-055 (TNIK inhibitor, IPF) reaching Phase II trials in 12 months at 1/10 traditional cost [99]. AI-enhanced VS platforms demonstrate 10–400× improvements in hit rates over traditional HTS (0.01–0.14%) in prospective campaigns, with AI-originated drug programs reaching 173 active clinical candidates by early 2026 [99].

Parameter	Serendipitous	High-throughput screening	SBDD / Rational design	AI-enhanced SBDD
Discovery approach	Accidental observation	Brute-force mass screening	Structure-guided design	Generative AI + 3D structure
Typical timeline	Unpredictable	10–15 yr	5–10 yr	1–4 yr
Hit rate	<0.01%	0.01–0.14%	1–10%	Up to 14%
Avg. cost per approved drug	Low / variable	~\$879M (incl. failures)	\$0.5–1.5B	Sharply reduced
Target knowledge required	None	None	High-res 3D structure	Predicted structure (AF3)
Key enabling tools	Observation, bioassay	Robotics, plate readers, FRET	X-ray, cryo-EM, docking	AlphaFold, GNN, GVAE, MD
Chemical space explored	Extremely limited	10 <sup>6</sup> –10 <sup>7</sup> compounds	Target-guided libraries	10 <sup>8</sup> –10 <sup>9</sup> (virtual)
Success predictability	Very low	Low	Moderate–high	High
Landmark example	Penicillin (1928)	Taxol (HTS-confirmed)	Imatinib (2001)	Nirmatrelvir (2022)
Key limitation	Irreproducible, unreliable	False positives, <0.1% HL conversion	Requires structural data	Data bias; interpretability

**Table 1.** Comparative analysis of drug discovery paradigms across four methodological eras. Hit rates, timelines, and cost data derived from [44, 47, 97, 98, 99]. HL = hit-to-lead; SBDD = structure-based drug design; GNN = graph neural network; GVAE = graph variational autoencoder; AF3 = AlphaFold3; MD = molecular dynamics.

## 10. FUTURE PERSPECTIVES AND EMERGING TRENDS

### 10.1 Next-Generation Technologies

The convergence of quantum computing, advanced artificial intelligence, and next-generation structural biology is poised to fundamentally reshape rational drug design over the coming decade. Quantum computing — leveraging superposition and entanglement to model molecular systems at the electronic Hamiltonian level — promises to overcome the computational intractability of high-accuracy quantum-chemical simulations underpinning binding affinity prediction, with hybrid classical-quantum variational approaches (VQE) already demonstrating proof-of-concept applications in molecular docking and protein folding within the Noisy Intermediate-Scale Quantum (NISQ) era; the AI in pharmaceuticals market, valued at \$1.8 billion in 2023, is projected to reach \$13.1 billion by 2030 (CAGR 18.8%) [100, 101]. DNA-encoded libraries (DELs) — collections of billions of small molecules individually tagged with amplifiable DNA barcodes — have emerged as a transformative hit-identification platform, breaking the traditional cost-per-well model of HTS by enabling pooled affinity selections at unprecedented chemical space coverage; integration of DEL datasets with machine learning (DEL+ML) pipelines has been prospectively validated for CK1 $\alpha$ / $\delta$  kinase targets, yielding nanomolar confirmed binders with 10% hit confirmation rates [102, 103]. Meanwhile, cryo-EM resolution advances and MicroED (micro-electron diffraction) continue to expand the structural landscape of druggable targets, complementing AlphaFold3's capacity for joint protein–ligand structure prediction established in 2024 [100].

### 10.2 Precision Medicine and Personalized Drug Design

Genomics-guided drug design — integrating CRISPR target validation, multi-omics patient stratification, and patient-specific structural modelling — is transitioning precision medicine from a clinical to a molecular design framework. Quantum computing further augments this vision: its capacity to integrate multi-omic datasets — genomic, transcriptomic, proteomic — at scales that challenge classical computing is anticipated to enable overnight AI-optimized treatment tailoring to individual metabolic signatures [101].

### 10.3 Sustainable and Green Chemistry

Computational drug design inherently promotes green chemistry principles by reducing the number of empirical synthesis trials required before an optimized lead is identified, thereby curtailing solvent consumption, reagent waste, and energy expenditure associated with iterative medicinal chemistry campaigns [100].

### 10.4 Collaborative and Open-Source Initiatives

The COVID Moonshot — a non-profit, open-science consortium of >200 scientists from 25 countries — redefined the possibilities of collaborative drug discovery by crowdsourcing 18,000 compound designs, synthesizing 2,400 compounds, and publicly releasing >500 ligand-bound X-ray structures and 10,000 assay measurements, all in the intellectual-property-free domain [104, 105]. The project advanced from a fragment screen to preclinical candidate in 18 months at a cost under \$1 million — an order-of-magnitude reduction relative to proprietary programmes — and directly informed the development of ensitrelvir (Xocova®), approved in Japan and Singapore [104]. Its successor candidate, ASAP-0017445, selected in 2025, demonstrates the model's continued pandemic preparedness applicability [105]. The Moonshot paradigm establishes a replicable blueprint for open, patent-free drug discovery: decoupling therapeutic innovation from commercial incentives and ensuring global affordability through straight-to-generic development pipelines.

## 11. CONCLUSION

The trajectory of drug discovery — from accidental observations and empirical screening to precision-guided computational design — represents one of the most consequential scientific transitions of the past century. As documented across the preceding sections, rational drug design has demonstrably outpaced its traditional predecessors: SBDD-derived candidates achieve higher mechanistic confidence entering clinical development, compressed lead-optimization timelines, and markedly improved attrition profiles relative to HTS-originated series [47, 90, 97]. The successes of imatinib, oseltamivir, venetoclax, nirmatrelvir, and the COVID Moonshot collectively validate that structure-guided, computationally integrated pipelines can deliver transformative therapeutics across oncology, infectious disease, and rare disorders — at fractions of conventional cost and time [92, 93, 94, 104]. Yet critical challenges remain: scoring function inaccuracies, protein flexibility limitations, data standardization across PDB and ChEMBL, and the persistent experimental-computational translation gap continue to constrain pipeline productivity [83, 86, 87]. Bridging these gaps demands genuine interdisciplinary convergence — medicinal chemists, structural biologists, data scientists, clinicians, and regulatory scientists working within open, collaborative frameworks modelled on the COVID Moonshot paradigm [104, 105]. Looking forward, the integration of quantum computing, AlphaFold3-enabled proteome-wide target mapping, DEL+ML hit discovery platforms, and generative AI into unified end-to-end drug design ecosystems positions rational design not merely as a tool, but as the central organizing principle of 21st-century pharmaceutical innovation [79, 100, 101]. The imperative now is not merely technological adoption but equitable access — ensuring that the efficiency gains of computational drug design translate into affordable, globally accessible therapeutics for populations that bear the highest burden of unmet disease

## SUPPLEMENTARY SECTIONS

### Abbreviations and Acronyms

#### Abbreviation Full Form

ADMET	Absorption, Distribution, Metabolism, Excretion, Toxicity
AF2 / AF3	AlphaFold2 / AlphaFold3
AI	Artificial Intelligence
AUC	Area Under the Receiver Operating Characteristic Curve
BBB	Blood–Brain Barrier
BCR-ABL	Breakpoint Cluster Region–Abelson Murine Leukemia Viral Oncogene
BRAF	B-Raf Proto-Oncogene Serine/Threonine Kinase
CADD	Computer-Aided Drug Design
ChEMBL	Chemical Database of Bioactive Molecules with Drug-like Properties
CML	Chronic Myeloid Leukaemia
CNS	Central Nervous System
CYP450	Cytochrome P450 Enzyme
DEL	DNA-Encoded Library
EF	Enrichment Factor
EGFR	Epidermal Growth Factor Receptor
FBDD	Fragment-Based Drug Discovery
FDA	Food and Drug Administration
FRET	Fluorescence Resonance Energy Transfer

**Abbreviation Full Form**

GNN	Graph Neural Network
GPCR	G Protein-Coupled Receptor
GVAE	Graph Variational Autoencoder
hERG	Human Ether-à-go-go-Related Gene
HIV	Human Immunodeficiency Virus
HTS	High-Throughput Screening
ITC	Isothermal Titration Calorimetry
KD	Equilibrium Dissociation Constant
LBVS	Ligand-Based Virtual Screening
LUDI	De Novo Ligand Design Program
MD	Molecular Dynamics
ML	Machine Learning
MM/GBSA	Molecular Mechanics/Generalized Born Surface Area
MM/PBSA	Molecular Mechanics/Poisson–Boltzmann Surface Area
MOE	Molecular Operating Environment
Mpro	Main Protease (SARS-CoV-2)
MRI	Magnetic Resonance Imaging
MST	Microscale Thermophoresis
MW	Molecular Weight
NMR	Nuclear Magnetic Resonance
NP	Natural Product
PAINS	Pan-Assay Interference Compounds
PDB	Protein Data Bank
PSA	Polar Surface Area
QSAR	Quantitative Structure–Activity Relationship
QSPR	Quantitative Structure–Property Relationship
RDD	Rational Drug Design
ROC	Receiver Operating Characteristic
SAR	Structure–Activity Relationship
SBDD	Structure-Based Drug Design
SBVS	Structure-Based Virtual Screening
SPR	Surface Plasmon Resonance
SVM	Support Vector Machine
VQE	Variational Quantum Eigensolver
VS	Virtual Screening

**Acknowledgments**

The authors sincerely acknowledge the open-access scientific community, database curators of the Protein Data Bank (RCSB PDB), ChEMBL, BindingDB, and DrugBank, whose freely accessible resources formed the foundational data infrastructure for this review. The contributions of the COVID Moonshot Consortium and the global open-science drug discovery community are gratefully recognized.

**Funding**

This review article was conducted independently without any external funding, institutional grant support, or industry sponsorship. All literature survey, data compilation, analysis, and writing were carried out entirely through the authors' own initiative and resources.

### Conflict of Interest Statement

The authors declare no conflict of interest. No financial relationships, consultancies, stock ownership, honoraria, patents, or personal affiliations exist with any commercial entity whose products or processes are discussed in this review. This work represents an independent academic review conducted solely in the interest of scientific communication.

### REFERENCES

1. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. *J Mol Biol.* 1982;161(2):269–288. [https://doi.org/10.1016/0022-2836\(82\)90153-X](https://doi.org/10.1016/0022-2836(82)90153-X)
2. Macarron R, Banks MN, Bojanic D, Burns DJ, Cirovic DA, Garyantes T, et al. Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov.* 2011;10(3):188–195. <https://doi.org/10.1038/nrd3368>
3. Bleicher KH, Böhm HJ, Müller K, Alanine AI. Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov.* 2003;2(5):369–378. <https://doi.org/10.1038/nrd1086>
4. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res.* 2000;28(1):235–242. <https://doi.org/10.1093/nar/28.1.235>
5. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res.* 2024;52(D1):D368–D375. <https://doi.org/10.1093/nar/gkad1011>
6. Sadybekov AV, Katritch V. Computational approaches streamlining drug discovery. *Nature.* 2023;616(7958):673–685. <https://doi.org/10.1038/s41586-023-05905-z>
7. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov.* 2019;18(6):463–477. <https://doi.org/10.1038/s41573-019-0024-5>
8. Food and Drug Administration. Considerations for the use of artificial intelligence to support regulatory decision-making for drug and biological products: draft guidance for industry. Silver Spring (MD): FDA; January 2025. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological>
9. Liu Q, Huang R, Hsieh J, Zhu H, Qu G, Liu Z, et al. Landscape analysis of the application of artificial intelligence and machine learning in regulatory submissions for drug development from 2016 to 2021. *Clin Pharmacol Ther.* 2023;113(4):771–774. <https://doi.org/10.1002/cpt.2668>
10. European Medicines Agency. Reflection paper on the use of artificial intelligence (AI) in the medicinal product lifecycle. Amsterdam: EMA; October 2024. [https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-use-artificial-intelligence-ai-medicinal-product-lifecycle\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-use-artificial-intelligence-ai-medicinal-product-lifecycle_en.pdf)
11. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat Biotechnol.* 2019;37(9):1038–1040. <https://doi.org/10.1038/s41587-019-0224-x>
12. Congressional Budget Office. Research and development in the pharmaceutical industry. Washington (DC): CBO; April 2021. <https://www.cbo.gov/publication/57126>
13. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. *Biostatistics.* 2019;20(2):273–286. <https://doi.org/10.1093/biostatistics/kxx069>
14. Wouters OJ, McKee M, Luyten J. Estimated research and development investment needed to bring a new medicine to market, 2009–2018. *JAMA.* 2020;323(9):844–853. <https://doi.org/10.1001/jama.2020.1166>
15. DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. *J Health Econ.* 2003;22(2):151–185. [https://doi.org/10.1016/S0167-6296\(02\)00126-1](https://doi.org/10.1016/S0167-6296(02)00126-1)
16. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: new estimates of R&D costs. *J Health Econ.* 2016;47:20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>
17. Tufts Center for the Study of Drug Development. Cost to develop and win marketing approval for a new drug is \$2.6 billion. Boston (MA): Tufts CSDD; 2014 Nov 18.
18. Sertkaya A, Jessup A, Sommers BD. Costs of drug development and research and development intensity in the US, 2000–2018. *JAMA Netw Open.* 2024;7(6):e2415445. <https://doi.org/10.1001/jamanetworkopen.2024.15445>
19. Prasad V, Mailankody S. Research and development spending to bring a single cancer drug to market and revenues after approval. *JAMA Intern Med.* 2017;177(11):1569–1575. <https://doi.org/10.1001/jamainternmed.2017.3601>
20. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: new estimates of R&D costs. *J Health Econ.* 2016;47:20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>
21. Medpath. Clinical trial complexity drives 30% cost increase: industry faces growing challenges in drug development [Internet]. Medpath; 2025 Feb 19. <https://trial.medpath.com/news/f9da185685ce1a58/>
22. Medpath. Clinical trial complexity drives 30% cost increase [Internet]. Medpath; 2025 Feb.
23. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? *Acta Pharm Sin B.* 2022;12(7):3049–3062. <https://doi.org/10.1016/j.apsb.2022.02.002>

24. Harrison RK. Phase II and phase III failures: 2013–2015. *Nat Rev Drug Discov*. 2016;15(12):817–818. <https://doi.org/10.1038/nrd.2016.184>
25. Thomas DW, Burns J, Audette J, Carroll A, Dow-Hygelund C, Hay M. *Clinical development success rates 2006–2015*. Washington (DC): BIO Industry Analysis; 2016.
26. Greenfield Chemical. The staggering cost of drug development: a look at the numbers [Internet]. Greenfield Chemical; 2023 Aug 10. <https://greenfieldchemical.com/2023/08/10/the-staggering-cost-of-drug-development-a-look-at-the-numbers/>
27. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. *Nat Biotechnol*. 2014;32(1):40–51. <https://doi.org/10.1038/nbt.2786>
28. Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: a review. *Contemp Clin Trials Commun*. 2018;11:156–164. <https://doi.org/10.1016/j.conctc.2018.08.001>
29. Medpath. Clinical trial complexity drives 30% cost increase [Internet]. Medpath; 2025 Feb.
30. ClinicalTrials.gov. Trends, charts, and maps [Internet]. Bethesda (MD): National Library of Medicine; 2024. <https://clinicaltrials.gov/>
31. Food and Drug Administration. Expedited programs for serious conditions — drugs and biologics. Silver Spring (MD): FDA; 2014 May (revised 2023).
32. Office of the Assistant Secretary for Planning and Evaluation. Drug development [Internet]. Washington (DC): US Department of Health and Human Services. <https://aspe.hhs.gov/reports/drug-development>
33. Krammer F. SARS-CoV-2 vaccines in development. *Nature*. 2020;586(7830):516–527. <https://doi.org/10.1038/s41586-020-2798-3>
34. Wouters OJ, Shadlen KC, Salcher-Konrad M, Pollard AJ, Larson HJ, Teerawattananon Y, et al. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. *Lancet*. 2021;397(10278):1023–1034. [https://doi.org/10.1016/S0140-6736\(21\)00306-8](https://doi.org/10.1016/S0140-6736(21)00306-8)
35. Pathak A, Theagarajan R, Rizqi MM, Nugraha AS, Boruah T, Kumar H, et al. AI-enabled drug and molecular discovery: computational methods, platforms, and translational horizons. *Discov Molecules*. 2025;2(1). <https://doi.org/10.1007/s44345-025-00037-5>
36. Bajorath J, Chávez-Hernández AL, Dunn WB, Engel T, Grygorenko OO, Heid E, et al. Artificial intelligence in drug discovery: into the great wide open. *Drug Discov Today*. 2022;27(7):1981–1986. <https://doi.org/10.1016/j.drudis.2022.03.002>
37. Goldstein I, Burnett AL, Rosen RC, Park PW, Stecher VJ. The serendipitous story of sildenafil: an unexpected oral therapy for erectile dysfunction. *Sex Med Rev*. 2019;7(1):115–128. <https://doi.org/10.1016/j.sxmr.2018.06.005>
38. Ban TA. The role of serendipity in drug discovery. *Dialogues Clin Neurosci*. 2006;8(3):335–344. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3181823/>
39. Ahmad I, Tahir MN, Hong L, Zia MA, Rafeeq H, Ahmad A, et al. Plant and marine-derived natural products: sustainable pathways for future drug discovery and therapeutic development. *Front Pharmacol*. 2025;15:1497668. <https://doi.org/10.3389/fphar.2024.1497668>
40. Pirintzos S, Panagiotopoulos A, Bariotakis M, Daskalakis V, Lionis C, Sourvinos G, et al. From traditional ethnopharmacology to modern natural drug discovery: a methodology discussion and specific examples. *Molecules*. 2022;27(13):4060. <https://doi.org/10.3390/molecules27134060>
41. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, et al. QSAR modelling: where have you been? Where are you going to? *J Med Chem*. 2014;57(12):4977–5010. <https://doi.org/10.1021/jm4004285>
42. Aldewachi H, Al-Zidan RN, Conner MT, Salman MM. High-throughput screening platforms in the discovery of novel drugs for neurodegenerative diseases. *Bioengineering*. 2021;8(2):30. <https://doi.org/10.3390/bioengineering8020030>
43. Murtaza I, Shah A, Nazeer RA, Saxena SK. High-throughput screening for drug discovery toward infectious diseases. In: Saxena SK, editor. *High-throughput screening for drug discovery*. London: IntechOpen; 2022. <https://doi.org/10.5772/intechopen.102988>
44. Böttcher J, Zak KM, Haselmayer P, Stump H, Brunschweiler A. Machine learning assisted hit prioritization for high throughput screening in drug discovery. *ACS Cent Sci*. 2024;10(2):314–325. <https://doi.org/10.1021/acscentsci.3c01517>
45. Lloyd M. High-throughput screening as a method for discovering new drugs. *Drug Target Rev* [Internet]. 2020 Dec 17. <https://www.drugtargetreview.com/article/61883/high-throughput-screening-as-a-method-for-discovering-new-drugs/>
46. Lyu C, Chen X, Yao M, Li R. Exploring the potential of marine natural products in drug development: a comprehensive review. *Eur J Med Chem*. 2024;265:116088. <https://doi.org/10.1016/j.ejmech.2024.116088>
47. Batool M, Ahmad B, Choi S. A structure-based drug discovery paradigm. *Int J Mol Sci*. 2019;20(11):2783. <https://doi.org/10.3390/ijms20112783>
48. Niu M, Zou Q, Wang C. AI approaches for the discovery and validation of drug targets. *Camb Prism Precis Med*. 2024;2:e7. <https://doi.org/10.1017/pcm.2024.4>
49. Halder AK, Dutta P, Mallick P, Basu S, Nasipuri M. Review of computational methods for drug target discovery and validation. *Curr Med Chem*. 2024;31(26):4152–4172. <https://doi.org/10.2174/0929867330666230418110442>
50. Cebi E, Lee J, Subramani VK, Bak N, Oh C, Kim KK. Cryo-electron microscopy-based drug design. *Front Mol Biosci*. 2024;11:1342179. <https://doi.org/10.3389/fmolb.2024.1342179>

51. Elmlund D, Le SN, Elmlund H. Advances in cryo-electron microscopy (cryoEM) for structure-based drug discovery. *Expert Opin Drug Discov.* 2025;20(2):163–176. <https://doi.org/10.1080/17460441.2025.2450636>
52. Aggarwal S, Bhatt DL. Advances in the translational science of drug discovery and development: from target identification to clinical trials. *J Am Coll Cardiol.* 2021;78(9):969–977. <https://doi.org/10.1016/j.jacc.2021.05.051>
53. Ferreira LG, Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. *Molecules.* 2015;20(7):13384–13421. <https://doi.org/10.3390/molecules200713384>
54. Moro S, Bacilieri M, Deflorian F. Molecular and structure-based drug design: from theory to practice. 2025. <https://pubmed.ncbi.nlm.nih.gov/40175038/>
55. Aplin CP, Geng J, Cerione RA. Evolving experimental techniques for structure-based drug design. *J Chem Inf Model.* 2022;62(18):4291–4300. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10161966/>
56. AlKharboush DF, Wells G, Porta EO. Fragment-based drug discovery: a graphical review. *Curr Res Pharmacol Drug Discov.* 2025;9:100233. <https://doi.org/10.1016/j.crphar.2025.100233>
57. Patel D, Barrientos A, Bhatt V. Fragment-based drug discovery: small fragments, big impact — success stories of approved oncology therapeutics. *Eur J Med Chem.* 2025. <https://doi.org/10.1016/j.ejmech.2025.117236>
58. Erlanson DA, Fesik SW, Hubbard RE, Jahnke W, Jhoti H. Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov.* 2016;15(9):605–619. <https://doi.org/10.1038/nrd.2016.109>
59. Martin RL, Heifetz A, Bodkin MJ, Townsend-Nicholson A. High-throughput structure-based drug design using drug docking, fragment molecular orbital calculations, and molecular dynamic techniques. *Methods Mol Biol.* 2024;2716:293–306. [https://doi.org/10.1007/978-1-0716-3449-3\\_13](https://doi.org/10.1007/978-1-0716-3449-3_13)
60. Noonan T, Denzinger K, Talagayev V, Chen Y, Puls K, Wolf CA, et al. Mind the gap — deciphering GPCR pharmacology using 3D pharmacophores and artificial intelligence. *Pharmaceuticals.* 2022;15(11):1304. <https://doi.org/10.3390/ph15111304>
61. Giordano D, Biancaniello C, Argenio MA, Facchiano A. Drug design by pharmacophore and virtual screening approach. *Pharmaceuticals.* 2022;15(5):646. <https://doi.org/10.3390/ph15050646>
62. Kalimuthu P, Mehra A, Gupta P. Ligand- and structure-based virtual screening in drug discovery. *Mol Divers.* 2024. <https://doi.org/10.1007/s11030-024-10979-6>
63. Boecker A, Schneider G, Terzyan SS. Navigating structure-based drug discovery with emerging innovations in physics- and knowledge-based approaches. *npj Drug Discov.* 2025. <https://doi.org/10.1038/s44386-025-00031-4>
64. Szwabowski GL, Cole JA, Baker DL, Parrill AL. Ligand-based G protein coupled receptor pharmacophore modeling: assessing the role of ligand function in model development. *Front Pharmacol.* 2024;14:1346270. <https://doi.org/10.3389/fphar.2024.1346270>
65. El-Atawneh S, Goldblum A. A machine learning algorithm suggests repurposing opportunities for targeting selected GPCRs. *Int J Mol Sci.* 2024;25(18):10230. <https://doi.org/10.3390/ijms251810230>
66. da Rocha MN, de Sousa DS, Mendes FRS, dos Santos HS, Marinho GS, Marinho MM, et al. Ligand and structure-based virtual screening approaches in drug discovery: minireview. *Mol Divers.* 2025;29(3):2799–2809. <https://doi.org/10.1007/s11030-024-10979-6>
67. Bhunia SS, Saxena M, Saxena AK. Ligand- and structure-based virtual screening in drug discovery. In: Saxena AK, editor. *Biophysical and computational tools in drug discovery. Topics in medicinal chemistry.* Cham: Springer; 2021. p. 281–339. [https://doi.org/10.1007/7355\\_2021\\_130](https://doi.org/10.1007/7355_2021_130)
68. Alanzi AR, Moussa AY, Alsalhi MS, Nawaz T, Ali I. Integration of pharmacophore-based virtual screening, molecular docking, ADMET analysis, and MD simulation for targeting EGFR: a comprehensive drug discovery study using commercial databases. *PLoS One.* 2024;19(12):e0311527. <https://doi.org/10.1371/journal.pone.0311527>
69. Gheidari D, Mehrdad M, Hoseini F. Virtual screening, molecular docking, MD simulation studies, DFT calculations, ADMET, and drug likeness of diaza-adamantane as potential MAPK-ERK inhibitors. *Front Pharmacol.* 2024;15:1360226. <https://doi.org/10.3389/fphar.2024.1360226>
70. Murugesan S, Mohan A, Ramesh M. A review of the current trends in computational approaches in drug design and metabolism. *Discov Public Health.* 2024;21:94. <https://doi.org/10.1186/s12982-024-00229-3>
71. Tropsha A, Isayev O, Varnek A, Schneider G, Cherkasov A. Integrating QSAR modelling and deep learning in drug discovery: the emergence of deep QSAR. *Nat Rev Drug Discov.* 2024;23(2):141–155. <https://doi.org/10.1038/s41573-023-00832-0>
72. Zhou G, Rusnac DV, Park H, Canzani D, Nguyen HM, Stewart L, et al. An artificial intelligence accelerated virtual screening platform for drug discovery. *Nat Commun.* 2024;15(1):7761. <https://doi.org/10.1038/s41467-024-52061-7>
73. Al-Khafaji K, Taskin-Tok T. Consensus holistic virtual screening for drug discovery: a novel machine learning model approach. *J Cheminform.* 2024;16:61. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11134635/>
74. Giordano D, Biancaniello C, Argenio MA, Facchiano A. Drug design by pharmacophore and virtual screening approach. *Pharmaceuticals.* 2022;15(5):646. <https://doi.org/10.3390/ph15050646>
75. Boecker A, Schneider G, Terzyan SS. Navigating structure-based drug discovery with emerging innovations in physics- and knowledge-based approaches. *npj Drug Discov.* 2025. <https://doi.org/10.1038/s44386-025-00031-4>

76. Murugesan S, Mohan A, Ramesh M. A review of the current trends in computational approaches in drug design and metabolism. *Discov Public Health*. 2024;21:94. <https://doi.org/10.1186/s12982-024-00229-3>
77. Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem Rev*. 2019;119(16):9478–9508. <https://doi.org/10.1021/acs.chemrev.9b00055>
78. Yau MQ, Liew CWY, Toh JH, Bhatt DL. A head-to-head comparison of MM/PBSA and MM/GBSA in predicting binding affinities for the CB1 cannabinoid ligands. *J Mol Model*. 2024;30:390. <https://doi.org/10.1007/s00894-024-06189-4>
79. Cieplinski T, Danel T, Podlewska S, Jastrzebski S. AlphaFold meets de novo drug design: leveraging structural protein information in multitarget molecular generative models. *J Chem Inf Model*. 2024;64(21):8113–8122. <https://doi.org/10.1021/acs.jcim.4c00309>
80. Liao Z, Peng X, Ma S, Zhang L. Structure-based drug design with a deep hierarchical generative model (DrugHIVE). *J Chem Inf Model*. 2024;64(21):7957–7968. <https://doi.org/10.1021/acs.jcim.4c01193>
81. Desai S, Rathod R, Jha P. Review of AlphaFold 3: transformative advances in drug design and therapeutics. *Cureus*. 2024;16(7):e64718. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11292590/>
82. Fang Z, Ran H, Zhang Y, Chen C, Lin P, Zhang X, et al. AlphaFold 3: an unprecedented opportunity for fundamental research and drug development. *Precis Clin Med*. 2025;8(3):pbaf015. <https://doi.org/10.1093/pcmedi/pbaf015>
83. Boecker A, Schneider G, Terzyan SS. Navigating structure-based drug discovery with emerging innovations in physics- and knowledge-based approaches. *npj Drug Discov*. 2025. <https://doi.org/10.1038/s44386-025-00031-4>
84. Zdrazil B, Felix E, Hunter F, Manners EJ, Blackshaw J, Corbett S, et al. The ChEMBL database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res*. 2024;52(D1):D1180–D1192. <https://doi.org/10.1093/nar/gkad1004>
85. Gilson MK, Liu T, Baitaluk M, Nicola G, Hwang L, Chong J. BindingDB in 2024: a FAIR knowledgebase of protein-small molecule binding data. *Nucleic Acids Res*. 2025;53(D1):D1633–D1641. <https://doi.org/10.1093/nar/gkae1025>
86. Dong L, Qu X, Wang B, Zhang Y. Improving protein–ligand docking and screening accuracies by incorporating a scoring function correction term. *Brief Bioinform*. 2022;23(3):bbac051. <https://doi.org/10.1093/bib/bbac051>
87. Sun M, Xu B, Wu Z, Liu L, Zhang C, Liu H, et al. Docking strategies for predicting protein-ligand interactions and their application to structure-based drug design. *Front Chem*. 2024;12:1430828. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11583305/>
88. Maschberger M, Krimm I, Lubin C, Renaud JP. High-throughput investigation of macromolecular interactions for drug development using spectral shift technology. *Biophys Rev*. 2025. <https://doi.org/10.1007/s12551-025-01359-x>
89. Holdgate GA, Freire E, Leavitt S, Pillai BK. Biophysical methods in drug discovery — SPR, ITC and MST: orthogonal strategies for hit characterisation. *Drug Discov Today*. 2019;24(7):1445–1459. <https://doi.org/10.1016/j.drudis.2019.02.006>
90. Macalino SJY, Gosu V, Hong S, Choi S. Role of computer-aided drug design in modern drug discovery. *Arch Pharm Res*. 2015;38(9):1686–1701. <https://doi.org/10.1007/s12272-015-0640-5>
91. Yonezawa A, Yoshida A, Ishiyama T, Murata M, Minami H, Matsumoto T, et al. Computational methods in drug discovery. *Beilstein J Org Chem*. 2016;12:2267–2277. <https://pmc.ncbi.nlm.nih.gov/articles/PMC5238551/>
92. Capdeville R, Buchdunger E, Zimmermann J, Matter A. Gleevec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nat Rev Drug Discov*. 2002;1(7):493–502. <https://doi.org/10.1038/nrd839>
93. Cohen P, Cross D, Jänne PA. Kinase drug discovery 20 years after imatinib: progress and future directions. *Nat Rev Drug Discov*. 2021;20(7):551–569. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8127496/>
94. Bege M, Borbás A. The design, synthesis and mechanism of action of Paxlovid, a protease inhibitor drug combination for the treatment of COVID-19. *Pharmaceutics*. 2024;16(2):217. <https://doi.org/10.3390/pharmaceutics16020217>
95. Hammond J, Fountaine RJ, Yunis C, Fleury-Aubusson E, Norrie JD, Mbaya H, et al. Nirmatrelvir for vaccinated or unvaccinated adult outpatients with COVID-19. *N Engl J Med*. 2024;390(13):1186–1195. <https://doi.org/10.1056/NEJMoa2309002>
96. Ombelet S, Barbé B, Verbruggen L, Ingelbeen B, van Griensven J, Jacobs J. Effectiveness of nirmatrelvir–ritonavir on severe outcomes of COVID-19 in the era of vaccination and Omicron: an updated meta-analysis. *J Med Virol*. 2024;96(2):e29434. <https://doi.org/10.1002/jmv.29434>
97. Sertkaya A, Beleche T, Jessup A, Sommers BD. Costs of drug development and research and development intensity in the US, 2000–2018. *JAMA Netw Open*. 2024;7(6):e2415445. <https://doi.org/10.1001/jamanetworkopen.2024.15445>
98. Schuhmacher A, Hinder M, Brief E, Gassmann O, Hartl D. Benchmarking R&D success rates of leading pharmaceutical companies: an empirical analysis of FDA approvals (2006–2022). *Drug Discov Today*. 2025;30(5):104326. <https://doi.org/10.1016/j.drudis.2025.104326>
99. Zhou G, Rusnac DV, Park H, Canzani D, Nguyen HM, Stewart L, et al. An artificial intelligence accelerated virtual screening platform for drug discovery. *Nat Commun*. 2024;15(1):7761. <https://doi.org/10.1038/s41467-024-52061-7>
100. Pathak A, Theagarajan R, Rizqi MM, Nugraha AS, Boruah T, Kumar H, et al. AI-enabled drug and molecular discovery: computational methods, platforms, and translational horizons. *Discov Molecules*. 2025;2(1). <https://doi.org/10.1007/s44345-025-00037-5>

101. Doga H, Bose A, Sahin ME, Bettencourt-Silva J, Pham A, Kim E, et al. Quantum computing and the implementation of precision medicine. *npj Genomic Med.* 2025;10. <https://doi.org/10.1038/s41525-025-00537-w>
102. Peterson AA, Liu DR. Small-molecule discovery through DNA-encoded libraries. *Nat Rev Drug Discov.* 2023;22(9):699–722. <https://doi.org/10.1038/s41573-023-00713-6>
103. Iqbal J, Gao X, Nies AT, Schwab M, Ecker GF. Evaluation of DNA encoded library and machine learning model combinations for hit discovery. *npj Drug Discov.* 2025. <https://doi.org/10.1038/s44386-025-00007-4>
104. Boby ML, Fearon D, Ferla M, Filep M, Koekemoer L, Robinson MC, et al; COVID Moonshot Consortium. Open science discovery of potent noncovalent SARS-CoV-2 main protease inhibitors. *Science.* 2023;382(6671):eabo7201. <https://doi.org/10.1126/science.abo7201>
105. Drugs for Neglected Diseases initiative. COVID Moonshot: open-science portfolio update — ASAP-0017445 pre-clinical candidate 2025 [Internet]. Geneva: DNDi; 2025. <https://dndi.org/research-development/portfolio/covid-moonshot/>

**Copyright & License:**

© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.