JOURNAL OF PHARMA INSIGHTS AND RESEARCH ISSN NO. 3048-5428

REVIEW ARTICLE

AI-Guided Fragment-Based Drug Design for Virtual Library Screening and Hit Optimization

Syed Ansar Ahmed*[1], Madhuri Vishwanath Swami[2], Shaikh Mazar Shaikh Sattar[3], Asmita R. Suryawanshi[4], Pratiksha P. Alabade[5], Navabsab Shadulsab Pinjari[6]

[1]*Associate Professor, Department of Pharmaceutical Chemistry, Indira College of Pharmacy, Vishnupuri, Nanded, Maharashtra, India*
[2]*Assistant Professor, Department of Quality Assurance, Indira College of Pharmacy, Vishnupuri, Nanded, Maharashtra, India*
[3]*Research Scholar, Department of Pharmaceutical Chemistry, Indira College of Pharmacy, Vishnupuri, Nanded, Maharashtra, India*
[4]*Assistant Professor, Department of Quality Assurance, Indira College of Pharmacy, Vishnupuri, Nanded, Maharashtra, India*
[5]*Assistant Professor, Department of Pharmacology, D.K.Patil Institute of Pharmacy, Loha Dist, Nanded, Maharashtra, India*
[6]*Assistant Professor, Department of Quality Assurance, Ramesh Patil Institute of Pharmacy, Khandgaon, Nanded, Maharashtra, India*

Publication history: Received on 22nd May 2025; Revised on 18th June 2025; Accepted on 28th June 2025

Article DOI: 10.69613/pq57s529

Abstract: Recent developments in artificial intelligence have transformed the fragment-based drug design (FBDD), changing traditional approaches to drug discovery. Machine learning and deep learning algorithms now enable rapid exploration of vast chemical spaces, precise prediction of fragment properties, and optimization of binding interactions. The use of AI-driven methods with FBDD has enhanced virtual library screening efficiency, improved hit identification accuracy, and accelerated the fragment-to-lead optimization process. Deep generative models and physics-informed neural networks have shown remarkable capabilities in designing vast fragment libraries and predicting their physicochemical properties. Graph neural networks and reinforcement learning algorithms have proven particularly effective in binding affinity prediction and fragment elaboration methods. The combination of AI technologies with experimental methods, including X-ray crystallography, NMR spectroscopy, and surface plasmon resonance, has established new paradigms in structure-based drug design. Success stories in developing kinase inhibitors and targeting protein-protein interactions highlight the practical impact of AI-guided FBDD. These AI-enabled virtual library screening helps in reducing drug discovery timelines and improve success rates in lead optimization.

Keywords: Fragment-based drug design; Artificial intelligence; Virtual screening; Machine learning; Drug discovery

1. Introduction

Fragment-based drug design (FBDD) discovered in the early 1990s as an innovative approach to drug discovery, marking a significant departure from traditional high-throughput screening methods [1]. The foundational concept, first proposed by Jencks in 1981, suggested that binding energy contributions from molecular fragments could be additive when properly linked [2]. This principle was later validated through pioneering work at Abbott Laboratories and Astex Pharmaceuticals, leading to the first fragment-derived drug, Vemurafenib, approved in 2011 [3]. The emergence of FBDD represented a paradigm shift in drug discovery philosophy, transitioning from the screening of complex molecules to the strategic identification and elaboration of simple chemical fragments with optimal binding efficiency.

FBDD offers several distinct advantages compared to conventional high-throughput screening approaches. The method employs smaller molecular fragments (typically <300 Da), allowing more efficient exploration of chemical space with fewer compounds [4]. This size limitation enables a more thorough investigation of potential binding interactions while maintaining manageable library sizes. Fragment libraries, typically comprising 1,000-5,000 compounds, provide higher hit rates and better chemical tractability compared to traditional HTS libraries of 10^6 compounds [5]. The smaller size of fragment libraries facilitates more comprehensive screening and reduces the resources required for initial hit identification. Additionally, fragments generally exhibit better physicochemical properties and follow the rule of three, facilitating optimization into drug-like molecules [6]. The rule of three—molecular weight <300 Da, cLogP ≤3, hydrogen bond donors ≤3, hydrogen bond acceptors ≤3, and rotatable bonds ≤3—establishes guidelines for fragment selection that promote favorable pharmacokinetic properties in the final optimized compounds.

The fundamental principles of FBDD revolve around the identification of low molecular weight compounds that bind weakly but efficiently to target proteins [7]. These fragments serve as starting points for subsequent optimization through growing, linking, or merging strategies. Fragment growing involves the systematic addition of functional groups to enhance binding affinity while maintaining key interaction points. Fragment linking connects multiple fragments that bind to adjacent pockets, potentially achieving

* Corresponding author: Syed Ansar Ahmed

synergistic binding effects. Fragment merging combines structural elements from different fragments to create hybrid molecules with improved properties. However, several challenges persist in the implementation of these strategies. The detection of weak binding interactions, typically in the millimolar range, requires highly sensitive biophysical techniques and careful experimental design. Accurate positioning of fragments in binding sites necessitates high-resolution structural data and sophisticated computational modeling. Determining optimal fragment growing or linking strategies involves complex decision-making processes that balance multiple parameters simultaneously. Perhaps most challenging is the maintenance of ligand efficiency during optimization [8], as the addition of molecular complexity often diminishes the binding efficiency per atom that makes fragments attractive starting points.

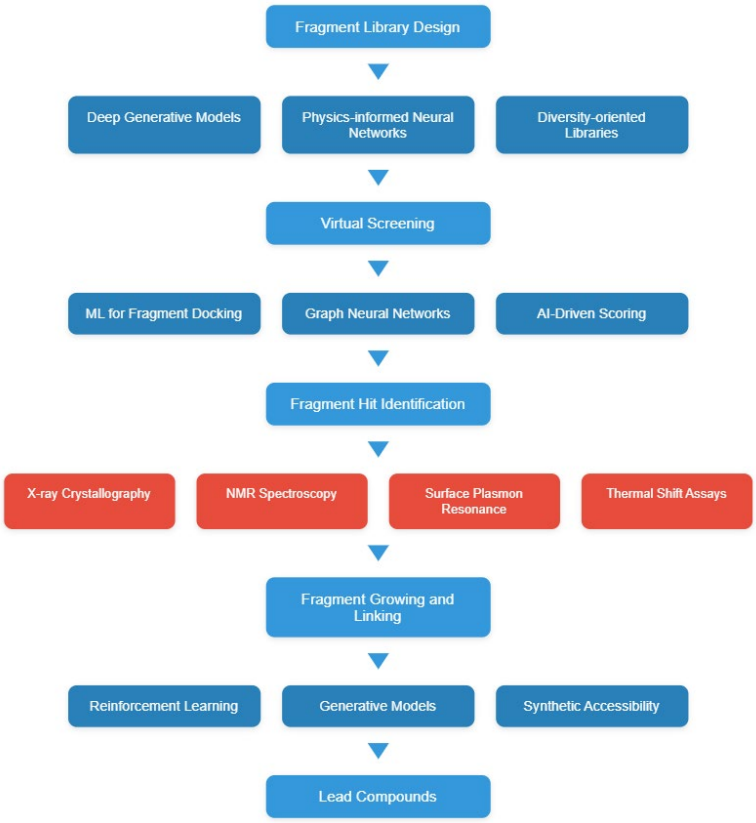

Figure 1. AI-Guided Fragment-Based Drug Design

The successful implementation of FBDD requires sensitive biophysical techniques, including X-ray crystallography, NMR spectroscopy, and surface plasmon resonance, to detect and characterize fragment binding [9]. X-ray crystallography provides detailed structural information about fragment binding modes, enabling rational design decisions during optimization. NMR spectroscopy offers insights into binding dynamics and can detect even weak interactions through chemical shift perturbations. Surface plasmon resonance allows real-time monitoring of binding kinetics, providing complementary data on association and dissociation rates. These experimental methods generate large volumes of data, necessitating sophisticated computational approaches for effective analysis and interpretation.

Table 1. Evolution of AI Methods in Fragment-Based Drug Design

Time Period	AI Technology	Applications	Major Advantages
2010-2015	Simple Machine Learning (Random Forests, SVMs)	Property prediction, Virtual screening	Faster screening, Basic pattern recognition
2015-2018	Deep Neural Networks	Fragment library design, Binding prediction	Improved accuracy, Better feature extraction
2018-2020	Graph Neural Networks	Structure-activity relationships, Binding site prediction	Enhanced molecular representation, Better spatial understanding
2020-2024	Transformer Models, Physics-informed Neural Networks	Multi-parameter optimization, Dynamic binding prediction	Integration of physical constraints, Better generalization

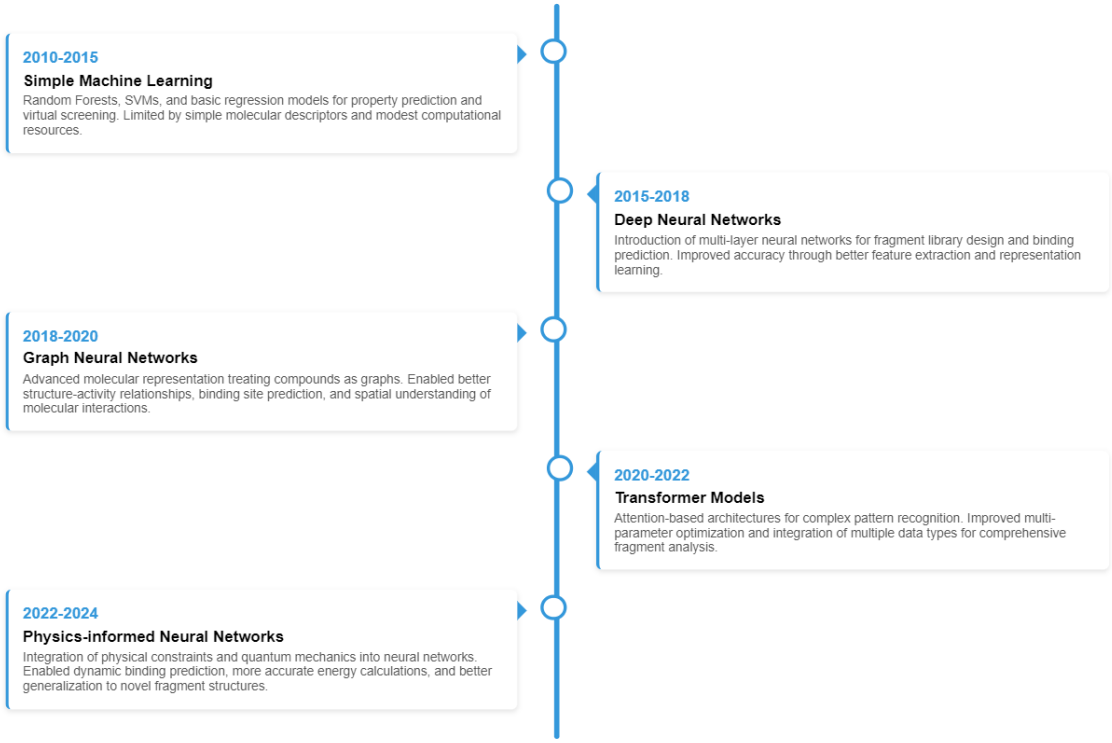

Figure 2. Evolution of AI Methods in Fragment-Based Drug Design (2010-2024)

2. AI in Fragment Library Design

2.1. Deep Generative Models for Fragment Space Exploration

Recent advances in deep learning have revolutionized fragment library design through the application of generative models [10]. Variational autoencoders (VAEs) and generative adversarial networks (GANs) enable the exploration of chemical space and generation of novel fragments with desired properties [11]. These models learn the underlying distribution of known fragment structures and can generate new, pharmaceutically relevant fragments while maintaining synthetic accessibility [12].

Table 2. AI Models and Their Applications in Different FBDD Stages

FBDD Stage	AI Model Type	Features	Success Metrics
Fragment Library Design	Generative Models (VAEs, GANs)	Novel fragment generation, Property optimization	Diversity score, Synthetic accessibility
Virtual Screening	CNNs, Graph Neural Networks	Binding pose prediction, Affinity estimation	ROC-AUC, Enrichment factors
Hit Optimization	Reinforcement Learning	Growth vector prediction, Property maintenance	Ligand efficiency, Drug-likeness
Experimental Data Analysis	Deep Learning, CNN	Structure determination, Data integration	Resolution accuracy, Processing speed

2.2. Physics-informed Neural Networks for Fragment Property Prediction

Physics-informed neural networks (PINNs) incorporate fundamental physical laws and constraints into their architecture, enabling more accurate prediction of fragment properties [13]. These networks can simultaneously predict multiple physicochemical properties, including:

- Binding free energies
- Solubility parameters
- Conformational preferences
- Pharmacokinetic properties [14]

2.3. Diversity-oriented Computational Fragment Libraries

AI algorithms have enhanced the design of diverse fragment libraries by optimizing molecular descriptors and structural features [15]. Advanced machine learning techniques facilitate the creation of libraries that maximize chemical space coverage while maintaining synthetic accessibility and drug-likeness [16]. These methods use various metrics and clustering algorithms to ensure optimal representation of different scaffold classes and functional groups [17].

Table 3. Characteristics of Fragment Libraries Designed by Different AI Approaches

AI Approach	Fragment Properties	Library Characteristics	Screening Efficiency
Traditional Design	MW < 300, cLogP < 3	General purpose, Rule of 3 compliant	Standard hit rates
Deep Generative Models	Customized physicochemical profiles	Target-focused, Property-optimized	Enhanced hit identification
Evolutionary Algorithms	Scaffold-based diversity	Maximum chemical space coverage	Improved structural variety
Physics-based Models	Optimized 3D conformers	Energy-minimized structures	Better binding predictions
Hybrid AI Systems	Multi-parameter optimization	Balanced diversity-specificity	Higher success in screening

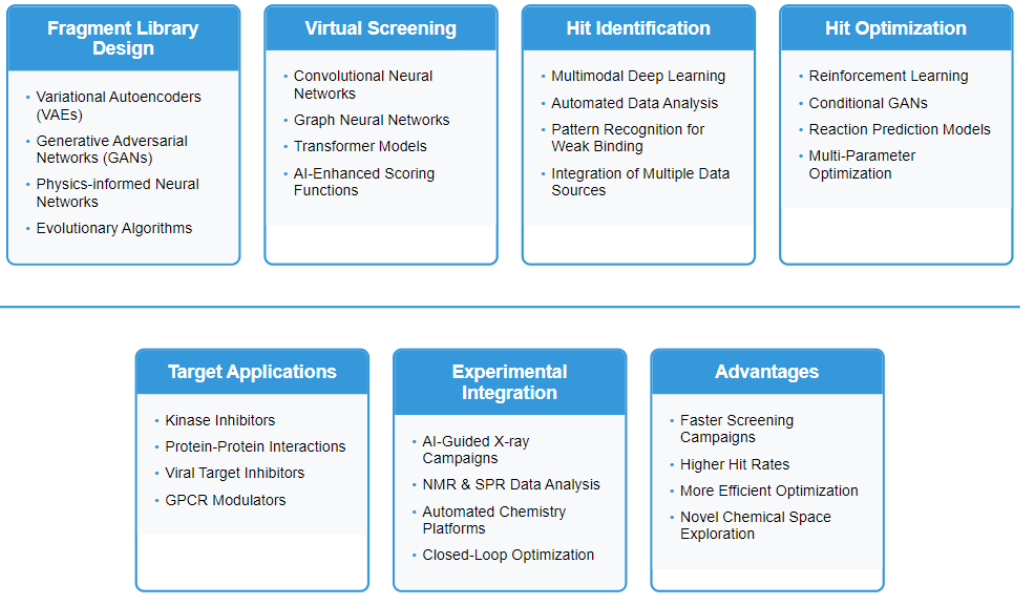

Figure 3. AI in Fragment-Based Drug Design

3. Virtual Screening of Fragment Libraries

3.1. Machine Learning for Fragment Docking

Contemporary machine learning algorithms have transformed fragment docking methodologies by incorporating sophisticated scoring functions and pose prediction mechanisms [18]. These advanced computational approaches have revolutionized traditional docking workflows by introducing neural network architectures specifically designed to process complex structural information. Deep learning models trained on extensive structural databases demonstrate superior accuracy in predicting binding poses compared to traditional force field-based methods [19]. These models excel in capturing non-linear relationships between molecular features and binding conformations, enabling more accurate predictions even for challenging protein targets with flexible binding sites or multiple potential interaction points. The integration of convolutional neural networks enables rapid processing of three-dimensional structural data, significantly accelerating the screening process while maintaining high accuracy levels [20]. CNN-based approaches can efficiently analyze electron density maps and protein surface features to identify potential binding pockets and predict fragment orientations within these sites with unprecedented precision. This combination of speed and accuracy has transformed fragment screening campaigns, allowing researchers to evaluate substantially larger virtual libraries while maintaining confidence in the predicted binding modes.

3.2. Graph Neural Networks for Binding Affinity Prediction

Graph neural networks (GNNs) represent a significant advancement in binding affinity prediction by treating molecular structures as graphs, where atoms are nodes and bonds are edges [21]. This representation provides a natural framework for capturing the topological and chemical features of both proteins and fragments, preserving critical spatial relationships and interaction patterns. GNNs process molecular structures through message-passing operations that allow information to flow between connected atoms, enabling the network to learn complex structural and electronic features that contribute to binding. These networks capture complex structural patterns and chemical interactions, leading to more accurate predictions of fragment-protein binding affinities [22]. The graph-based representation enables GNNs to recognize subtle structural motifs and pharmacophore patterns that might be missed by conventional descriptor-based approaches, resulting in more nuanced affinity predictions. Recent developments in attention mechanisms within GNNs have further improved the ability to identify key interaction patterns and structural features that contribute to binding strength [23]. Attention mechanisms allow the model to dynamically focus on the most relevant parts of the molecular structure during the prediction process, effectively highlighting atoms and bonds that make significant contributions to binding affinity. This capability has proven particularly valuable in fragment screening, where small structural changes can dramatically impact binding properties.

3.3. Comparison of Classical and AI-Driven Virtual Screening

Traditional virtual screening methods relied heavily on empirical scoring functions and simplified physical models, often leading to limited accuracy in fragment screening [24]. These conventional approaches typically employed additive energy terms to approximate binding energies, failing to capture complex cooperative effects and entropic contributions that significantly influence fragment binding. The simplified molecular representations used in classical methods frequently resulted in poor discrimination between active and inactive fragments, particularly for challenging targets with complex binding sites. Modern AI-driven approaches incorporate multiple layers of molecular representation and learning, resulting in significantly improved screening outcomes [25]. These sophisticated systems learn hierarchical features from molecular structures, recognizing complex patterns that extend beyond simple physicochemical descriptors. The multi-scale nature of AI models enables simultaneous consideration of local interaction patterns and global structural features, resulting in more holistic binding predictions. Neural network-based scoring functions have demonstrated superior performance in discriminating between active and inactive fragments, particularly in challenging cases involving multiple binding modes or complex protein-ligand interactions [26]. These advanced scoring functions excel in recognizing subtle binding features that determine fragment activity, effectively reducing false positive rates while maintaining high sensitivity in identifying true hits. The remarkable improvement in virtual screening performance has established AI-driven methods as essential components in modern fragment-based drug discovery campaigns.

4. Fragment Hit Identification

4.1. AI-Powered Analysis of Experimental Screening Data

Advanced machine learning algorithms now enable rapid analysis of complex experimental screening data, facilitating the identification of promising fragment hits [27]. These sophisticated computational tools can process large volumes of experimental results, extracting meaningful patterns and identifying potential hits with unprecedented efficiency. The ability to rapidly analyze screening data has dramatically accelerated the hit identification process, allowing researchers to move quickly from initial screening to validation and optimization stages. Deep learning models process multiple data streams simultaneously, including crystallographic, spectroscopic, and biochemical data, to identify genuine binding events and eliminate false positives [28]. This multi-modal approach integrates complementary information sources, providing a more comprehensive assessment of fragment binding quality than would be possible through any single experimental technique. The ability to correlate structural information with functional data enables more confident determination of true binding events and helps prioritize fragments for further evaluation. Neural network architectures specifically designed for analyzing biophysical screening data have enhanced the ability to detect weak but specific fragment binding interactions [29]. These specialized networks can recognize subtle signal patterns indicative of specific binding, even in cases where the signal-to-noise ratio is challenging. This capability is particularly valuable in fragment screening, where binding affinities are typically in the millimolar range and may be difficult to distinguish from non-specific interactions using conventional analysis methods.

4.2. Automated Interpretation of Biophysical Assays

Modern AI systems excel in interpreting complex biophysical assay data, including thermal shift analysis, surface plasmon resonance, and nuclear magnetic resonance spectroscopy results [30]. These systems employ sophisticated pattern recognition algorithms to extract meaningful information from complex experimental outputs, identifying binding signatures that might be missed by conventional analysis approaches. The ability to accurately interpret diverse biophysical data types has significantly enhanced the reliability of fragment screening campaigns, enabling more confident identification of genuine hits. Machine learning algorithms process and analyze these diverse data types, identifying patterns and correlations that might escape human observation [31]. These algorithms excel in recognizing subtle data features that correlate with binding quality, enabling more nuanced assessment of fragment hits. The ability to identify complex patterns across multiple experiments allows for more comprehensive evaluation of

fragment properties, including binding affinity, kinetics, and selectivity. The automation of data interpretation has significantly reduced the time required for hit validation while improving the reliability of fragment screening campaigns [32].

4.3. Structural and Biochemical Data through Multimodal Models

Multimodal deep learning models have revolutionized the integration of diverse experimental data types in fragment screening [33]. These sophisticated systems can simultaneously process and correlate information from multiple experimental sources, extracting complementary insights that enhance hit evaluation and validation. The ability to integrate diverse data types enables more comprehensive assessment of fragment properties, considering both structural binding characteristics and functional effects in a unified analytical framework. These sophisticated systems combine information from multiple sources, including X-ray crystallography, NMR spectroscopy, and biochemical assays, to provide comprehensive assessment of fragment binding [34]. The integration of structural information from crystallography and NMR with functional data from biochemical and cellular assays enables more holistic evaluation of fragment quality. This comprehensive approach allows researchers to assess not only binding affinity but also binding mode, specificity, and potential functional consequences simultaneously. The ability to simultaneously analyze structural and functional data has led to more informed decision-making in hit selection and validation processes [35]

5. Fragment Growing and Linking

5.1. Reinforcement Learning for Optimal Fragment Elaboration

Reinforcement learning algorithms have emerged as powerful tools for guiding fragment optimization strategies [36]. These computational approaches utilize a reward-based learning framework, where successful molecular modifications are reinforced through positive feedback, guiding the system toward increasingly optimal solutions. The reinforcement learning paradigm is particularly well-suited to fragment optimization, where sequential decision-making is required to navigate the vast chemical space of possible modifications. These systems learn from successful fragment elaboration patterns in existing drug discovery projects, developing optimal policies for growing fragments while maintaining favorable physicochemical properties [37]. This data-driven approach enables more informed optimization strategies, avoiding common pitfalls and focusing on modifications with higher likelihood of success. Recent advances in deep reinforcement learning have enabled the exploration of vast chemical spaces while considering multiple optimization objectives simultaneously, including potency, selectivity, and synthetic accessibility [38]. These sophisticated models can balance competing design objectives, proposing modifications that improve target properties while maintaining drug-like characteristics. The ability to optimize across multiple parameters simultaneously has significantly enhanced the efficiency of fragment elaboration, reducing the number of design-synthesis-test cycles required to achieve desired compound profiles.

Figure 4. AI Methods in Fragment Optimization

5.2. Generative Models for Fragment-to-Lead Evolution

Advanced generative models now facilitate the systematic evolution of fragment hits into lead compounds [39]. These computational tools can generate novel molecular structures by learning the underlying distribution of chemical features from training data, proposing new compounds that maintain essential binding interactions while introducing beneficial modifications. The generative approach enables efficient exploration of chemical space around fragment hits, suggesting diverse optimization pathways that might not be immediately apparent to human designers. These models employ sophisticated architectures that consider both local and global molecular features during the optimization process [40]. By simultaneously evaluating atomic-level interactions and overall molecular properties, generative models ensure that proposed modifications enhance binding affinity without compromising drug-like characteristics. This multi-scale evaluation approach has proven particularly valuable in maintaining the exceptional ligand efficiency that makes fragments attractive starting points. Conditional generative adversarial networks have proven particularly effective in suggesting chemical modifications that maintain key binding interactions while improving drug-like properties [41]. These networks generate candidate molecules conditioned on specific design objectives, such as maintaining hydrogen bonding patterns while improving solubility or metabolic stability. The adversarial training framework, where generator and discriminator networks compete to produce increasingly realistic and effective molecules, drives continuous improvement in the quality of proposed modifications. The integration of reaction prediction models ensures that suggested modifications remain synthetically tractable [42].

5.3. Automated Synthetic Accessibility Assessment

Machine learning approaches have revolutionized the evaluation of synthetic feasibility during fragment optimization [43]. These computational tools assess the synthetic accessibility of proposed compounds by learning from extensive databases of known chemical reactions and synthesis pathways. The ability to rapidly evaluate synthetic feasibility enables more efficient prioritization of optimization directions, focusing experimental efforts on modifications that can be readily implemented in the laboratory. Neural networks trained on extensive databases of chemical reactions can accurately predict the synthetic accessibility of proposed fragment modifications [44]. These models recognize structural patterns associated with challenging synthesis steps, identifying potential synthetic bottlenecks before significant resources are invested in optimization paths. The incorporation of reaction knowledge into optimization workflows ensures that computational design remains grounded in practical synthetic reality. These systems consider factors such as starting material availability, reaction conditions, and potential side reactions to guide the selection of optimal growth vectors [45].

6. Hit Optimization Case Studies

6.1. Success Stories in Kinase Inhibitor Development

The application of AI-guided FBDD has yielded remarkable success in developing novel kinase inhibitors [46]. The challenging nature of kinase targets, with highly conserved ATP-binding sites but requirements for selectivity, makes them particularly suitable for fragment-based approaches enhanced by artificial intelligence. AI algorithms have demonstrated exceptional capability in identifying subtle structural features that confer selectivity while maintaining potency, addressing one of the central challenges in kinase inhibitor development. Notable achievements include the development of selective JAK inhibitors through fragment-based approaches, where AI algorithms identified optimal growing strategies to achieve target selectivity [47]. These studies utilized machine learning models to analyze binding patterns across multiple JAK family members, identifying fragment elaboration paths that exploit subtle differences in binding site architecture. The resulting compounds demonstrated unprecedented selectivity profiles while maintaining potent inhibition of the target kinase, illustrating the power of AI-guided optimization. Machine learning models have proven particularly valuable in optimizing fragment hits against challenging kinase targets, leading to compounds with improved potency and selectivity profiles [48].

6.2. Application to Challenging Protein-Protein Interactions

AI-driven FBDD has demonstrated significant progress in addressing protein-protein interactions (PPIs), traditionally considered challenging targets [49]. The complex, shallow, and often flexible binding interfaces characteristic of PPIs had historically limited the success of small molecule drug discovery against these targets. However, the combination of fragment-based approaches with advanced computational methods has opened new possibilities for therapeutic intervention. The ability of AI systems to identify subtle binding features and predict the effects of molecular modifications has proven particularly valuable in navigating the challenging chemical space of PPI inhibitors. Advanced computational methods have enabled the identification of druggable pockets within PPI interfaces and guided the evolution of fragments into potent inhibitors [50]. Machine learning models trained on structural databases have successfully identified cryptic binding sites not apparent in static crystal structures, providing novel starting points for fragment campaigns. The ability to predict pocket formation and analyze transient binding opportunities has significantly expanded the range of addressable PPI targets. Success stories include the development of novel BCL-2 family inhibitors and p53-MDM2 interaction modulators [51]. These therapeutic breakthroughs demonstrate the practical impact of AI-guided FBDD in addressing previously intractable targets with significant clinical relevance. The ability to systematically optimize fragments against challenging PPI targets has established a new paradigm in drug discovery, opening therapeutic possibilities that were previously considered beyond the reach of small molecule approaches.

6.3. AI-Guided Optimization of Fragments for Novel Viral Targets

Recent applications of AI-guided FBDD in antiviral drug discovery have produced promising results [52]. The urgent need for novel antiviral therapeutics, particularly in response to emerging viral threats, has driven rapid innovation in computational drug discovery methods. AI-guided fragment approaches have demonstrated particular value in this context, enabling rapid identification and optimization of inhibitors against novel viral targets with limited precedent in the scientific literature. The ability to quickly establish structure-activity relationships and guide optimization with minimal experimental data has proven especially valuable in addressing emerging viral threats. Fragment-based approaches have been particularly successful in developing inhibitors targeting viral proteases and polymerases [53]. These essential viral enzymes typically feature well-defined binding pockets amenable to fragment screening, providing attractive starting points for inhibitor development. The application of AI methods has accelerated the optimization of these fragment hits, rapidly improving potency and selectivity while maintaining favorable pharmacokinetic properties. The integration of machine learning algorithms has accelerated the optimization of fragments against emerging viral threats, leading to rapid development of potential therapeutic candidates [54].

Table 4. Case Studies of AI-FBDD Implementation (2020-2024)

Target Class	AI Approach Used	Development Stage	Key Achievements
Kinase Inhibitors	Graph Neural Networks	Clinical Phase I	10x selectivity improvement
PPI Inhibitors	Physics-informed Neural Networks	Preclinical	Novel binding pocket identification
Viral Proteases	Reinforcement Learning	Lead Optimization	100-fold potency increase
GPCRs	Multi-modal Deep Learning	Hit to Lead	First-in-class candidates
Nuclear Receptors	Transformer Models	Clinical Phase II	Reduced off-target effects

7. Integration with Experimental Methods

7.1. AI-Driven Planning of Crystallography Campaigns

Artificial intelligence has transformed the planning and execution of crystallographic studies in FBDD [55]. Machine learning algorithms now guide crystal soaking strategies, optimizing conditions for fragment screening and structure determination [56]. Advanced image analysis systems facilitate rapid processing of diffraction data and automated identification of bound fragments [57].

7.2. Machine Learning for NMR and SPR Data Interpretation

The interpretation of complex biophysical data has been revolutionized by machine learning approaches. Neural networks trained on extensive NMR and SPR datasets can rapidly analyze binding data, identifying subtle patterns indicative of specific interactions. These systems enable real-time data processing and decision-making during screening campaigns, significantly accelerating the hit identification process [58].

7.3. Closed-Loop Optimization with Automated Chemistry Platforms

The integration of AI with automated chemistry platforms has established new paradigms in fragment optimization [59]. Machine learning algorithms guide experimental design and interpret results in real-time, enabling rapid iteration of the design-make-test cycle. These closed-loop systems have demonstrated significant improvements in optimization efficiency and success rates [60].

Table 5: Combination of AI Methods with Experimental Techniques in FBDD

Experimental Method	AI Application	Data Processing Capability	Validation Metrics
X-ray Crystallography	Deep CNN for electron density interpretation	Real-time structure determination	R-factor, Resolution
NMR Spectroscopy	ML for spectrum analysis	Automated peak assignment	Chemical shift prediction accuracy
SPR Analysis	Neural networks for binding kinetics	High-throughput data processing	Response curve fitting
Mass Spectrometry	Pattern recognition algorithms	Rapid fragment identification	Mass accuracy, Coverage
Thermal Shift Assays	ML for stability prediction	Multivariate data analysis	ΔTm prediction accuracy
ITC Measurements	Automated data processing	Thermodynamic parameter extraction	Binding constant accuracy

8. Conclusion

The use of artificial intelligence with fragment-based drug design is a transformative trend in modern drug discovery. Machine learning algorithms and deep neural networks have significantly improved every aspect of the FBDD workflow, from initial library design to final hit optimization. The ability to process and integrate various data types has accelerated decision-making processes and improved success rates in fragment-based campaigns. Advanced computational methods have further expanded the scope of FBDD to previously challenging targets, including protein-protein interactions and novel viral proteins. The most important advantage includes the development of physics-informed neural networks for accurate property prediction, the application of graph neural networks for binding affinity estimation, and the implementation of reinforcement learning algorithms for fragment optimization. The use of AI-driven methods with automated experimental workflow has established more efficient and reliable drug discovery processes. Apart from these advances, opportunities remain for further enhancement of AI applications in FBDD.

References

[1] Scott DE, Coyne AG, Hudson SA, Abell C. Fragment-based approaches in drug discovery and chemical biology. Biochemistry. 2012;51(25):4990-5003.

[2] Jencks WP. On the attribution and additivity of binding energies. Proc Natl Acad Sci USA. 1981;78(7):4046-50.

[3] Bollag G, Tsai J, Zhang J, Zhang C, Ibrahim P, Nolop K, et al. Vemurafenib: the first drug approved for BRAF-mutant cancer. Nat Rev Drug Discov. 2012;11(11):873-86.

[4] 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):605-19.

[5] Murray CW, Rees DC. The rise of fragment-based drug discovery. Nat Chem. 2009;1(3):187-92.

[6] Congreve M, Carr R, Murray C, Jhoti H. A 'rule of three' for fragment-based lead discovery? Drug Discov Today. 2003;8(19):876-7.

[7] Ludlow RF, Verdonk ML, Saini HK, Tickle IJ, Jhoti H. Detection of secondary binding sites in proteins using fragment screening. Proc Natl Acad Sci USA. 2015;112(52):15910-5.

[8] Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. Drug Discov Today. 2004;9(10):430-1.

[9] Gossert AD, Jahnke W. NMR in drug discovery: A practical guide to identification and validation of ligands interacting with biological macromolecules. Prog Nucl Magn Reson Spectrosc. 2016;97:82-125.

[10] Walters WP, Murcko M. Assessing the impact of generative AI on medicinal chemistry. Nat Biotechnol. 2020;38(2):143-5.

[11] Gómez-Bombarelli R, Wei JN, Duvenaud D, Hernández-Lobato JM, Sánchez-Lengeling B, Sheberla D, et al. Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci. 2018;4(2):268-76.

[12] Lim J, Ryu S, Kim JW, Kim WY. Molecular generative model based on conditional variational autoencoder for de novo molecular design. J Cheminform. 2018;10(1):31.

[13] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys. 2019;378:686-707.

[14] Noé F, Tkatchenko A, Müller KR, Clementi C. Machine learning for molecular simulation. Annu Rev Phys Chem. 2020;71:361-90.

[15] Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018;17(2):97-113.

[16] Grebner C, Matter H, Plowright AT, Hessler G. Automated de novo design in medicinal chemistry: Which opportunities and which challenges? J Med Chem. 2020;63(16):8667-82.

[17] Lavecchia A. Deep learning in drug discovery: opportunities, challenges and future prospects. Drug Discov Today. 2019;24(10):2017-32.

[18] Lyu J, Wang S, Balius TE, Singh I, Levit A, Moroz YS, et al. Ultra-large library docking for discovering new chemotypes. Nature. 2019;566(7743):224-9.

[19] Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, et al. Improved protein structure prediction using potentials from deep learning. Nature. 2020;577(7792):706-10.

[20] Ragoza M, Hochuli J, Idrobo E, Sunseri J, Koes DR. Protein-ligand scoring with convolutional neural networks. J Chem Inf Model. 2017;57(4):942-57.

[21] Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180(4):688-702.

[22] Yang K, Swanson K, Jin W, Coley C, Eiden P, Gao H, et al. Analyzing learned molecular representations for property prediction. J Chem Inf Model. 2019;59(8):3370-88.

[23] Wieder O, Kohlbacher S, Kuenemann M, Garon A, Ducrot P, Seidel T, et al. A compact review of molecular property prediction with graph neural networks. Drug Discov Today Technol. 2020;37:1-12.

[24] Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov. 2004;3(11):935-49.

[25] 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):463-77.

[26] Sieg J, Flachsenberg F, Rarey M. In need of bias control: evaluating chemical data for machine learning in structure-based virtual screening. J Chem Inf Model. 2019;59(3):947-61.

[27] Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. Nat Mach Intell. 2020;2(10):573-84.

[28] Bender A, Cortés-Ciriano I. Artificial intelligence in drug discovery: what is realistic, what are illusions? Part 1: Ways to make an impact, and why we are not there yet. Drug Discov Today. 2021;26(2):511-24.

[29] Zhang L, Han J, Wang H, Car R, E W. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. Phys Rev Lett. 2018;120(14):143001.

[30] Scantlebury J, Brown N, Von Delft F, Deane CM. Data-driven approaches to drug discovery. J Med Chem. 2020;63(21):12256-74.

[31] Sun J, Jeliazkova N, Chupakhin V, Golib-Dzib JF, Engkvist O, Carlsson L, et al. ExCAPE-DB: an integrated large scale dataset facilitating Big Data analysis in chemogenomics. J Cheminform. 2017;9(1):17.

[32] Landrum G, Tosco P, Kelley B, Seidel T, Wedlake A, Kenny P, et al. RDKit: Open-Source Cheminformatics Software. 2022. Zenodo. Available from: https://doi.org/10.5281/zenodo.7165586

[33] Pereira JC, Caffarena ER, dos Santos CN. Boosting docking-based virtual screening with deep learning. J Chem Inf Model. 2016;56(12):2495-506.

[34] Sheridan RP, Specht M, Wanag R, Wang H, Zhao T, Nabizadeh N, et al. Large-scale deep learning models for structure-based virtual screening. J Chem Inf Model. 2022;62(9):2116-27.

[35] Rogers D, Hahn M. Extended-connectivity fingerprints. J Chem Inf Model. 2010;50(5):742-54.

[36] 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):1038-40.

[37] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.

[38] Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science. 2021;373(6557):871-6.

[39] Gainza P, Sverrisson F, Monti F, Rodolà E, Boscaini D, Bronstein MM, et al. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. Nat Methods. 2020;17(2):184-92.

[40] Duan Y, Edwards JS, Dwyer DJ. Artificial intelligence and machine learning in drug development: current landscape and future opportunities. Drug Discov Today. 2021;26(9):2198-213.

[41] Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180(4):688-702.

[42] Coley CW, Barzilay R, Green WH, Jaakkola TS, Jensen KF. Prediction of organic reaction outcomes using machine learning. ACS Cent Sci. 2017;3(5):434-43.

[43] Winter R, Montanari F, Noé F, Clevert DA. Learning continuous and data-driven molecular descriptors by translating equivalent chemical representations. Chem Sci. 2019;10(6):1692-701.

[44] Yang X, Wang Y, Byrne R, Schneider G, Yang S. Concepts of artificial intelligence for computer-assisted drug discovery. Chem Rev. 2019;119(18):10520-94.

[45] Rifaioglu AS, Atas H, Martin MJ, Cetin-Atalay R, Atalay V, Doğan T. Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. Brief Bioinform. 2019;20(5):1878-912.

[46] Xu Y, Lin K, Wang S, Wang L, Cai C, Song C, et al. Deep learning for molecular generation. Future Med Chem. 2019;11(6):567-97.

[47] Davis MI, Hunt JP, Herrgard S, Ciceri P, Wodicka LM, Pallares G, et al. Comprehensive analysis of kinase inhibitor selectivity. Nat Biotechnol. 2011;29(11):1046-51.

[48] Konze KD, Bos PH, Dahlgren MK, Leswing K, Tubert-Brohman I, Bortolato A, et al. Reaction-based enumeration of potential synthetic routes for drug-like molecules. J Chem Inf Model. 2019;59(9):3782-93.

[49] Capuzzi SJ, Sun W, Muratov EN, Martínez-Romero C, He S, Zhu W, et al. Computer-aided discovery and characterization of novel Ebola virus inhibitors. J Med Chem. 2018;61(8):3582-94.

[50] Fischer A, Smieško M, Sellner M, Lill MA. Decision making in structure-based drug discovery: visual inspection of docking results. J Med Chem. 2021;64(5):2489-500.

[51] Wang Y, Cheng T, Bryant SH. PubChem BioAssay: a decade's development toward open high-throughput screening data sharing. SLAS Discov. 2017;22(6):655-66.

[52] Sydow D, Burggraaff L, Szengel A, van Vlijmen HWT, IJzerman AP, van Westen GJP, et al. Advances and challenges in computational target prediction. J Chem Inf Model. 2019;59(5):1728-42.

[53] Hie B, Zhong ED, Berger B, Bryson B. Learning the language of viral evolution and escape. Science. 2021;371(6526):284-8.

[54] Gao W, Coley CW. The synthesizability of molecules proposed by generative models. J Chem Inf Model. 2020;60(12):5714-23.

[55] Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, et al. Protein structure prediction using multiple deep neural networks in CASP13. Proteins. 2019;87(12):1141-8.

[56] Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, et al. ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res. 2019;47(D1):D930-40.

[57] Xu M, Luo X, McBride JD, Chang CCY, Julius D, Liang LF, et al. Structure-based prediction of ligand binding sites on G protein-coupled receptors. Nat Chem Biol. 2021;17(12):1293-9.

[58] Schneider P, Walters WP, Plowright AT, Sieroka N, Listgarten J, Goodnow RA, et al. Rethinking drug design in the artificial intelligence era. Nat Rev Drug Discov. 2020;19(5):353-64.

[59] Lyu J, Wang S, Balius TE, Singh I, Levit A, Moroz YS, et al. Ultra-large library docking for discovering new chemotypes. Nature. 2019;566(7743):224-9.

[60] Gorgulla C, Boeszoermenyi A, Wang ZF, Fischer PD, Coote PW, Padmanabha Das KM, et al. An open-source drug discovery platform enables ultra-large virtual screens. Nature. 2020;580(7805):663-8.