



Medicinal Chemistry in Recent Drug Discovery: Rational Design, Structure–Activity Relationships, and Future Perspectives

Qureshi Hudabiya Dawood^{1*}

¹Anjuman-I-Islam's Kalsekar Technical Campus, New Panvel, Navi Mumbai, M.H., PIN-410206

*Corresponding Author E-mail: huudaquuraish704@gmail.com

Abstract:

Medicinal chemistry occupies the interface between chemical structure and biological function, and its central working method – the iterative design–synthesis–test cycle guided by structure–activity relationships (SAR) – has remained conceptually stable even as the tools available to execute it have been transformed. This review surveys the current state of rational drug design across three converging strands. First, we examine the principal design paradigms in use today: structure-based drug design guided by target crystal or predicted structures, ligand-based and quantitative structure–activity relationship (QSAR) approaches, and fragment-based drug discovery, alongside the quantitative drug-likeness and ligand-efficiency criteria that guide compound optimisation. Second, we survey how SAR analysis has been extended beyond classical single-target potency optimisation to encompass bioisosteric replacement, natural product scaffold interrogation, and the deliberate, mechanism-guided design of targeted covalent inhibitors, illustrated by the rapid clinical translation of covalent KRAS(G12C) inhibitors. Third, we examine the growing integration of artificial intelligence and machine learning across this pipeline, from QSAR and absorption-distribution-metabolism-excretion-toxicity (ADMET) property prediction through to AI-guided protein structure prediction and generative molecular design, including a recently reported randomised clinical trial of a generative-AI-discovered candidate. We further discuss emerging design paradigms that extend medicinal chemistry beyond classical small-molecule occupancy, including proteolysis-targeting chimeras (PROTACs) and small-molecule RNA-targeting strategies. Finally, we critically examine the methodological and translational challenges that continue to constrain the field – including the persistent gap between computational prediction and experimental validation, the physicochemical property challenges posed by newer modalities, and the difficulty of generalising machine learning models beyond their training data – and outline priority directions for the continued evolution of rational drug design.

Keywords: medicinal chemistry; rational drug design; structure–activity relationship; structure-based drug design; fragment-based drug discovery; QSAR; ADMET; PROTAC; covalent inhibitors; artificial intelligence

Received: July 19, 2026

Revised: Aug. 29, 2026

Accepted: Sep. 12, 2026

Published: Oct. 30, 2026

DOI: <https://doi.org/10.64062/IJPCAT.Vol2.Issue5.9>

<https://ijpcat.com/index.php/1/issue/archive>

This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers. (<https://creativecommons.org/licenses/by-nc/4.0/>)

1. Introduction

Medicinal chemistry is fundamentally concerned with a deceptively simple question: how does a small change in chemical structure translate into a change in biological activity? The systematic pursuit of this question, formalised over the past half-century into the discipline of structure–activity relationship (SAR) analysis, has underpinned the discovery of the overwhelming majority of small-molecule drugs currently in clinical use. What has changed substantially over the past two to three decades is not the conceptual centrality of this question but the tools available to answer it: the field has progressed from largely empirical, trial-and-error synthesis and testing toward a rational design paradigm informed by target protein structures, quantitative physicochemical property models, and, most recently, machine learning models trained on increasingly large chemical and biological datasets.

This transition toward rational design was itself made possible by a series of conceptual and methodological milestones that established quantitative, generalisable criteria for compound optimisation in place of purely empirical medicinal chemist intuition. The formulation of drug-likeness criteria based on molecular weight, lipophilicity, and hydrogen-bonding capacity provided the field's first widely adopted quantitative filter for oral bioavailability potential¹, subsequently refined and extended by additional criteria addressing rotatable bond count and polar surface area². In parallel, the recognition that ligand potency should be evaluated relative to molecular size and complexity, rather than in absolute terms, gave rise to ligand efficiency metrics that remain central to modern lead optimisation decision-making³. These quantitative frameworks, together with the structure-based, fragment-based, and computational design paradigms discussed in Section 2, collectively define what is now understood as rational drug design.

This review is organised around three converging strands of the contemporary field. Section 2 surveys the principal rational design paradigms—structure-based, ligand-based, and fragment-based approaches. Section 3 examines how SAR analysis has been extended beyond classical potency optimisation to bioisosteric replacement, natural product interrogation, and covalent inhibitor design. Section 4 surveys the growing role of artificial intelligence and machine learning across this pipeline. Section 5 discusses emerging design paradigms, including PROTACs and RNA-targeting small molecules, that extend medicinal chemistry beyond classical small-molecule target occupancy, before Sections 6 through 8 present an illustrative case study, critically examine remaining challenges, and outline future perspectives for the field.

2. Principles of Rational Drug Design

2.1 Structure-Based Drug Design

Structure-based drug design (SBDD) uses experimentally determined or computationally predicted three-dimensional structures of a target protein, typically in complex with a bound ligand, to guide the design of new or optimised compounds through molecular docking, structure-guided fragment growth, and free-energy-based binding affinity prediction. Applied

extensively to kinase targets in oncology, in silico SBDD methods including homology modelling for targets lacking experimental structures, virtual screening of large compound libraries, and molecular dynamics simulation to assess binding pose stability have become standard components of the modern medicinal chemist's toolkit for prioritising synthetic effort toward compounds with the highest predicted probability of target engagement⁴. The reliability of SBDD depends critically on the quality of the underlying structural model; the recent availability of high-accuracy AI-predicted structures for targets lacking experimental crystal structures, discussed further in Section 4.2, has substantially expanded the range of targets amenable to this design paradigm.

2.2 Ligand-Based Design and Quantitative Structure–Activity Relationships

Where a validated three-dimensional target structure is unavailable, ligand-based design approaches instead build predictive models directly from the structures and measured activities of known active and inactive compounds. Quantitative structure–activity relationship (QSAR) modelling, the most established ligand-based methodology, correlates molecular descriptors encoding size, shape, electronic distribution, and lipophilicity with measured biological activity using statistical or machine learning regression methods, providing both a predictive tool for prioritising untested compounds and, through descriptor interpretation, mechanistic insight into which structural features drive activity. Contemporary QSAR practice increasingly incorporates machine learning regression and classification algorithms in place of classical linear regression, a development discussed further in Section 4.1.

2.3 Fragment-Based Drug Discovery

Fragment-based drug discovery (FBDD) begins from a deliberately different starting point than either SBDD or classical high-throughput screening: rather than screening drug-sized compounds directly, FBDD screens libraries of very small, low-molecular-weight chemical fragments (typically under 300 Da) that individually bind weakly but efficiently, following the empirical “rule of three” guideline that well-behaved fragment libraries should satisfy correspondingly scaled-down versions of classical drug-likeness criteria⁶. A decade-later reassessment of this rule of three concluded that, while useful as an initial filter, the criterion should be applied flexibly rather than as a rigid cut-off, reflecting broader lessons learned as the field matured⁷. Selected fragments are then elaborated grown, linked, or merged with other fragments into higher-affinity leads, an approach whose twenty-year track record has been credited with contributing to several approved drugs and numerous clinical candidates across diverse target classes⁸.

A comprehensive historical and forward-looking review of FBDD methodology traces the field's evolution from its origins in nuclear magnetic resonance (NMR)-based fragment screening through to contemporary integration with cryo-electron microscopy, DNA-encoded library technology, and computational fragment growth prediction, and outlines several emerging directions likely to shape the field's next phase⁹. Practical guidance on fragment screening and hit validation methodology, essential given the weak binding affinities characteristic of fragment hits and the correspondingly high risk of false-positive identification, has been comprehensively catalogued in recent methodological reviews¹⁰, while an annual survey of fragment-to-lead medicinal chemistry publications documents the continued year-

on-year productivity of this design paradigm across a broad range of therapeutic areas and target classes¹¹.

FBDD has proven particularly valuable for historically challenging target classes. Fragment-based approaches applied to RNA targets, which present flatter, more solvent-exposed and dynamic binding surfaces than the well-defined hydrophobic pockets typical of most protein drug targets, have identified RNA-binding fragment starting points for subsequent medicinal chemistry elaboration¹², while fragment-based NMR screening applied to allosteric sites on protein kinases such as NIK has identified starting-point ligands for targets that have proven difficult to address using orthosteric, active-site-directed approaches¹³.

Table 1. Comparison of the principal rational drug design paradigms discussed in Section 2

Design paradigm	Starting point / input	Key strength	Key limitation
Structure-based (SBDD)	Experimental or AI-predicted 3D target structure	Direct structural rationale for binding; efficient for well-characterised, tractable pockets	Requires reliable structural model; less effective for flexible or poorly defined binding sites
Ligand-based / QSAR	Structure–activity data for known active/inactive compounds	Applicable without a target structure; captures empirical SAR trends	Predictive only within the chemical space of the training data; limited mechanistic insight
Fragment-based (FBDD)	Library of low-molecular-weight (<300 Da) fragments	High ligand efficiency; explores chemical space efficiently; suited to challenging/allosteric sites	Weak initial binding affinity; requires sensitive biophysical detection methods; longer optimisation runway
AI/generative design	Large bioactivity, structural, or generative-model training data	Rapid, scalable candidate generation; can propose novel chemical matter	Model generalisation and interpretability limitations; requires experimental validation

3. Structure–Activity Relationships: From Classical SAR to Covalent Design

3.1 Classical SAR and Bioisosteric Replacement

Classical SAR analysis proceeds through systematic, incremental structural variation of a lead compound substituent addition, removal, or replacement at defined positions correlated against measured changes in potency, selectivity, and physicochemical properties, building an empirical map of which structural features are essential for activity and which can be modified to improve drug-like properties without loss of potency. Bioisosteric replacement, in which a functional group is substituted with another group of similar size, shape, or electronic character, is one of the most widely used tactics within this SAR-driven optimisation process, allowing medicinal chemists to address metabolic liabilities, improve selectivity, or circumvent intellectual property constraints while preserving the essential pharmacophoric features responsible for target engagement.

3.2 Structure–Activity Relationships in Natural Product Chemistry

Natural products present a distinctive SAR challenge relative to fully synthetic chemical matter: their structural complexity, frequent stereochemical intricacy, and biosynthetic rather

than synthetic origin have historically made systematic analogue synthesis for SAR interrogation far more resource-intensive than for simpler synthetic scaffolds. A comprehensive recent review of methods for interrogating natural product SAR surveys both classical semisynthetic analogue generation and newer approaches including biosynthetic pathway engineering to generate analogues directly in the producing organism, and chemoenzymatic synthesis combining enzymatic and synthetic steps that are beginning to make natural product SAR studies more tractable at a scale comparable to fully synthetic medicinal chemistry programmes¹⁶. This is a consequential development given natural products' historically disproportionate contribution to approved drugs across nearly every major therapeutic area, and given the multi-target polypharmacological character of many natural scaffolds, which network pharmacology methodology has increasingly been used to characterise systematically^{17,18}, albeit with methodological limitations around database annotation quality and experimental validation that have been the subject of recent critical review¹⁹.

3.3 Targeted Covalent Inhibitor Design

A structurally distinct and increasingly prominent SAR paradigm involves the deliberate design of covalent inhibitors, in which a reactive electrophilic “warhead” moiety, appropriately positioned by the surrounding reversible binding scaffold, forms an irreversible or slowly reversible covalent bond with a nucleophilic amino acid residue most commonly cysteine within the target binding site. Although covalently acting drugs such as aspirin and β -lactam antibiotics have a long history in medicine, deliberate, structure-guided covalent inhibitor design was for decades viewed with suspicion within the field owing to concerns about off-target reactivity and immunogenic haptization; a comprehensive review traces how improved understanding of warhead reactivity–selectivity relationships, quantified through kinetic parameters describing the rate and equilibrium of covalent bond formation, has enabled a substantial resurgence of rationally designed covalent drugs over the past two decades²⁰.

This resurgence is most vividly illustrated by the covalent targeting of KRAS(G12C), a GTPase long considered an “undruggable” oncology target owing to its lack of a deep, well-defined small-molecule binding pocket and its picomolar affinity for endogenous GTP. Covalent inhibitors exploiting a mutation-specific cysteine adjacent to a shallow allosteric pocket enabled selective, irreversible inactivation of the mutant but not wild-type protein, and the resulting compound, sotorasib, became the first FDA-approved KRAS inhibitor in 2021, alongside the subsequently approved adagrasib, demonstrating that covalent chemistry can expand the druggable target space to proteins previously considered inaccessible to classical reversible small-molecule inhibition^{21,22}. A systematic taxonomy of covalent inhibition mechanisms, cataloguing the distinct chemical strategies by which different electrophilic warheads engage nucleophilic residues, provides the mechanistic framework underlying this and subsequent covalent drug design programmes²³, while a broader historical review of the field's resurgence documents how rational warhead selection and residence-time optimisation have progressively expanded covalent chemistry's applicability well beyond its original, largely serendipitous origins²⁴.

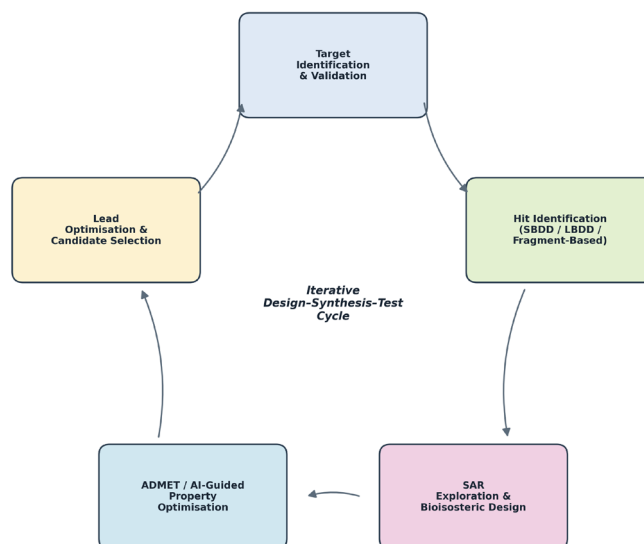


Figure 1. The iterative rational drug design cycle in modern medicinal chemistry, integrating the design paradigms discussed in Section 2 with the SAR, ADMET, and AI-guided optimisation strategies discussed in Sections 3–4.

4. Computational and Artificial Intelligence-Driven Approaches

4.1 Machine Learning in QSAR and ADMET Prediction

Machine learning methods have substantially extended classical QSAR methodology, replacing linear regression on hand-selected descriptors with learned molecular representations including graph neural networks operating directly on molecular structure capable of capturing more complex, non-linear structure–activity relationships. A systematic study of the key elements underlying molecular property prediction models found that data quality, splitting strategy, and representation choice each contribute substantially and sometimes counter-intuitively to reported model performance, underscoring that headline accuracy figures for a given machine learning architecture cannot be interpreted independently of these methodological choices²⁷.

ADMET property prediction has become one of the most actively developed applications of machine learning in medicinal chemistry, given the central role early ADMET failure has historically played in costly late-stage clinical attrition. Systematic denoising of historical ADMET training data, correcting for measurement variability and inconsistent experimental protocols across data sources, has been shown to measurably improve downstream model prediction accuracy²⁵, while industrial prospective validation studies in which model predictions are tested against compounds synthesised and measured after model training, rather than only against retrospectively held-out data provide a more stringent and practically relevant assessment of real-world predictive utility than standard cross-validation alone²⁶. Recently proposed method-comparison protocols aim to establish minimum standards of practical significance, rather than purely statistical significance, when benchmarking competing machine learning approaches for small-molecule property prediction, addressing a

longstanding methodological gap in how the field evaluates and compares newly published models²⁸.

Integrated, multi-endpoint ADMET prediction platforms combining self-supervised pretraining with knowledge transfer across related property-prediction tasks have demonstrated robust performance across diverse absorption, distribution, metabolism, and excretion endpoints simultaneously²⁹, while therapeutic-area-specific applications, such as machine-learning-guided QSAR modelling for antidiabetic α -glucosidase inhibitor discovery, illustrate how these general-purpose methodological advances are being adapted to specific disease-focused medicinal chemistry programmes³⁰. A fully automated AI/ML virtual screening cascade recently implemented at a drug discovery centre in Africa further illustrates how these computational tools are lowering the infrastructural barrier to conducting rational, machine-learning-guided medicinal chemistry outside traditionally resource-concentrated research hubs³¹, and complementary in silico ADMET review work continues to consolidate best-practice guidance for applying these tools within standard medicinal chemistry workflows³².

4.2 AI-Guided Structure Prediction

A central practical constraint on structure-based drug design, discussed in Section 2.1, has historically been the availability of a sufficiently accurate three-dimensional structure for the target protein, particularly in complex with a bound ligand. The release of AlphaFold 3, capable of predicting the structure of protein–ligand, protein–nucleic acid, and protein–protein complexes with accuracy approaching experimental structure determination for many complex types, has substantially expanded the number of medicinal chemistry targets for which a high-quality structural model is available to guide design, including targets that lack an experimentally solved crystal structure³³. This development is expected to be particularly consequential for challenging target classes including many RNA-binding proteins and allosteric sites discussed in Sections 2.3 and 5.2 for which experimental structure determination has historically proven especially difficult.

4.3 Generative Molecular Design

Beyond property prediction and structure determination, generative AI models can propose entirely novel candidate structures optimised computationally for predicted potency, selectivity, and drug-likeness. The clinical significance of this approach was demonstrated by a randomised phase 2a trial of a small-molecule TNIK inhibitor for idiopathic pulmonary fibrosis discovered using a generative AI platform, representing one of the first published examples of an AI-generated drug candidate progressing to a randomised controlled human efficacy trial³⁴. This result provides a proof-of-concept precedent directly relevant to the broader medicinal chemistry field's growing use of generative design methods, and is likely to accelerate adoption of similar generative approaches across the design paradigms discussed throughout this review.

Table 2. Representative applications of artificial intelligence and machine learning across the medicinal chemistry pipeline

Pipeline stage	AI/ML method class	Representative contribution	Ref.
QSAR / property prediction	Graph neural networks; learned molecular representations	Improved capture of non-linear structure–activity relationships beyond classical descriptor-based QSAR	27
ADMET prediction	Multi-task deep learning; self-supervised knowledge transfer	Robust, endpoint-extensible prediction across absorption, distribution, metabolism, excretion	25,26,29
Model evaluation methodology	Prospective validation; practical-significance benchmarking protocols	More stringent, real-world-relevant assessment of predictive model utility	26,28
Structure prediction	Deep learning structure prediction (AlphaFold 3)	High-accuracy prediction of protein and protein–ligand complex structures	33
Generative design	Generative models (VAEs, diffusion/transformer architectures)	De novo candidate generation; clinical translation demonstrated for a generative-AI-discovered TNiK inhibitor	34
Decentralised/global drug discovery	Automated AI/ML virtual screening cascades	Lowering infrastructural barriers to rational design outside traditional research hubs	31

5. Emerging Design Paradigms Beyond Classical Occupancy

5.1 Targeted Protein Degradation: PROTACs

Proteolysis-targeting chimeras (PROTACs) represent a qualitatively distinct medicinal chemistry design paradigm from the classical occupancy-based approaches discussed in Section 2. A PROTAC is a heterobifunctional molecule linking a target-protein ligand to an E3 ubiquitin ligase recruiter, inducing target ubiquitination and proteasomal degradation rather than sustained occupancy-dependent inhibition. Because PROTACs must simultaneously present two independent binding pharmacophores connected through an appropriately spaced linker, their design and synthesis pose distinct SAR challenges around linker length and geometry optimisation that differ qualitatively from classical single-pharmacophore SAR³⁵. Click chemistry has proven particularly well suited to addressing these PROTAC-specific synthetic and SAR-exploration challenges, enabling modular, combinatorial assembly and screening of candidate degrader molecules from independently synthesised building blocks^{36,37}.

5.2 Small-Molecule Targeting of RNA

A second emerging design paradigm extends medicinal chemistry beyond the traditional focus on protein targets to small molecules that selectively bind structured RNA elements, exploiting the recognition that many disease-relevant RNAs adopt specific secondary and tertiary structures capable of selective small-molecule recognition. A comprehensive review of small-molecule approaches to targeting RNA surveys both structure-based and fragment-based strategies applied to this target class, while complementary work on targeting RNA-binding proteins with small molecules, including bifunctional degrader-like approaches analogous to PROTACs, illustrates how the induced-proximity design logic discussed in Section 5.1 is being

generalised beyond classical E3-ligase-mediated protein degradation to a broader family of RNA- and RNA-protein-complex-directed therapeutic strategies^{38,39}.

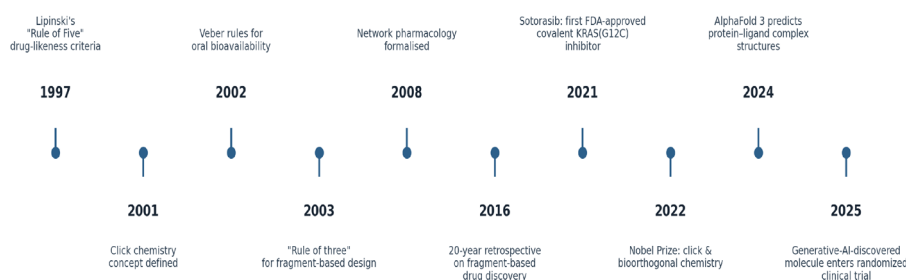


Figure 2. Timeline of key conceptual and technological milestones in rational drug design, from early drug-likeness criteria through to AI-guided structure prediction and generative design.

6. Illustrative Case Study: Structure-Guided Design of Covalent KRAS(G12C) Inhibitors

The development of sotorasib and adagrasib, discussed in Section 3.3, offers a particularly instructive case study illustrating how the design paradigms surveyed across this review converge in a single successful drug discovery programme. KRAS(G12C) was for decades considered undruggable using classical reversible small-molecule approaches, since the protein's smooth surface topology, lack of a deep binding pocket, and picomolar affinity for endogenous GTP defeated conventional structure-based and fragment-based screening strategies applied on their own²². The eventual solution combined fragment-based screening, discussed in Section 2.3, to identify weakly binding starting points that engaged a previously unrecognised shallow allosteric pocket adjacent to the mutant cysteine residue, with covalent warhead SAR optimisation, discussed in Section 3.3, to convert this weak, reversible fragment interaction into a potent, selective, irreversible inhibitor²⁴. Only eight years elapsed between the first published demonstration of covalent binding to this site and FDA approval of sotorasib, a translational timeline substantially shorter than the historical average for first-in-class oncology therapeutics, illustrating the potential for rational, structure- and mechanism-guided design to accelerate drug development relative to less systematic discovery approaches²⁴. This case study exemplifies the iterative design–synthesis–test cycle summarised in Figure 1, integrating fragment-based hit identification, structure-guided SAR optimisation, and covalent mechanism engineering within a single, rationally executed programme⁴⁰.

7. Challenges and Limitations

Despite the substantial methodological progress surveyed across this review, several challenges continue to constrain the practical impact of rational drug design.

Table 3. Principal challenges currently facing rational drug design and SAR-driven medicinal chemistry

Challenge	Description	Illustrative consequence
Prediction-validation gap	Computational SBDD, QSAR, and generative design predictions frequently outpace experimental validation capacity	A substantial proportion of computationally prioritised compounds fail experimental confirmation
Model generalisation limits	ML/QSAR models trained on historical chemical space may generalise poorly to genuinely novel scaffolds	Reduced predictive reliability for the most innovative, furthest-from-precedent candidate structures
Newer-modality physicochemical properties	PROTACs and other heterobifunctional/induced-proximity molecules routinely exceed classical drug-likeness criteria	Constrained oral bioavailability; smaller accessible target/indication space for first-generation molecules
Covalent selectivity and safety	Electrophilic warheads carry inherent risk of off-target reactivity toward untargeted nucleophilic residues	Requires careful kinetic (kinact/KI) optimisation; narrower acceptable reactivity window than reversible inhibitors
Natural product SAR tractability	Structural complexity and biosynthetic origin complicate systematic natural product analogue synthesis	Slower, more resource-intensive SAR interrogation relative to fully synthetic chemical matter
Benchmarking and reproducibility	Inconsistent data splitting, descriptor choice, and evaluation protocols across published ML/QSAR studies	Difficulty comparing reported model performance fairly across the literature

A further, cross-cutting challenge concerns the integration of these individually maturing methodologies into a coherent, efficient overall workflow. Structure-based design, fragment-based screening, machine-learning-guided property prediction, and generative design have each been extensively validated in isolation, yet relatively few published case studies document their systematic, sequential combination within a single discovery programme with the rigor applied to the KRAS(G12C) case study in Section 6; realising the full productivity gains this convergence of methodologies theoretically offers will require closer methodological integration than is currently standard practice across much of the field.

8. Future Perspectives

Several directions are likely to shape the continued evolution of rational drug design. First, continued expansion of AI-predicted structural coverage to historically challenging target classes including RNA-binding proteins, membrane proteins, and intrinsically disordered regions should progressively extend structure-based design to targets currently accessible only through ligand-based or fragment-based approaches³³. Second, closer integration of generative molecular design with explicit multi-objective optimisation simultaneously targeting potency, selectivity, and favourable ADMET properties rather than optimising these properties sequentially represents a natural extension of the generative design methods discussed in Section 4.3, building on the demonstrated clinical feasibility of generative-AI-discovered candidates.

Third, continued refinement of induced-proximity and covalent design paradigms PROTACs, molecular glues, RNA-targeting small molecules, and mechanism-optimised covalent inhibitors is likely to extend rational, structure- and mechanism-guided medicinal chemistry to an increasing proportion of the historically “undruggable” target space, following the template established by the KRAS(G12C) case study discussed in Section 6³⁵³⁸. Fourth,

prospective, pre-registered validation of computational predictions against genuinely novel experimental data, rather than retrospective benchmarking against training-adjacent datasets, will be essential to establishing confidence in the real-world predictive validity of the machine learning methods increasingly embedded throughout the modern medicinal chemistry pipeline^{28,39}.

Finally, achieving the full productivity potential of this increasingly computational and multi-modal field will depend on training medicinal chemists who are equally fluent in classical SAR reasoning, structural biology, and machine learning evaluation methodology, so that computational predictions are neither uncritically accepted nor reflexively dismissed, but are integrated as one additional, rigorously interpreted source of evidence within the same iterative design–synthesis–test cycle that has defined rational medicinal chemistry since its earliest formalisation.

9. Conclusion

Medicinal chemistry's core method – the iterative, SAR-guided design–synthesis–test cycle summarised in Figure 1 – has proven remarkably durable even as the tools available to execute each stage of that cycle have been repeatedly transformed, from the formulation of quantitative drug-likeness criteria in the late 1990s through structure-based and fragment-based design paradigms to the current integration of machine learning and generative AI across nearly every stage of the pipeline. Structure–activity relationship analysis has similarly expanded well beyond its classical single-target potency-optimisation origins to encompass bioisosteric design, natural product interrogation, and the deliberate, mechanism-guided engineering of covalent and induced-proximity therapeutic modalities, illustrated concretely by the rapid, rationally executed translation of covalent KRAS(G12C) inhibitors from first mechanistic demonstration to FDA approval. Persistent challenges around the gap between computational prediction and experimental validation, the physicochemical properties of newer therapeutic modalities, and model generalisation and benchmarking rigor will continue to constrain the pace of translation, but the trajectory traced across this review – toward increasingly rational, structure- and data-guided design applied to an expanding range of previously inaccessible target classes – represents the clearest current path toward continued productivity gains in medicinal chemistry and modern drug discovery.

References

1. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 23(1–3), 3–25. [https://doi.org/10.1016/S0169-409X\(96\)00423-1](https://doi.org/10.1016/S0169-409X(96)00423-1)
2. Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, 45(12), 2615–2623. <https://doi.org/10.1021/jm020017n>

3. Hopkins, A. L., Keserü, G. M., Leeson, P. D., Rees, D. C., & Reynolds, C. H. (2014). The role of ligand efficiency metrics in drug discovery. *Nature Reviews Drug Discovery*, 13(2), 105–121. <https://doi.org/10.1038/nrd4163>
4. Gagic, Z., Ruzic, D., Djokovic, N., Djikic, T., & Nikolic, K. (2020). In silico methods for design of kinase inhibitors as anticancer drugs. *Frontiers in Chemistry*, 7, Article 873. <https://doi.org/10.3389/fchem.2019.00873>
5. Sheng, C., & Zhang, W. (2013). Fragment informatics and computational fragment-based drug design: An overview and update. *Medicinal Research Reviews*, 33(3), 554–598. <https://doi.org/10.1002/med.21255>
6. Congreve, M., Carr, R., Murray, C., & Jhoti, H. (2003). A ‘rule of three’ for fragment-based lead discovery? *Drug Discovery Today*, 8(19), 876–877. [https://doi.org/10.1016/S1359-6446\(03\)02831-9](https://doi.org/10.1016/S1359-6446(03)02831-9)
7. Jhoti, H., Williams, G., Rees, D. C., & Murray, C. W. (2013). The ‘rule of three’ for fragment-based drug discovery: Where are we now? *Nature Reviews Drug Discovery*, 12(8), 644–645. <https://doi.org/10.1038/nrd3926-c1>
8. Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W., & Jhoti, H. (2016). Twenty years on: The impact of fragments on drug discovery. *Nature Reviews Drug Discovery*, 15(9), 605–619. <https://doi.org/10.1038/nrd.2016.109>
9. Xu, W., & Kang, C. (2025). Fragment-based drug design: From then until now, and toward the future. *Journal of Medicinal Chemistry*, 68(15), 15234–15256. <https://doi.org/10.1021/acs.jmedchem.5c00424>
10. Kirkman, T., dos Santos Silva, C., Tosin, M., & Bertacine Dias, M. V. (2024). How to find a fragment: Methods for screening and validation in fragment-based drug discovery. *ChemMedChem*, 19(24), Article e202400342. <https://doi.org/10.1002/cmdc.202400342>
11. Holvey, R. S., Erlanson, D. A., de Esch, I. J. P., Farkaš, B., Jahnke, W., Nishiyama, T., & Woodhead, A. J. (2025). Fragment-to-lead medicinal chemistry publications in 2023. *Journal of Medicinal Chemistry*, 68(2), 986–1001. <https://doi.org/10.1021/acs.jmedchem.4c02593>
12. Suresh, B. M., Taghavi, A., Childs-Disney, J. L., & Disney, M. D. (2023). Fragment-based approaches to identify RNA binders. *Journal of Medicinal Chemistry*, 66(10), 6523–6541. <https://doi.org/10.1021/acs.jmedchem.3c00034>
13. Anderson, J. J., Grillo, M. J., & Harki, D. A. (2023). Development of allosteric NIK ligands from fragment-based NMR screening. *ACS Medicinal Chemistry Letters*, 14(12), 1815–1820. <https://doi.org/10.1021/acsmedchemlett.3c00429>
14. Patani, G. A., & LaVoie, E. J. (1996). Bioisosterism: A rational approach in drug design. *Chemical Reviews*, 96(8), 3147–3176. <https://doi.org/10.1021/cr950066q>
15. Bleicher, K. H., Böhm, H.-J., Müller, K., & Alanine, A. I. (2003). Hit and lead generation: Beyond high-throughput screening. *Nature Reviews Drug Discovery*, 2(5), 369–378. <https://doi.org/10.1038/nrd1086>

16. Ancajas, C. M. F., Oyedele, A. S., Butt, C. M., & Walker, A. S. (2024). Advances, opportunities, and challenges in methods for interrogating the structure activity relationships of natural products. *Natural Product Reports*, 41, 1543–1578. <https://doi.org/10.1039/D4NP00009A>
17. Li, S., Fan, T.-P., Jia, W., Lu, A., & Zhang, W. (2014). Network pharmacology in traditional Chinese medicine. *BioMed Research International*, 2014, Article 138460. <https://doi.org/10.1155/2014/138460>
18. Hopkins, A. L. (2008). Network pharmacology: The next paradigm in drug discovery. *Nature Chemical Biology*, 4(11), 682–690. <https://doi.org/10.1038/nchembio.118>
19. Panossian, A. (2025). Trends and pitfalls in the progress of network pharmacology research on natural products. *Pharmaceuticals*, 18(4), Article 538. <https://doi.org/10.3390/ph18040538>
20. Serafim, R. A. M., Gehringer, M., & Borsari, C. (2024). Targeted covalent inhibitors in drug discovery, chemical biology and beyond. *Pharmaceuticals*, 17(2), Article 206. <https://doi.org/10.3390/ph17020206>
21. Singh, J., Petter, R. C., Baillie, T. A., & Whitty, A. (2011). The resurgence of covalent drugs. *Nature Reviews Drug Discovery*, 10(4), 307–317. <https://doi.org/10.1038/nrd3410>
22. Tuley, A., & Fast, W. (2018). The taxonomy of covalent inhibitors. *Biochemistry*, 57(24), 3326–3337. <https://doi.org/10.1021/acs.biochem.8b00315>
23. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A., & Shokat, K. M. (2013). K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*, 503(7477), 548–551. <https://doi.org/10.1038/nature12796>
24. Ostrem, J. M. L., & Shokat, K. M. (2016). Direct small-molecule inhibitors of KRAS: From structural insights to mechanism-based design. *Nature Reviews Drug Discovery*, 15(11), 771–785. <https://doi.org/10.1038/nrd.2016.139>
25. Adrian, M., Chung, Y., & Cheng, A. C. (2024). Denoising drug discovery data for improved absorption, distribution, metabolism, excretion, and toxicity property prediction. *Journal of Chemical Information and Modeling*, 64(16), 6324–6337. <https://doi.org/10.1021/acs.jcim.4c00639>
26. Fang, C., Wang, Y., Grater, R., Kapadnis, S., Black, C., Trapa, P., & Sciabola, S. (2023). Prospective validation of machine learning algorithms for absorption, distribution, metabolism, and excretion prediction: An industrial perspective. *Journal of Chemical Information and Modeling*, 63(11), 3263–3274. <https://doi.org/10.1021/acs.jcim.3c00160>
27. Deng, J., Yang, Z., Wang, H., Ojima, I., Samaras, D., & Wang, F. (2023). A systematic study of key elements underlying molecular property prediction. *Nature Communications*, 14, Article 6395. <https://doi.org/10.1038/s41467-023-41948-6>
28. Ash, J. R., Wognum, C., Rodríguez-Pérez, R., Aldeghi, M., Cheng, A. C., Clevert, D.-A., Engkvist, O., Fang, C., Price, D. J., Hughes-Oliver, J. M., & Walters, W. P. (2025). Practically significant method comparison protocols for machine learning in small

- molecule drug discovery. *Journal of Chemical Information and Modeling*, 65(18), 9398–9411. <https://doi.org/10.1021/acs.jcim.5c01352>
29. Zhang, S., Yan, Z., Huang, Y., Liu, L., He, D., Wang, W., Fang, X., Zhang, X., Wang, F., Wu, H., & Wang, H. (2022). HelixADMET: A robust and endpoint extensible ADMET system incorporating self-supervised knowledge transfer. *Bioinformatics*, 38(13), 3444–3453. <https://doi.org/10.1093/bioinformatics/btac342>
 30. Odugbemi, A. I., Nyirenda, C., Christoffels, A., & Egieyeh, S. A. (2024). Artificial intelligence in antidiabetic drug discovery: The advances in QSAR and the prediction of α -glucosidase inhibitors. *Computational and Structural Biotechnology Journal*, 23, 2964–2977. <https://doi.org/10.1016/j.csbj.2024.07.003>
 31. Turon, G., Hlozek, J., Woodland, J. G., Kumar, A., Chibale, K., & Duran-Frigola, M. (2023). First fully-automated AI/ML virtual screening cascade implemented at a drug discovery centre in Africa. *Nature Communications*, 14, Article 5736. <https://doi.org/10.1038/s41467-023-41512-2>
 32. Agrawal, S., Pandey, P., Jacob, A., Tiwari, V., Dixit, R. K., & Kumar, R. (2026). In silico ADMET: Supporting new drug discovery. *Future Health*, 4, 49–56. https://doi.org/10.25259/FH_100_2025
 33. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
 34. Xu, Z., Ren, F., Wang, P., Cao, J., Tan, C., Ma, D., et al. (2025). A generative AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis: A randomized phase 2a trial. *Nature Medicine*, 31, 2602–2610. <https://doi.org/10.1038/s41591-025-03743-2>
 35. Osman, J., Thompson, P. E., Jörg, M., & Scanlon, M. J. (2025). Methods to accelerate PROTAC drug discovery. *Biochemical Journal*, 482(13), 921–937. <https://doi.org/10.1042/BCJ20243018>
 36. Yang, C., Tripathi, R., & Wang, B. (2024). Click chemistry in the development of PROTACs. *RSC Chemical Biology*, 5(3), 189–197. <https://doi.org/10.1039/d3cb00199g>
 37. Kolb, H. C., Finn, M. G., & Sharpless, K. B. (2001). Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie International Edition*, 40(11), 2004–2021. [https://doi.org/10.1002/1521-3773\(20010601\)40:11<2004::AID-ANIE2004>3.0.CO;2-5](https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5)
 38. Kovachka, S., Panosetti, M., Grimaldi, B., Azoulay, S., Di Giorgio, A., & Duca, M. (2024). Small molecule approaches to targeting RNA. *Nature Reviews Chemistry*, 8(2), 120–135. <https://doi.org/10.1038/s41570-023-00569-9>
 39. Li, Q., & Kang, C. (2023). Targeting RNA-binding proteins with small molecules: Perspectives, pitfalls and bifunctional molecules. *FEBS Letters*, 597(16), 2031–2047. <https://doi.org/10.1002/1873-3468.14710>

40. Agard, N. J., Prescher, J. A., & Bertozzi, C. R. (2004). A strain-promoted [3 + 2] azide–alkyne cycloaddition for covalent modification of biomolecules in living systems. *Journal of the American Chemical Society*, 126(46), 15046–15047. <https://doi.org/10.1021/ja044996f>