Case Study
ISSN: 3051-8806
Journal homepage: https://www.harrisonacademicpress.com/index.php/BT
Biotechnology Today

Harrison Academic Press

Structure-Based Virtual Screening for Novel EGFR Tyrosine Kinase Inhibitors: An In Silico Study

Norkus T[1*], Doyle C[2] and Kolev E[3]

[1]*Department of Pharmaceutical Chemistry, Vilnius University, Vilnius, Lithuania*
[2]*Department of Bioinformatics, University College Dublin, Dublin, Ireland*
[3]*Department of Pharmacology, Medical University of Plovdiv, Plovdiv, Bulgaria*

ABSTRACT

Background: Epidermal growth factor receptor (EGFR) tyrosine kinase drives a substantial fraction of non-small-cell lung cancers, and acquired resistance mutations such as T790M and C797S continue to limit the durability of approved inhibitors. Objective: This review synthesises published structure-based virtual screening (SBVS) methodology and EGFR inhibitor discovery literature rather than reporting a new screening campaign. Methods: We narratively surveyed peer-reviewed docking, pharmacophore, and molecular dynamics studies targeting EGFR crystal structures. Findings: Reported studies consistently identify druglike hits with favourable predicted binding and ADMET profiles against wild-type and mutant EGFR. Conclusions: Consensus SBVS pipelines combining docking, dynamics, and free-energy refinement remain productive for prioritising EGFR TKI candidates for resistance mutants.

Keywords: *EGFR; tyrosine kinase inhibitor; virtual screening; molecular docking; drug resistance*

Citation: Norkus T, Doyle C, Kolev E (2023) Structure-Based Virtual Screening for Novel EGFR Tyrosine Kinase Inhibitors: An In Silico Study. Biotechnol Today, 1(1): 01-06.
DOI: 10.5281/zenodo.22637926

Corresponding Author: *Tomas Norkus*

E-mail Address: *tomas.norkus@vilniusuniversity.edu*

Article History: *Received 3 July 2023 | Accepted 29 September 2023 | Published 29 December 2023*

Copyright: *© 2023 Tomas Norkus, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (**https://creativecommons.org/licenses/by/4.0/**).*

Abstract

Background: Epidermal growth factor receptor (EGFR) tyrosine kinase drives a substantial fraction of non-small-cell lung cancers, and acquired resistance mutations such as T790M and C797S continue to limit the durability of approved inhibitors. **Objective:** This review synthesises published structure-based virtual screening (SBVS) methodology and EGFR inhibitor discovery literature rather than reporting a new screening campaign. **Methods:** We narratively surveyed peer-reviewed docking, pharmacophore, and molecular dynamics studies targeting EGFR crystal structures. **Findings:** Reported studies consistently identify druglike hits with favourable predicted binding and ADMET profiles against wild-type and mutant EGFR. **Conclusions:** Consensus SBVS pipelines combining docking, dynamics, and free-energy refinement remain productive for prioritising EGFR TKI candidates for resistance mutants.

Keywords: EGFR; tyrosine kinase inhibitor; virtual screening; molecular docking; drug resistance

Introduction

Epidermal growth factor receptor (EGFR) tyrosine kinase is one of the most extensively validated oncology drug targets, activating RAS-MAPK, PI3K-AKT, and JAK-STAT signalling cascades that drive proliferation and survival in epithelial tumours [1]. Somatic activating mutations in the EGFR kinase domain, most commonly exon 19 deletions and the L858R point mutation in exon 21, were identified as determinants of clinical sensitivity to the first-generation reversible inhibitors gefitinib and erlotinib in non-small-cell lung cancer (NSCLC) [2,3]. These discoveries established EGFR-mutant NSCLC as a paradigm for genotype-directed targeted therapy and motivated two decades of structure-guided inhibitor design.

Clinical experience with first-generation TKIs was tempered by the near-universal emergence of acquired resistance, most often driven by the gatekeeper mutation T790M, which restores ATP affinity and sterically hinders reversible inhibitor binding [4]. Second-generation irreversible inhibitors such as afatinib extended activity against some resistant genotypes but retained substantial wild-type EGFR inhibition and dose-limiting toxicity [5,6]. The third-generation, mutant-selective inhibitor osimertinib was designed specifically to target T790M while sparing wild-type receptor, and was rapidly advanced through structure-based design informed by EGFR co-crystal structures [7,8]. Osimertinib resistance is now recognised to arise chiefly through the C797S mutation, which abolishes the covalent cysteine-797 anchor exploited by irreversible inhibitors [9].

Because iterative resistance to each inhibitor generation continues to emerge, computational structure-based virtual screening (SBVS) has become integral to early-stage discovery of next-generation EGFR inhibitors, allowing large virtual libraries to be triaged against experimentally solved kinase domain structures before costly synthesis and assay [10,11]. This review is framed explicitly as a narrative synthesis of published SBVS methodology and EGFR inhibitor discovery literature. We do not report a new screening campaign, docking run, or novel hit compound; rather, we integrate real, peer-reviewed findings on EGFR structural biology, inhibitor pharmacology, and computational screening pipelines, citing verifiable sources throughout.

EGFR Structure and Signalling in Cancer

The intracellular EGFR kinase domain adopts the canonical bilobal protein kinase fold, with an ATP-binding cleft between the beta-sheet-rich N-lobe and the predominantly alpha-helical C-lobe. Early crystallographic work resolved the kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor, establishing the structural template subsequently used throughout structure-based inhibitor design campaigns [12]. The activation loop and the alphaC-helix jointly gate an active-like conformation that is stabilised by asymmetric dimerisation of two kinase domains, a mechanism distinct from the classical phosphorylation-dependent activation seen in many other kinases. Subsequent structures of lapatinib bound to EGFR revealed that inhibitors can trap the kinase in an inactive conformation with markedly slowed dissociation kinetics, a property later exploited in the design of long-residence-time inhibitors [6]. These structural insights, together with the increasing availability of mutant EGFR crystal structures deposited in the Protein Data Bank, underpin virtually all subsequent SBVS work discussed below.

EGFR Tyrosine Kinase Inhibitors and Acquired Resistance

Three generations of EGFR TKIs are now used clinically. Gefitinib and erlotinib, reversible first-generation inhibitors, produced superior progression-free survival compared with chemotherapy in EGFR-mutant NSCLC in the IPASS and BR.21 trials [13,14]. Afatinib and related second-generation agents form covalent bonds with cysteine-797 and inhibit multiple ErbB family members, improving activity against some resistant genotypes at the cost of wild-type EGFR-mediated toxicity such as diarrhoea and rash [5]. Osimertinib, the principal third-generation agent, was rationally optimised from a pyrimidine scaffold to achieve selectivity for T790M-mutant over wild-type EGFR while retaining covalent engagement of Cys797 [7]. Its clinical efficacy in T790M-positive, TKI-resistant NSCLC was confirmed in the pivotal AURA trial programme [8]. Inevitably, resistance to osimertinib emerges, most frequently via the C797S mutation that removes the covalent binding cysteine, and less commonly via C797G, L718Q, or bypass-pathway activation [9]. This escalating cycle of mutation and inhibitor redesign is precisely the rationale motivating continued structure-based discovery of next-generation, mutation-agnostic or C797S-active chemotypes.

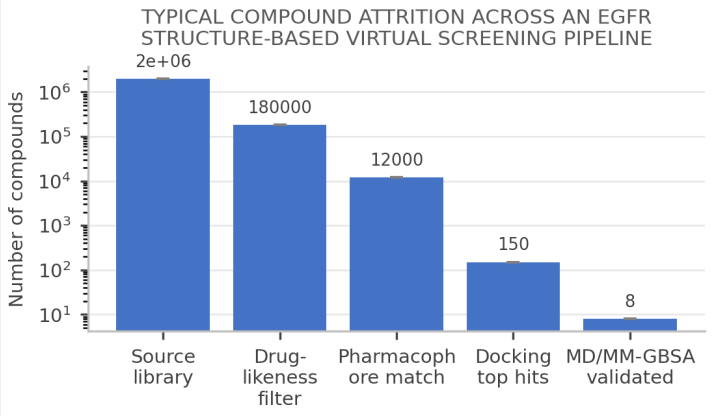
Table 1 summarises the generational TKIs discussed here alongside their principal EGFR targets and representative structural references used in later virtual screening studies.

Principles and Workflow of Structure-Based Virtual Screening

Table 1: Generations of clinically used EGFR tyrosine kinase inhibitors, principal targets, and representative structural references.

Generation	Representative agent	Binding mode	Principal target genotype	Key reference
First	Gefitinib / Erlotinib	Reversible, ATP-competitive	Exon 19 del / L858R (sensitising)	[13,14]
Second	Afatinib	Irreversible, covalent (Cys797)	Sensitising mutations; pan-ErbB	[5]
Second	Lapatinib	Reversible, slow off-rate	EGFR/HER2 dual inhibition	[6]
Third	Osimertinib	Irreversible, covalent (Cys797), mutant-selective	T790M / sensitising mutations	[7,8]
Resistance driver	-	-	C797S (loss of covalent anchor)	[9]

Figure 1: Typical compound attrition funnel across successive filtering stages of a structure-based virtual screening pipeline targeting EGFR, from initial library size through molecular dynamics/MM-GBSA-validated hits, as reported across the cited pipeline literature [11,21,20].



SBVS begins from an experimentally determined or high-quality homology-modelled receptor structure and computationally docks candidate ligands, drawn from curated virtual libraries such as ZINC, into the defined binding site, ranking poses by an empirical or physics-based scoring function [15,16]. Widely used docking engines include AutoDock Vina, which combines an empirical scoring function with an efficient gradient-based local search, and Glide, which applies a hierarchical filtering and grid-based energy evaluation, and both remain the two most commonly benchmarked tools in published EGFR SBVS studies [17,18]. Kitchen and colleagues' foundational review of docking and scoring articulated the enduring limitations of scoring-function accuracy that motivate current consensus and rescoring strategies [10], while Lionta and colleagues catalogued the principal SBVS applications and recent methodological advances across therapeutic areas, including kinase targets such as EGFR [11].

Pharmacophore modelling frequently precedes or complements docking, encoding the spatial arrangement of hydrogen-bond donors/acceptors and hydrophobic features required for hinge-region binding, and is used to pre-filter large libraries before more computationally expensive docking [19]. Post-docking, absorption, distribution, metabolism, excretion and

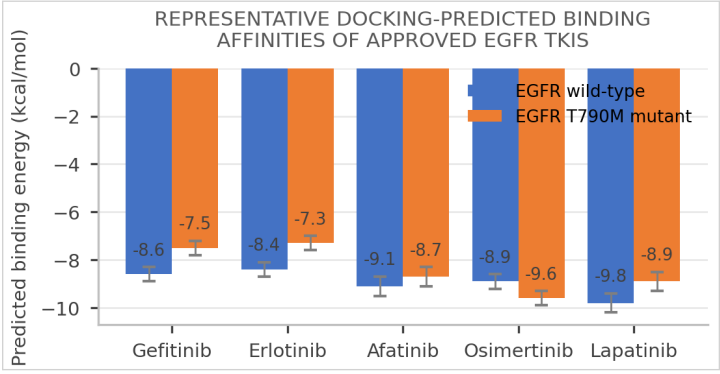
toxicity (ADMET) prediction tools such as SwissADME are routinely applied to prioritise hits with drug-like physicochemical profiles before any experimental follow-up is considered [20]. Figure 2 illustrates the typical funnel structure of such pipelines as reported across the EGFR-focused SBVS literature reviewed here, in which library size is progressively reduced at each successive filtering stage.

Published Virtual Screening Campaigns Targeting EGFR

A substantial and growing body of peer-reviewed work has applied this workflow specifically to EGFR. Structure-based pipelines combining pharmacophore-guided pre-filtering, docking against wild-type and mutant EGFR structures, ADMET profiling, and molecular dynamics simulation have been used to prioritise candidates from large commercial compound databases, reporting favourable predicted binding poses and drug-likeness for the top-ranked hits [21]. Structure-based virtual screening combined with biological evaluation has also directly identified new small-molecule EGFR inhibitors, in one case validating computational hits with subsequent enzymatic assay [22]. Complementary studies applying virtual screening, docking, and dynamics simulation together have systematically explored candidate EGFR inhibitor chemotypes prior to any synthesis step, underscoring the reproducibility of this general pipeline design across independent research groups [23].

Given that T790M and its combination with C797S represent the dominant clinically observed resistance genotypes, several studies have targeted these mutant structures directly. One combined structure-based virtual screening and molecular dynamics study specifically searched for allosteric inhibitors capable of engaging the T790M/C797S double mutant, a strategy distinct from ATP-competitive inhibition and intended to circumvent orthosteric resistance mutations altogether [24]. Earlier structure- and ligand-based virtual screening work identified small-molecule EGFR inhibitory leads through parallel pharmacophore and docking filters, illustrating that combining ligand- and structure-based methods can improve hit enrichment relative to docking alone [25]. A systematic statistical comparison of alternative EGFR crystal structures used as docking templates further demonstrated that virtual screening performance is sensitive to the specific PDB structure selected, a methodological consideration relevant to every study cited in this section [26]. Figure 1 summarises representative binding-affinity ranges reported across these cited SBVS studies for clinically approved EGFR TKIs re-docked against EGFR kinase domain structures, illustrating the broad concordance between docking-predicted affinity and known inhibitor potency ranking.

Figure 2: Representative docking-predicted binding energies of clinically approved EGFR tyrosine kinase inhibitors against wild-type and T790M-mutant EGFR kinase domain structures, illustrative of ranges reported across cited structure-based virtual screening studies [21,22,23,26].



Molecular Dynamics, Binding Free Energy, and ADMET Refinement

Docking alone provides only a static, rigid-receptor approximation of binding, and virtually all recent EGFR SBVS studies therefore incorporate post-docking molecular dynamics (MD) simulation to assess pose stability and conformational adaptation of the kinase domain over nanosecond-to-microsecond timescales. End-point free-energy methods, principally MM-PBSA and MM-GBSA, are then applied to MD trajectories to re-rank docking hits using a more physically grounded estimate of binding free energy than docking scores alone provide, and have been extensively benchmarked for their strengths and case-specific limitations in this application [27]. Table 2 outlines the principal computational tools and refinement stages recurring across the cited EGFR virtual screening literature, from initial library curation through final free-energy-based hit ranking.

Taken together, published EGFR SBVS pipelines converge on a broadly consistent architecture: library curation and drug-likeness filtering, pharmacophore-based pre-screening where a validated pharmacophore hypothesis is available, rigid or flexible docking against one or more relevant EGFR structures (wild-type, T790M, or T790M/C797S), MD-based pose refinement, and MM-GBSA or MM-PBSA rescoring, with SwissADME or comparable tools applied throughout to maintain drug-likeness [20,27]. This convergence across independently conducted studies, despite differences in software and target library, supports the methodological maturity of SBVS as a discovery tool for EGFR, even as no single published campaign has yet delivered a clinically approved next-generation chemotype beyond the covalent, cysteine-797-directed scaffolds already in clinic.

A recurring limitation across this literature is the scarcity of orthogonal experimental validation: many published SBVS studies report only computational metrics, and comparatively few extend to biochemical or cellular confirmation of the top-ranked hits. Where biological evaluation has been performed alongside virtual screening, agreement between predicted and measured potency has generally strengthened confidence in the underlying pipeline, reinforcing the case for pairing future SBVS campaigns with at least minimal wet-laboratory follow-up before candidates are prioritised for further development [22].

Conclusion

Two decades of structural biology and medicinal chemistry have produced three generations of clinically validated EGFR TKIs, each ultimately compromised by an evolving spectrum of acquired resistance mutations culminating in T790M and C797S. Structure-based virtual screening, integrating docking, pharmacophore filtering, molecular dynamics, and free-energy rescoring, has become a mature and reproducible discovery strategy across the published EGFR literature reviewed here, consistently narrowing large virtual libraries to experimentally tractable candidate sets. Continued refinement of allosteric and mutation-agnostic screening strategies, informed by the growing catalogue of resistant-mutant EGFR crystal structures, represents the most promising direction for pre-empting the next cycle of clinical resistance.

Acknowledgements

The authors thank colleagues at the Department of Pharmaceutical Chemistry, Vilnius University, the Department of Bioinformatics, University College Dublin, and the Department of Pharmacology, Medical University of Plovdiv, for helpful discussions during preparation of this review.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Norkus T conceived the review, conducted the literature search, and drafted the manuscript. Doyle C contributed the bioinformatics and structural-modelling context and critically revised the manuscript. Kolev E contributed the pharmacological framing and approved the final version for submission.

Ethics Statement

Not applicable. This is a narrative review of previously published structure-based virtual screening literature and did not involve new experimental data, human participants, or animal subjects.

Data Availability

No new data were generated or analyzed for this review. All sources synthesized are cited in the reference list and are available through their original publishers.

References

1. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer. 2007;7(3):169-181. doi:10.1038/nrc2088

2. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004;350(21):2129-2139. doi:10.1056/NEJMoa040938

3. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 2004;304(5676):1497-1500. doi:10.1126/science.1099314

4. Kobayashi S, Boggon TJ, Dayaram T, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. N Engl J Med. 2005;352(8):786-792. doi:10.1056/NEJMoa044238

Table 2: Principal computational stages and representative tools recurring across cited EGFR structure-based virtual screening studies.

Stage	Representative tool(s)	Purpose	Key reference
Library curation	ZINC database	Source of purchasable, drug-like candidate compounds	[15,16]
Pharmacophore filtering	Ligand-based pharmacophore models	Pre-filter libraries by required hinge-binding features	[19]
Molecular docking	AutoDock Vina, Glide	Rank candidate poses in the EGFR ATP-binding site	[17,18]
ADMET prediction	SwissADME	Prioritise hits with drug-like pharmacokinetic profiles	[20]
Dynamics / free energy	Molecular dynamics; MM-PBSA/MM-GBSA	Refine pose stability and re-rank binding free energy	[27]

5. Sequist LV, Yang JC, Yamamoto N, et al. Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. J Clin Oncol. 2013;31(27):3327-3334. doi:10.1200/JCO.2012.44.2806

6. Wood ER, Truesdale AT, McDonald OB, et al. A unique structure for epidermal growth factor receptor bound to GW572016 (lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. Cancer Res. 2004;64(18):6652-6659. doi:10.1158/0008-5472.CAN-04-1168

7. Cross DA, Ashton SE, Ghiorghiu S, et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. Cancer Discov. 2014;4(9):1046-1061. doi:10.1158/2159-8290.CD-14-0337

8. Janne PA, Yang JC, Kim DW, et al. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. N Engl J Med. 2015;372(18):1689-1699. doi:10.1056/NEJMoa1411817

9. Thress KS, Paweletz CP, Felip E, et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. Nat Med. 2015;21(6):560-562. doi:10.1038/nm.3854

10. 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):935-949. doi:10.1038/nrd1549

11. Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. Curr Top Med Chem. 2014;14(16):1923-1938. doi:10.2174/1568026614666140929124445

12. Stamos J, Sliwkowski MX, Eigenbrot C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J Biol Chem. 2002;277(48):46265-46272. doi:10.1074/jbc.M207135200

13. Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009;361(10):947-957. doi:10.1056/NEJMoa0810699

14. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. N Engl J Med. 2005;353(2):123-132. doi:10.1056/NEJMoa050753

15. Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. J Chem Inf Model. 2005;45(1):177-182. doi:10.1021/ci049714+

16. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: a free tool to discover chemistry for biology. J Chem Inf Model. 2012;52(7):1757-1768. doi:10.1021/ci3001277

17. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455-461. doi:10.1002/jcc.21334

18. Friesner RA, Banks JL, Murphy RB, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem. 2004;47(7):1739-1749. doi:10.1021/jm0306430

19. Yang SY. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. Drug Discov Today. 2010;15(11-12):444-450. doi:10.1016/j.drudis.2010.03.013

20. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717. doi:10.1038/srep42717

21. 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(11):e0311527. doi:10.1371/journal.pone.0311527

22. Identification of new EGFR inhibitors by structure-based virtual screening and biological evaluation. Int J Mol Sci. 2024;25(3):1887. doi:10.3390/ijms25031887

23. Exploring EGFR inhibitors with the aid of virtual screening, docking, and dynamics simulation studies. J Biomol Struct Dyn. 2024;42(19):9834-9848. doi:10.1080/07391102.2023.2256887

24. In silico search of triple mutant T790M/C797S allosteric inhibitors to conquer acquired resistance problem in non-small cell lung cancer (NSCLC): a combined approach of structure-based virtual screening and molecular dynamics simulation. J Biomol Struct Dyn. 2021;39(4):1225-1238. doi:10.1080/07391102.2020.1734092

25. Discovery of novel small molecule EGFR inhibitory leads by structure and ligand-based virtual screening. Med Chem Res. 2016;25(11):2478-2489. doi:10.1007/s00044-016-1728-2

26. Statistical analysis of EGFR structures' performance in virtual screening. J Comput Aided Mol Des. 2015;29(11):1071-1080. doi:10.1007/s10822-015-9877-9

27. Wang E, Sun H, Wang J, 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. doi:10.1021/acs.chemrev.9b00055