DOI 10.24412/1829-0450-fm-2025-2-82-92
УДК 577 .29(577.322.9)

Поступила: 23.10.2025г.
Сдана на рецензию: 24.10.2025г.
Подписана к печати: 31.10.2025г.

EFFECT OF EGFR CONFORMATIONAL PLASTICITY ON THE ACCURACY OF COVALENT EGFR INHIBITOR DOCKING: AN IMPROVED APPROACH

S. Grabska

L.A. Orbeli Institute of Physiology NAS RA
e-mail: siranuysh_grabska@outlook.com
ORCID: https://orcid.org/0000-0001-9291-3357

ABSTRACT

Selecting the appropriate crystallographic conformation of a target protein is essential for reliable structure-based virtual screening. In this study, conformational variability of the epidermal growth factor receptor (EGFR) was investigated and how it influences docking accuracy of Afatinib, a dual covalent inhibitor of EGFR, and HER2. Using five EGFR crystal structures (4G5J, 4I22, 3W33, 3POZ, and 5U8L), self- and cross-docking analyses were performed in ICM-Pro to assess binding precision, pose stability, and scoring performance. Self-docking of Afatinib into its native complex (4G5J) reproduced the experimental pose with an RMSD of 0.95 Å, validating the protocol. Among alternative conformations, 4I22 yielded the closest structural agreement, demonstrating an RMSD of 1.02 Å and superior docking scores, while other structures showed substantial deviations in key active-site residues, resulting in distorted poses and lower binding scores. These findings reveal that subtle rearrangements of residues surrounding Cys797, Met793, and Lys745 critically affect ligand accommodation and covalent bond formation. The results emphasize that protein conformational selection, particularly from high-resolution structures - profoundly influences docking fidelity. Incorporating multiple conformations through cross- docking enhances predictive robustness and better reflects the dynamic nature of protein–ligand recognition in virtual screening workflows.

Keywords: EGFR, afatinib, covalent inhibitor, cross-docking, protein conformation.

Introduction

The selection of an appropriate crystallographic structure of a protein is a critical step in any structure-based virtual screening project. The accuracy and predictive power of docking results strongly depend on the conformational state of the protein model used. Proteins are inherently dynamic and exist in an ensemble of conformations rather than in a single static structure. This intrinsic conformational plasticity reflects their ability to adopt multiple structural states in response to environmental conditions or ligand binding, profoundly influencing ligand recognition and binding affinity [1].

A single protein can display distinct conformations depending on the nature of its bound molecules, such as small-molecule modulators, peptides, cofactors, and other factors. Each binding event can stabilize a particular structural state, altering the geometry and physicochemical properties of the binding pocket. Consequently, the success of virtual screening is closely tied to how accurately the selected protein structure represents the biologically relevant conformation for ligand binding. Choosing an inappropriate conformation may lead to the exclusion of potentially active compounds or, conversely, to the prioritization of false positives [1].

This issue is particularly relevant for both non-covalent and covalent ligand screening, where the spatial arrangement and flexibility of key residues determine the feasibility of binding. In recent years, the importance of considering protein flexibility, either through ensemble docking, cross-docking, or molecular dynamic simulations-based approaches, has become increasingly recognized. Among these, cross-docking provides a practical means to evaluate the compatibility of ligands across multiple receptor conformations, thereby identifying the most suitable structural models for virtual screening.

In this study, a systematic cross-docking analysis using Afatinib, a dual covalent inhibitor of the epidermal growth factor receptor (EGFR) [2] and human epidermal growth factor receptor 2 (HER2), was performed [3]. As a model system, the co-crystal structure of Afatinib bound to EGFR was selected, along with four additional EGFR crystallographic structures complexed with different ligands. This approach was used to assess how variations in EGFR conformations influence docking accuracy and ligand accommodation. By comparing docking results across multiple receptor conformations, highlighting the impact of structural selection on virtual screening outcomes and providing insights into best practices for receptor preparation in structure-based drug discovery.

Materials and Methods

The crystallographic structure of the epidermal growth factor receptor (EGFR) in complex with Afatinib was obtained from the Protein Data Bank (PDB ID: 4G5J) [4]. This structure represents the covalent complex between Afatinib and EGFR, in which the acrylamide warhead of Afatinib forms an irreversible covalent bond with the Cys797 residue located in the ATP-binding pocket of the kinase domain. This interaction stabilizes the inactive conformation of EGFR and serves as a reference for evaluating ligand-receptor compatibility in subsequent docking studies.

To assess the influence of receptor conformation on docking accuracy, four additional EGFR crystallographic structures bound to different ligands were selected: 4I22 [5], 3W33 [6], 3POZ [7], and 5U8L [8]. The selection criteria were based on (i) high structural resolution, ensuring accurate representation of side-chain orientations and binding-site geometry, and (ii) the inclusion of either covalent or non-covalent ligands, providing a diverse set of conformational states. These structures encompass multiple EGFR conformations, thereby enabling a comparative analysis of ligand binding across distinct receptor conformations.

All computational procedures, including ligand preparation, docking, and visualization, were performed using MolSoft ICM-Pro (version 3.9-4a; Molsoft LLC, La Jolla, CA, USA) [9]. The receptor structures were preprocessed by removing water molecules and heteroatoms not directly involved in ligand binding. Covalent docking was applied to model the interaction between the Afatinib acrylamide group and the sulfur atom of Cys797. Cross-docking experiments were then conducted by docking Afatinib into the selected EGFR conformations to evaluate binding consistency and conformational adaptability.

Visualization and post-docking analyses, including examination of covalent bond geometry, and root-mean-square deviation (RMSD) values between predicted and experimental poses, were also performed in ICM-Pro.

Results

A docking analysis of Afatinib was performed using five crystallographic structures of EGFR, one of which corresponds to the experimentally co-crystallized complex (PDB ID: 4G5J). The docking outcomes for all receptor conformations are summarized in Table 1.

As expected, the self-docking of Afatinib into its native crystal structure (4G5J) reproduced the experimentally observed ligand pose with high precision, yielding RMSD of 0.95 Å between the predicted and co-crystallized conformations (Figure 1). This result confirms that the docking protocol and parameterization in ICM-Pro were well-suited for modeling covalent binding interactions, accurately capturing the geometry of the acrylamide-Cys797 linkage.

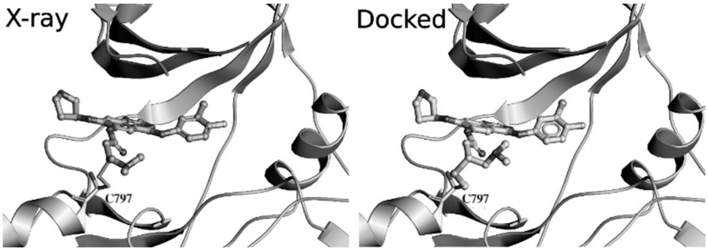

Figure 1. Conformation of co-crystallized and docked Afatinib.

Among the cross-docking experiments, Afatinib displayed the best agreement with the receptor conformation from PDB ID: 4I22, achieving an RMSD of 2.39 Å. Notably, the binding-site residues of 4I22 exhibit a spatial arrangement highly similar to those of the 4G5J structure, particularly in the vicinity of Cys797, Met793, and Lys745. This structural similarity likely accounts for the favorable docking score and pose alignment obtained with 4I22, which in some cases surpassed the score observed in the self-docking experiment. These findings indicate that 4I22 represents a receptor conformation well-suited for modeling Afatinib binding and may serve as a viable alternative for structure-based virtual screening of related inhibitors.

In contrast, the remaining three EGFR structures 3W33, 3POZ, and 5U8L demonstrated substantial conformational deviations in the residues forming the Afatinib binding pocket. These deviations resulted in significant distortions of the predicted poses and markedly lower docking scores. The altered orientation of side chains within the hinge region and the displacement of the catalytic loop residues appear to prevent the formation of the characteristic hydrogen-bonding and covalent interactions observed in the 4G5J complex.

Collectively, these results emphasize that the choice of crystallographic conformation has a decisive impact on docking performance, and that cross-docking analyses can effectively identify receptor models most compatible with the ligand of interest.

Docking performance of Afatinib across different EGFR crystallographic conformations. Summary of docking and cross-docking results for Afatinib using five EGFR crystal structures. The co-crystal structure (PDB ID: 4G5J) served as the reference for self-docking, while four additional structures (4I22, 3W33, 3POZ, and 5U8L) were used for cross-docking experiments. “Ligand interacting atoms RMSD” represents the deviation of key interacting residues relative to the 4G5J conformation, while “Ligand RMSD” indicates the deviation of the predicted ligand pose from the crystallographic reference. Docking performance is evaluated by ICM-Pro Score, RTCNN score, and Average Score, with lower (more negative) values indicating stronger predicted binding affinity and better overall docking quality.

Table 1.

PDB ID	Resolution (Å)	Type of docking	Ligand interacting atoms RMSD (Å)	Ligand RMSD (Å)	Score	RTCNN score	Average Score
4G5J	2.8	Self	0	0.95	-29	-30.72	-29.86
4I22	1.71	Cross	2.39	1.02	-31.14	-44.11	-37.62
3W33	1.7	Cross	2.74	4.29	-9.12	-38.03	-23.57
3POZ	1.5	Cross	2.56	4.37	-14.21	-31.28	-22.75
5U8L	1.6	Cross	2.83	4.21	-8.72	-23.24	-15.98

A comparative analysis of the co-crystallized Afatinib and its cross-docked conformations reveals that the orientation of the 3-chloro-4-fluoroanilino group is influenced by the conformation of residue M766 (Figure 2).

The amino acids R841 and D800 play a crucial role in stabilizing the 4-dimethylamino-trans-but-2-enamide tail of Afatinib, which contains the Michael acceptor warhead responsible for forming a covalent bond with C797. Their proper positioning ensures optimal alignment of the reactive site and facilitates effective covalent inhibition of the EGFR kinase domain.

Furthermore, the conformation of the (S)-tetrahydrofuran-3-yloxy group of Afatinib is largely influenced by the spatial orientation and flexibility of residue K728. Variations in K728 orientation can lead to significant differences in ligand fitting and overall docking accuracy across EGFR conformations.

It was observed that the conformations of these key residues in the co-crystallized Afatinib–EGFR complex (PDB ID: 4G5J) closely resemble those in the EGFR structure 4I22, which likely explains why 4I22 yielded the most accurate docking pose and best scoring performance. This structural similarity indicates that 4I22 provides a biologically relevant receptor conformation for reproducing experimentally observed binding modes.

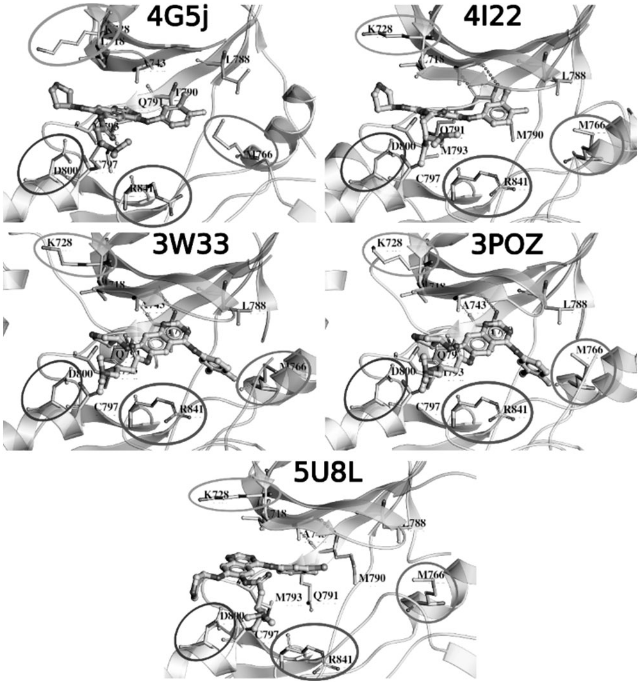

Figure 2. Co-crystallized afatinib (PDB ID: 4G5J) and predicted afatinib conformations docked to various crystallographic EGFR structures (4I22, 3W33, 3POZ, 5U8L). Amino acid residues exhibiting conformational differences among the five EGFR crystal structures are highlighted with circles.

In contrast, the corresponding residues in the other EGFR structures (3W33, 3POZ, and 5U8L) exhibit pronounced conformational variability, particularly in flexible loop regions surrounding the binding pocket. These deviations disrupt optimal ligand orientation, leading to reduced scoring accuracy and less reliable docking predictions across different EGFR conformations.

Thus, the predicted ligand conformation strongly depends on the conformational state of the amino acid residues surrounding the binding site, emphasizing the importance of local structural context. Accurate docking results require careful consideration of receptor flexibility and selection of the most representative protein conformation for computational modeling.

Conclusions

This study highlights the critical importance of selecting an appropriate crystallographic conformation of a target protein for structure-based virtual screening. Using Afatinib, a dual covalent inhibitor of EGFR and HER2, as a model compound, its docking performance was systematically compared across five EGFR crystal structures. The results clearly demonstrate that variations in receptor conformation strongly influence the predicted binding pose, docking score, and overall quality of ligand–receptor interactions.

The self-docking of Afatinib into its native complex (PDB ID: 4G5J) reproduced the experimental pose with sub-angstrom accuracy (RMSD 0.95 Å), validating the reliability of the docking protocol. Among the cross-docking experiments, structure 4I22 exhibited the closest alignment of binding-site residues to the 4G5J conformation, resulting in the most accurate and energetically favorable docking of Afatinib. Conversely, the remaining structures (3W33, 3POZ, 5U8L) displayed significant deviations in the binding pocket, correlating with poorer docking scores and distorted ligand poses. Here are the recommendations for conducting Virtual Screening:

1. It is advisable to use protein conformations obtained from high-resolution crystallographic structures, as they provide more accurate atomic positions and reliable hydrogen-bonding geometries. High-resolution data minimizes coordinate uncertainty, improves docking precision, and reduces false-positive predictions by ensuring that the binding site is modeled with maximal structural fidelity and minimal distortion.
2. For virtual screening, it is recommended to employ multiple diverse conformations of the target protein simultaneously, representing different biologically relevant states. This ensemble docking strategy increases the likelihood of identifying true binders, captures receptor flexibility, and improves hit diversity by accounting for dynamic changes in the active site that may influence ligand accommodation.

3. When selecting crystallographic structures for docking, careful attention should be paid to the specific ligand co-crystallized with the protein. Ligands can induce local conformational changes that significantly affect binding pocket geometry; therefore, choosing structures complexed with chemically similar ligands enhances predictive accuracy and biological relevance of virtual screening outcomes.

Collectively, these findings confirm that protein conformational plasticity plays a decisive role in determining docking outcomes. Even subtle rearrangements within the ATP-binding site can markedly alter ligand accommodation and scoring performance. Therefore, careful structural selection or the use of ensemble and cross-docking strategies are essential to improve the robustness and predictive power of virtual screening workflows.

Acknowledgements. *The author is thankful to Prof. Ruben Abagyan for fruitful and helpful discussion and access to computational infrastructure and for providing access to the MolSoft ICM-Pro software package, which was instrumental in conducting this study. The author is also thankful to Hovakim Grabski for valuable feedback. The research was supported by the Science Committee of MESCS RA, in the frames of the research projects № 25FAST-1F002.*

REFERENCES

1. *Waszkowycz B., Clark D., Gancia E.* Outstanding challenges in protein–ligand docking and structure-based virtual screening // WIREs Comput Mol Sci. 2011. Vol. 1, № 2. PP. 229–259.
2. *Uribe M., Marrocco I., Yarden Y.* EGFR in Cancer: Signaling Mechanisms, Drugs, and Acquired Resistance // Cancers. 2021. Vol. 13, № 11. P. 2748.
3. *Cheng X.* A Comprehensive Review of HER2 in Cancer Biology and Therapeutics // Genes. 2024. Vol. 15, № 7. P. 903.
4. *Solca F. et al.* Target Binding Properties and Cellular Activity of Afatinib (BIBW 2992), an Irreversible ErbB Family Blocker // The Journal of Pharmacology and Experimental Therapeutics. 2012. Vol. 343, № 2. PP. 342–350.
5. *Gajiwala K.S. et al.* Insights into the Aberrant Activity of Mutant EGFR Kinase Domain and Drug Recognition // Structure. 2013. Vol. 21, № 2. PP. 209–219.

6. *Kawakita Y. et al.* Design and synthesis of novel pyrimido[4,5-b]azepine derivatives as HER2/EGFR dual inhibitors // Bioorganic & Medicinal Chemistry. 2013. Vol. 21, № 8. PP. 2250–2261.
7. *Aertgeerts K. et al.* Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein // J Biol Chem. 2011. Vol. 286, № 21. PP. 18756–18765.
8. *Zhao Q. et al.* Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes // J Am Chem Soc. 2017. Vol. 139, № 2. PP. 680–685.
9. *Abagyan R., Totrov M., Kuznetsov D.* ICM-A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation // J. Comput. Chem. 1994. Vol. 15, № 5. PP. 488–506.

ВЛИЯНИЕ КОНФОРМАЦИОННОЙ ПЛАСТИЧНОСТИ EGFR НА ТОЧНОСТЬ ДОКИНГА КОВАЛЕНТНОГО ЛЕКАРСТВА EGFR: УСОВЕРШЕНСТВОВАННЫЙ ПОДХОД

С.В. Грабска

Институт Физиологии им. Л.А.Орбели НАН РА

АННОТАЦИЯ

Выбор подходящей кристаллографической конформации белка имеет решающее значение для надёжности структурно-основанного виртуального скрининга. В данном исследовании изучалось, как конформационная изменчивость рецептора эпидермального фактора роста (EGFR) влияет на точность докинга Афатиниба, двойного ковалентного ингибитора EGFR и HER2. Используя пять кристаллических структур EGFR (4G5J, 4I22, 3W33, 3POZ и 5U8L), был проведен анализ само- и кросс-докинга с помощью программы ICM-Pro для оценки точности связывания, стабильности поз и эффективности оценочных функций. Само-докинг Афатиниба в его нативный комплекс (4G5J) воспроизвёл экспериментальную позу с RMSD 0,95 Å, что подтвердило корректность протокола. Среди альтернативных конформаций структура 4I22 показала наибольшее структурное соответствие, продемонстрировав RMSD 1,02 Å и лучшие значения оценочных функций, тогда как другие структуры имели существенные отклонения в ключевых аминокислотных остатках активного центра, что приводило к искажённым позам и снижению оценки связывания. Эти результаты показывают, что незначительные отклонения аминокислот Cys797, Met793 и Lys745, критически влияют на конформацию лиганда и формирование ковалентной связи. Полученные данные подчёркивают, что выбор конформации белка, особенно при использовании структур с высоким разрешением, существенно влияет на точность докинга. Включение нескольких конформаций при кросс-докинге повышает предсказательную надёжность и лучше отражает динамическую природу взаимодействия белок–лиганд в процессах виртуального скрининга.

Ключевые слова: EGFR, Афатиниб, ковалентный ингибитор, кросс-докинг, конформация белка.