

Mechanism of Drug Efficacy Within the EGF Receptor Revealed by Microsecond Molecular Dynamics Simulation

Shunzhou Wan, David W. Wright, and Peter V. Coveney

Abstract

The EGF receptor (EGFR) regulates important cellular processes including proliferation, differentiation, and apoptosis. EGFR is frequently overexpressed in a range of cancers and is associated with disease progression and treatment. Clinical studies have shown that EGFR mutations confer tumor sensitivity to tyrosine kinase inhibitors in patients with non-small cell lung cancer. In this study, we have conducted molecular dynamics simulations over several microseconds for wild-type and L858R mutant forms of EGFR in the ligand-free state. Close inspection of the conformations and interactions within the binding pocket reveals, converse to the wild type, that the mutant EGFR prefers to bind gefitinib, a targeted anticancer drug, rather than ATP, offering an explanation for why gefitinib is more effective in patients with EGFR mutations than those without. *Mol Cancer Ther*; 11(11); 2394–400. ©2012 AACR.

Introduction

The EGF receptors (EGFR) are key regulators of cellular function, linking extracellular signaling to the cell-cycle machinery. Many carcinomas (1), including lung (2) and breast (3) cancers, exhibit elevated levels of EGFR and related growth factor receptors. Activation of the intrinsic EGFR kinase results in the recruitment and phosphorylation of several intracellular substrates and transduces signals through several pathways mediated by specific adaptor molecules (4). The signaling cascade regulates cell proliferation, survival, and migration.

Mutations in the tyrosine kinase domain of EGFR have been identified in several human cancers (1, 2). In lung cancer, the predominant mutations are a small in-frame deletion $\Delta E746-A750$ at the N-terminal end of the αC -helix, and a point substitution L858R at the start of the activation loop (A-loop; ref. 2; see Fig. 1A). They account for around 90% of all EGFR mutations in non-small cell lung cancer (NSCLC; ref. 2). The numbering system hereafter includes the 24-residue signal peptide of EGFR. An alternative sequence numbering system that excludes the 24 amino acid peptide is also frequently used in the literature, in which L858R is identified as L834R, and $\Delta E746-A750$ as $\Delta E722-A726$. These muta-

tions increase the activities of EGFR kinase, consequently altering the downstream intracellular signaling cascades. Most EGFR kinase domain mutations are clustered around the ATP-binding pocket of the enzyme (Fig. 1A). They are expected to distort the binding cleft and hence the specificity and affinity of ATP (Fig. 1B) as well as drug binding. Clinical data have shown that gefitinib treatment (Fig. 1C) is significantly more beneficial for patients with NSCLC who harbor these EGFR mutations compared with those who do not (2). Through the targeting of EGFR mutation positive patients only, one can think of gefitinib as something akin to a personalized medicine in NSCLC treatment.

Knowledge of the extensive protein kinase family, whose genes constitute about 2% of the entire human genome, has been advanced significantly, thanks to the rapid development of modern analysis techniques (5). With the aid of multiple sequence alignments, it is possible to study sequence patterns conserved within the protein kinase superfamily and to identify sequence motifs that have biologic significance. There are other functional motifs whose presence is detected from the protein structure rather than the sequence: a combination of secondary structures called the conventional structural motif (6), and an ensemble of noncontiguous hydrophobic residues referred to as the "unconventional structural motif" (7). The unconventional structural motifs usually shape into spines or cores in the interior of a protein, which stabilize one of its specific conformational states (7–9). Highly conserved hydrophobic spines and cores have been revealed within protein kinases, and are considered important for defining the catalytic state (7). Extension of a hydrophobic spine by mutation of an adjacent residue can further strengthen the conformation selection (10), whereas disruption of the hydrophobic spine assembly by mutagenesis destabilizes the conformation (8).

Authors' Affiliation: Centre for Computational Science, Department of Chemistry, University College London, London, United Kingdom

Note: Supplementary data for this article are available at Molecular Cancer Therapeutics Online (<http://mct.aacrjournals.org/>).

Corresponding Author: Peter V. Coveney, Centre for Computational Science, Department of Chemistry, University College London, WC1H 0AJ, United Kingdom. Phone: 44-20-7679-4850; Fax: 44-20-7679-1501. E-mail: p.v.coveney@ucl.ac.uk

doi: 10.1158/1535-7163.MCT-12-0644-T

©2012 American Association for Cancer Research.

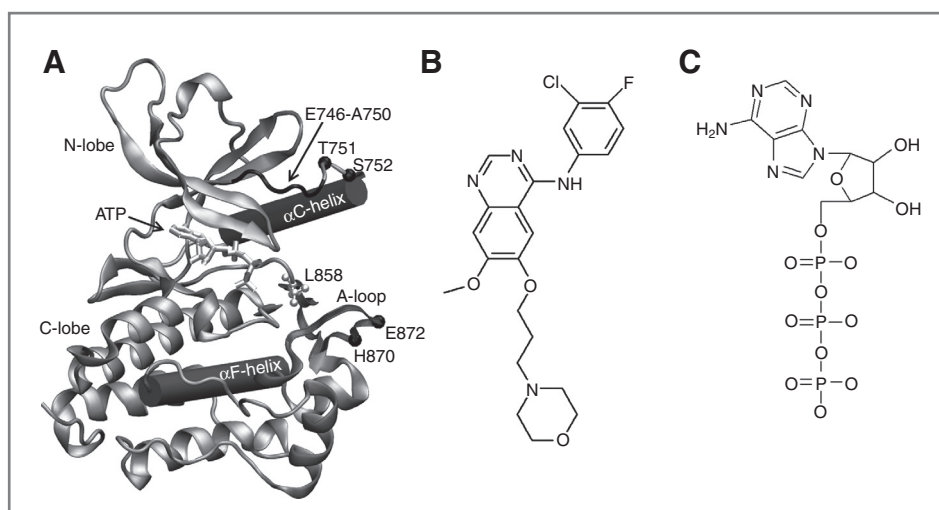


Figure 1. Structures of EGFR and the ligands studied in this paper. A, the intracellular tyrosine kinase domain of the EGFR in complex with ATP [protein data bank (PDB) id: 2GS6]. EGFR is depicted in cartoon and ATP in bond representation. The principal elements used in this study are labeled: the activation loop (A-loop, dark gray ribbon), the α C- and α F-helices (cylinders), the most common mutations Δ E746-A750 (black thin tube), and L858R (ball-and-stick), and the positions of 4 residues (T751, S752, H870, and E872, black balls) closely involved in this study. The chemical structures of gefitinib and ATP are shown in B and C, respectively.

Protein kinases present a highly conserved catalytic domain in which the ATP-binding cleft is located between the N-lobe and C-lobe (Fig. 1A). As a consequence, a number of strategies have been pursued in designing ATP-competitive agents for multiple targets (11). Subtle but significant differences in the binding pocket, however, make it possible to design kinase inhibitors with great selectivity (12). Proteins exist in different conformational substates that endow them with intrinsic degrees of flexibility. The conformational plasticity of the binding site allows multiple diverse ligands to bind at a single site (13), with differing binding affinities (14). Proteins are able to adjust the populations of these substates as they undergo such various binding events. Ligand binding can be viewed as producing a shift in preexisting dynamic equilibria between conformational substates rather than inducing a new conformation (15). Proteins must possess structural flexibility, both globally (via interdomain motions) and locally (through binding site rearrangements), to allow substrate access and product exit.

The existence of conformational substates was first showed directly using molecular dynamics simulations 24 years ago, on a timescale of 300 ps (16). With the advent of increasingly powerful software and hardware, it is now possible to investigate the preexisting equilibrium of multiple conformations on a much longer, microsecond timescale (17). In a recent article (18), we described the flexibility of EGFR in the active and inactive conformations and the effects of L858R mutation on the distribution of conformational substates. Our study was based on 4 replicas of 200 ns molecular dynamics simulations for each molecular system (18). Access to Anton (19) has allowed us to use simulations of 10 μ s to explore a much larger area of conformational space.

Materials and Methods

Molecular dynamics simulation

The simulations reported here were initiated from the final snapshots of 2 ns equilibration runs of our previous study (18). The system configurations were converted from Amber into Mae format using Maestro, and the OPLS all-atom force field was applied (20). Van der Waals and short-range electrostatic interactions were cut off at 9 Å; long-range electrostatics were computed using a *k*-space Gaussian split Ewald method (21), with a $64 \times 64 \times 64$ grid. An additional 10 ns equilibration was conducted using Desmond (22). Fully atomistic, explicit solvent, molecular dynamics simulations were run for both wild-type and the L858R mutant on Anton (19) at Pittsburgh Supercomputing Center generating 10 μ s of trajectory for each system. All simulations were conducted in the isothermal-isobaric (NPT) ensemble with the pressure held at 1 bar and the temperature at 300 K. As a special purpose massively parallel supercomputer (19), Anton is able to produce approximately 5.6 μ s/day trajectories for the molecular systems with approximately 50,000 atoms using all of its 512 nodes.

Analyses

To assess the closeness of the conformations observed in our simulations to the active and inactive forms of EGFR, principal component analysis was conducted on the concatenated trajectories from current simulations and those from our previous study (18). The analysis of the combined trajectory generated similar principal components to those observed previously (see Fig. 2 and Fig. 6 in our recent publication; ref. 18). The capacity of the active site conformations observed in our simulations to accommodate ATP and the inhibitor gefitinib was

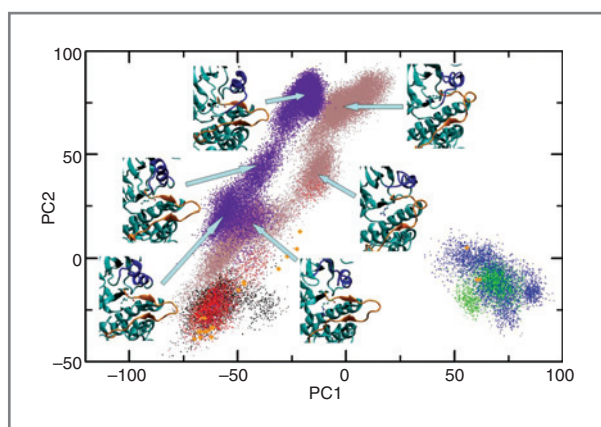


Figure 2. Results of principal component analysis on the EGFR kinase domain. Projections of molecular dynamics trajectories and the X-ray structures onto the principal planes defined by the 2 most significant principal components. Representative structures from the molecular dynamics simulations are displayed around the ATP binding site, with EGFR depicted in cyan, the A-loop in orange, the α C-helix in blue, and L858R in ball-and-stick representation. The results for 10 μ s simulation of wild-type EGFR are shown in indigo and those for L858R mutation in brown. Our previous simulations (18) are included for the definition of the principal component space and are displayed in the figure for reference: black, wild type at the active state; green, wild type at the inactive state; red, L858R at the active state; blue, L858R at the inactive state; and orange, X-ray structures.

assessed by the difference in volume between the protein conformation in each molecular dynamics snapshot and the same conformer with the appropriate ligand fitted into the active site. The capacity was further quantified by docking ATP and gefitinib to each snapshot of the molecular dynamics trajectories using the University of California, San Francisco (San Francisco, CA) Dock 6.5 package (23). The analyses were carried out using the Mavrino Cluster at the Center for Computational Science at University College London (London, United Kingdom). The docking poses were clustered, based on key distances between the hinge residue M793 and the nitrogen atoms in ATP/gefitinib. The numbers of "correct" bound associations were then compared, which indicated the preferences of EGFRs for the binding of ATP or gefitinib. Full details of the simulation protocol and analysis techniques used are provided in the Supplementary Information.

Results

Extensive conformations are sampled from microsecond simulations

Our previous study (18) shows that the free energy landscape is rugged and populated by an ensemble of conformational isomers. In this study, long timescale simulations furnish greatly extended sampling of the landscape and explore a range of new conformational substates, indicative of the flexibility of EGFR on microsecond timescales. The ensemble of conformations produced represents a repertoire of molecular states available for the binding of either ATP or inhibitory drugs such as gefitinib.

The principal component analysis of our current and previous work (18) identifies 2 main principal components that represent the dominant motions of EGFR (Fig. 2). The first principal component distinguishes between the active and inactive conformations of EGFR, whereas the second represents the largest movement within the active or inactive state. Both these modes involve relative motion between the C- and N-lobe, and hence affect the shape and volume of the binding cleft (Supplementary Fig. S1). The wild-type and mutant EGFRs explore similar space along the second PC, whereas the L858R mutation shifts the conformations toward states with more positive eigenvalues on the first PC (PC1 in Fig. 2). This motion is associated with the loss of secondary structure in the α C-helix and conformational changes in the A-loop (Fig. 2). The conformation observed is similar to that seen in crystal structures of the inactive protein kinase B (24). Nonetheless, both systems remain closer to the active rather than the inactive conformation, indicating that no transition occurs during the simulation and the conformations explored remain candidates for ligand binding. The conformational space sampled is, however, significantly extended from our previous 200 ns simulations (18). Several stable intermediate states are reached in these simulations, after about 3 μ s for wild-type and 5 μ s for L858R mutant EGFR (Supplementary Fig. S1). It is interesting to note that the substate of the L858R mutation, around (coordinates $-30, 37$) on the first 2 principal components, is only reached at about 3.4 μ s; the same conformation was observed in one of our previous four 200 ns simulations (18), after only 100 ns simulation. Although multiple short (ensemble) simulations have been shown to explore local conformational space more effectively (14, 25), long timescale simulations are needed for large-scale conformational changes such as ligand binding (26) and protein folding (27).

Wild-type and mutant EGFRs have different preferences for ATP and ligand

It is well understood that a binding event involves a repertoire of selection and adjustment processes (28). Proteins prefer to bind a ligand whose shape and properties are highly complementary to the binding site. To quantify EGFR's preference for binding ATP or gefitinib, the capacity is defined as the percentage volume of a ligand that does not overlap with the receptor. The capacity of the binding cleft was estimated for all EGFR conformations sampled in the simulation for ATP/ligand binding. ATP and gefitinib from X-ray crystallography were fitted into the binding cleft, after aligning the residues around the cleft. As apo EGFR was used in the simulations, the binding cleft changed its shape and volume, shrinking slightly during the course of the simulations in the absence of ATP/ligand. As a result, both ATP and ligand have some degree of overlap with the protein. Nonoverlapped percentages of the ATP/ligand volumes are shown in Fig. 3, as well as the relative binding capacity between ATP and ligand displayed as the ratio of these percentages.

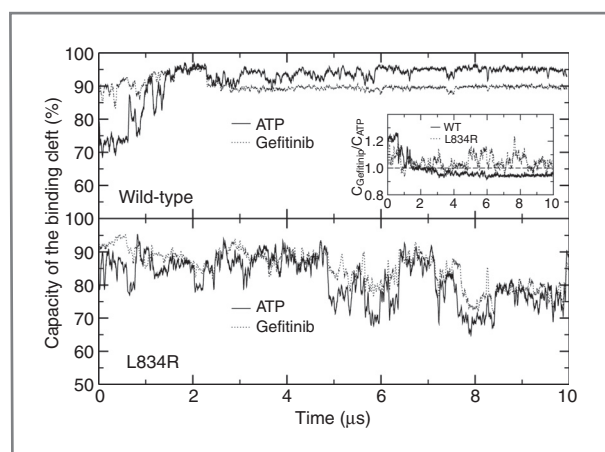


Figure 3. Capacity of the protein cleft for ATP/ligand binding. ATP and gefitinib taken from the X-ray structures are put into the binding cleft of EGFR. The percentages of nonoverlapped volumes are compared for ATP and gefitinib. One hundred percent means ATP/ gefitinib can be allocated without any strain, whereas 0 means there is no space for them at all. The inset figure shows the relative binding preference for ATP or gefitinib, which is inverted between wild-type and mutant EGFRs.

Binding clefts are inherent sites of conformational flexibility. The variations of the groove spaces for ATP/ gefitinib indicate the existence of an ensemble of substates, which can be related to the binding of multiple ligands with different sizes and shapes (13). The cleft must open to allow access of ATP/ligands to the hydrophobic pockets within the protein's interior. The dynamic "breathing" of the binding cavity in L858R EGFR may well assist the opening of the ligand binding cleft, and hence affect the binding/releasing of ATP and inhibitors. The greater flexibility of the L858R mutant EGFR implies that it possesses a larger ensemble of conformations than the wild type. Although a more rigid protein usually indicates more selective binding, an increased number of isoforms could bind a larger range of ligands whose structures are complementary to the ensemble. The less rigid mutant EGFR has more preexisting substates that diverse ligands can target (29). This may be the reason why the only 2 available active-state-targeted drugs, gefitinib and erlotinib (30), are both more effective in patients with EGFR mutations than those with wild type (31).

The extent to which the binding cleft is capable of binding ATP and gefitinib determines the preference for ligand binding and thus EGFR function. The preference is estimated here by the ratio of the relative available volumes for gefitinib and ATP. After reaching the most stable conformations, the wild-type EGFR consistently has larger relative capacity for ATP than that for gefitinib, which indicates its preference for the natural substrate over the inhibitor. In contrast with the wild-type, the L858R variant prefers gefitinib binding, as indicated by the larger relative space available in the groove. These observations are consistent with clinical studies that show

that gefitinib has a greater response among patients with mutations (32).

The capacity of EGFR for ATP/ gefitinib binding was further quantified by docking studies. Preferred position and orientation of flexible ATP and gefitinib were predicted for each snapshot of the trajectories. Cluster analyses were conducted to find the most common bound positions of ATP/ gefitinib from the large quantities of docked conformations. ATP and gefitinib each had one cluster that resembled their X-ray poses in the binding site (Supplementary Fig. S2). We designate EGFR conformations in these 2 clusters as being capable of binding the relevant ligand. The frequencies of observing binding competent conformations for ATP and gefitinib were then compared for the wild-type and mutant EGFRs (Table 1). Comparable numbers of conformations capable of binding ATP and gefitinib were observed for the wild-type EGFR. Although the L858R mutation does not change EGFR's binding capacity for gefitinib, it significantly decreases the preference for ATP. The docking result for ATP is in agreement with the *in vitro* kinase activity experiment (10) that reveals a marked increase in the Michaelis constant (K_m) for ATP in the L858R mutant as compared with the wild-type EGFR. The gefitinib result is also consistent with the binding assay experiment (33) and our previous simulation (14), in which gefitinib shows comparable binding affinities to both the wild-type and L858R mutant EGFRs. Consequently, the L858R-mutated EGFR prefers to bind gefitinib rather than ATP, which explains the efficacy of gefitinib for patients who are L858R mutation positive. The preference could be even more significant if the L858R mutant increases the binding affinity for gefitinib, as indicated in another inhibition assay experiment (10). It should be noted that, although gefitinib does induce dramatic responses in cancer patients with L858R mutation, the occurrence of a relapse is usually detected after an initial gefitinib treatment because of the development of resistance mutations in EGFR (34) and/or other protein components of the ErbB signaling network (35). A multiscale approach would be needed to address the complex nature of cancer and to open up new options for cancer treatment (36).

Table 1. Percentages of conformations capable of binding ATP and gefitinib

Variant EGFR	ATP	Gefitinib
WT	17.2%	16.7%
L858R	9.5%	15.5%

NOTE: Conformations were sampled from $2 \times 10 \mu\text{s}$ molecular dynamics simulations for wild-type and L858R mutant EGFRs.

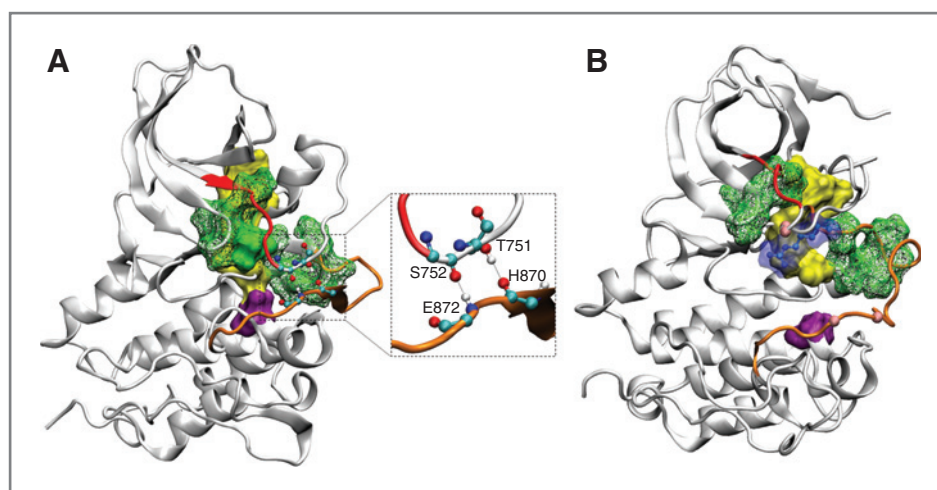


Figure 4. A, the R-spine (yellow) and N-spine (green) are well preserved in the simulations of wild-type EGFR. The R-spine is anchored by an aspartic acid (D896, purple) into the α F-helix and traverses both lobes. The N-spine links one end of the β -sheet to the N-terminal end of the A-loop, of which the L858 residue (green, solid surface) sits in the middle. The deletion mutation Δ E746–A750 is displayed in cartoon representation (red). The inset figure shows the 2 hydrogen bonds formed within the A-loop (orange) and the N-terminal end of the α C-helix, between T751 and H870, and between S752 and E872 (colored by atom type: carbon, cyan; oxygen, red; nitrogen, blue, and hydrogen, white). The spines and hydrogen bonds stabilize the conformations of wild-type EGFR. B, L858R mutation introduces a charged arginine residue (ball-and-stick, with transparent surface), which disturbs the N-spine (green) and detaches the R-spine (yellow) from the α F-helix (anchor residue D896, purple). The positions of the residues T751, H870, S752, and E872 are displayed in pink. The lack of hydrophobic and electrostatic interactions confers dynamic “breathing” on the conformation.

Hydrophobic spines stabilize the conformations of wild-type EGFR

Residues in a hydrophobic spine form a conserved spatial structure that is referred to as the unconventional structural motif (7). A regulatory spine (R-spine) is observed in all eukaryotic protein kinases (8) in which hydrophobic residues assemble together to connect the N- and C-lobes. The R-spine is well preserved in the wild-type simulation, and is anchored to the α F-helix (Fig. 1A) through a conserved aspartate residue (D896; Fig. 4A).

There is a second hydrophobic spine found in the wild-type simulation, referred to as an N-lobe spine (N-spine) here as it is built from residues in the N-lobe. It serves to stabilize the p-loop, the N-terminal end of the α C-helix, and the N-terminal region of the A-loop. The L858 residue sits in the middle of this spine (Fig. 4A). The N-spine forms the roof of the binding cleft and stabilizes its conformation. The A-loop residues H870 and E872 form hydrogen bonds with residues T751 and S752 at the N-terminal end of the α C-helix, which further stabilize these conformations (Supplementary Fig. S3).

L858R mutation disturbs the hydrophobic spines

From a structural standpoint, comparison of the wild-type EGFR to its L858R variant counterpart illustrates their variabilities (37). It has been shown that mutation of the spine residues leads to increased flexibility of local and global conformations of proteins (9). Residue L858 lies in the middle of the N-spine and the change from leucine to arginine introduces a bulky, charged residue. The mutation completely destroys the

N-spine, and also detaches the R-spine from the α F-helix (Fig. 4B). The conformational changes also separate the A-loop and the N-terminal of the α C-helix, breaking the hydrogen bonds between them (Fig. 4). The loss of hydrophobic spines and electrostatic interactions makes the L858R mutant more flexible (Supplementary Fig. S3), which is expected to assist the binding and release of some ligands.

It is interesting to note that another predominant mutant, the deletion Δ E746–A750 (2), is located adjacent to T751 and S752 (Fig. 4A). This deletion mutation leads to the shortening of a segment that adjoins the α C-helix and is likely to interrupt the electrostatic interactions between the N-lobe and the α C-helix (Fig. 4A). In the region spanning from E746 to I759, some deletion mutations (38) involve the residues T751 and S752 and hence directly remove the electrostatic interactions between N- and C-lobes, as shown in Fig. 4. The lack of hydrogen bonds may well increase the flexibility of the binding site and affect ATP/ligand binding.

Conclusions

Protein structure and flexibility are key to understanding ligand binding, upon which drug design paradigms are built. Although mutations rarely change global structures significantly, they can cause important variations within the local structure and flexibility around functional sites. Although L858R variant EGFR has a very similar X-ray structure to the wild-type, it confers substantially enhanced clinical responses to gefitinib in patients with NSCLC. As gefitinib is an ATP

competition tyrosine kinase inhibitor, the favorable response indicates L858R mutant EGFR preferentially binds the inhibitor over ATP.

In the absence of any ligands, the EGFR binding cleft is dynamic, shifting among various substates. Comparison of the simulations for the wild-type and L858R mutant EGFRs provides an explanation for the mechanism of gefitinib efficacy. The L858R mutation provides a series of conformational states that are more capable of binding gefitinib than ATP. The L858R mutation disrupts the hydrophobic spines found in the wild-type, enabling larger interlobe movement. The resultant changes in the shape and volume of the binding cleft alter the accessibility of the cleft for different ligands, favoring the binding of gefitinib over ATP. The drug efficacy of the $\Delta E746-A750$ mutation may also be related to interlobe mobility through disruption of the hydrogen bonds between 2 lobes.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

- Nicholson RI, Gee JMW, Harper ME. EGFR and cancer prognosis. *Eur J Cancer* 2001;37:S9–15.
- Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 2007; 7:169–81.
- Sainsbury JRC, Farndon JR, Needham GK, Malcolm AJ, Harris AL. Epidermal-growth-factor receptor status as predictor of early recurrence of and death from breast-cancer. *Lancet* 1987;1:1398–402.
- Yarden Y. The EGFR family and its ligands in human cancer: signalling mechanisms and therapeutic opportunities. *Eur J Cancer* 2001;37: S3–S8.
- Johnson SA, Hunter T. Kinomics: methods for deciphering the kinome. *Nat Methods* 2005;2:17–25.
- Branden C, Tooze J. Introduction to protein structure. New York: Garland; 1999.
- Kornev AP, Taylor SS. Defining the conserved internal architecture of a protein kinase. *Biochim Biophys Acta* 2010;1804:440–4.
- Garriz A, Qiu HF, Dey M, Seo EJ, Dever TE, Hinnebusch AG. A network of hydrophobic residues impeding helix alpha C rotation maintains latency of kinase Gcn2, which phosphorylates the alpha subunit of translation initiation factor 2. *Mol Cell Biol* 2009;29:1592–607.
- Shin AJ, Telesco SE, Choi SH, Lemmon MA, Radhakrishnan R. Molecular dynamics analysis of conserved hydrophobic and hydrophilic bond-interaction networks in ErbB family kinases. *Biochem J* 2011;436:241–51.
- Yun CH, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong KK, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci U S A* 2008;105: 2070–5.
- Morphy R. Selectively nonselective kinase inhibition: striking the right balance. *J Med Chem* 2010;53:1413–37.
- Noble MEM, Endicott JA, Johnson LN. Protein kinase inhibitors: insights into drug design from structure. *Science* 2004;303:1800–5.
- Ma BY, Shatsky M, Wolfson HJ, Nussinov R. Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci* 2002;11:184–97.
- Wan S, Coveney PV. Rapid and accurate ranking of binding affinities of epidermal growth factor receptor sequences with selected lung cancer drugs. *J R Soc Interface* 2011;8:1114–27.
- Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* 2009; 5:789–96.
- Elber R, Karplus M. Multiple conformational states of proteins—a molecular-dynamics analysis of myoglobin. *Science* 1987;235: 318–21.
- Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, et al. Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature* 2011;469:236–40.
- Wan S, Coveney PV. Molecular dynamics simulation reveals structural and thermodynamic features of kinase activation by cancer mutations within the epidermal growth factor receptor. *J Comput Chem* 2011;32:2843–52.
- Shaw DE, Deneroff MM, Dror RO, Kuskin JS, Larson RH, Salmon JK, et al. Anton, a special-purpose machine for molecular dynamics simulation. *CACM* 2008;51:91–7.
- Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 1996;118:11225–36.
- Shan YB, Klepeis JL, Eastwood MP, Dror RO, Shaw DE. Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. *J Chem Phys* 2005;122:054101.
- Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. In: Proceedings of the ACM/IEEE Conference on Supercomputing (SC06), Tampa, Florida; 2006.
- Lang PT, Brozell SR, Mukherjee S, Pettersen EF, Meng EC, Thomas V, et al. DOCK 6: combining techniques to model RNA-small molecule complexes. *RNA* 2009;15:1219–30.
- Yang J, Cron P, Good VM, Thompson V, Hemmings BA, Barford D. Crystal structure of an activated Akt/protein kinase B ternary complex with GSK3-peptide and AMP-PNP. *Nat Struct Biol* 2002;9: 940–4.
- Sadiq SK, Wright DW, Kenway OA, Coveney PV. Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases. *J Chem Inf Model* 2010;50: 890–905.
- Dror RO, Pan AC, Arlow DH, Borhani DW, Maragakis P, Shan YB, et al. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc Natl Acad Sci U S A* 2011;108:13118–23.

Authors' Contributions

Conception and design: S. Wan, D. W. Wright, P. V. Coveney

Development of methodology: S. Wan

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): S. Wan, D. W. Wright

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S. Wan, D. W. Wright

Writing, review, and/or revision of the manuscript: S. Wan, D. W. Wright, P. V. Coveney

Study supervision: P. V. Coveney

Acknowledgments

The authors thank Dr. Benjamin Hall for helpful discussions.

Grant Support

This work was supported by EU FP7 ContraCancrum (ICT-2007.5.30) and p-medicine (ICT-2009.5.3) grants, an NIH grant (NIH RC2GM093307) that provided an allocation (MCB110017) on Anton at the Pittsburgh Supercomputing Center (PSC), and a TeraGrid/XSEDE NSF grant (TG-MCB090174) for access to Ranger at the Texas Advanced Computing Center (TACC).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 25, 2012; revised July 20, 2012; accepted July 27, 2012; published OnlineFirst August 3, 2012.

27. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. *Science* 2011;334:517–20.
28. Csermely P, Palotai R, Nussinov R. Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci* 2010;35:539–46.
29. Carlson HA. Protein flexibility and drug design: how to hit a moving target. *Curr Opin Chem Biol* 2002;6:447–52.
30. Dowell J, Minna JD, Kirkpatrick P. Erlotinib hydrochloride. *Nat Rev Drug Discov* 2005;4:13–4.
31. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, et al. EGFR receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A* 2004;101:13306–11.
32. Zhou Q, Zhang XC, Chen ZH, Yin XL, Yang JJ, Xu CR, et al. Relative abundance of EGFR mutations predicts benefit from gefitinib treatment for advanced non-small-cell lung cancer. *J Clin Oncol* 2011; 29:3316–21.
33. Fabian MA, Biggs WH, Treiber DK, Atteridge CE, Azimioara MD, Benedetti MG, et al. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat Biotechnol* 2005;23: 329–36.
34. Kobayashi S, Boggon TJ, Dayaram T, Janne PA, Kocher O, Meyerson M, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2005;352: 786–92.
35. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2001;2:127–37.
36. Wright DW, Wan S, Shublaq N, Zesada S, Coveney PV. From base pair to bedside: molecular simulation and the translation of genomics to personalized medicine. *WIREs Syst Biol Med* 2012. doi: 10.1002/wsbm.1186. [Epub ahead of print].
37. Yun CH, Boggon TJ, Li YQ, Woo MS, Greulich H, Meyerson M, et al. Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. *Cancer Cell* 2007;11:217–27.
38. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004;304:1497–500.