Effects of oncogenic mutations on the conformational free-energy landscape of EGFR kinase

Ludovico Sutto[a] and Francesco Luigi Gervasio[a,b,c,1]

[a]Institute of Structural and Molecular Biology and [b]Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom; and [c]Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre, 28029 Madrid, Spain

Edited by José N. Onuchic, Rice University, Houston, TX, and approved May 16, 2013 (received for review December 16, 2012)

Activating mutations in the epidermal growth factor receptor (EGFR) tyrosine kinase are frequently found in many cancers. It has been suggested that changes in the equilibrium between its active and inactive conformations are linked to its oncogenic potential. Here, we quantify the effects of some of the most common single (L858R and T790M) and double (T790M-L858R) oncogenic mutations on the conformational free-energy landscape of the EGFR kinase domain by using massive molecular dynamics simulations together with parallel tempering, metadynamics, and one of the best force-fields available. Whereas the wild-type EGFR catalytic domain monomer is mostly found in an inactive conformation, our results show a clear shift toward the active conformation for all of the mutants. The L858R mutation stabilizes the active conformation at the expense of the inactive conformation and rigidifies the αC-helix. The T790M gatekeeper mutant favors activation by stabilizing a hydrophobic cluster. Finally, T790M with L858R shows a significant positive epistasis effect. This combination not only stabilizes the active conformation, but in nontrivial ways changes the free-energy landscape lowering the transition barriers.

conformational changes | resistance-causing mutations

The epidermal growth factor receptor (EGFR) is one of the most studied members of the receptor tyrosine kinases family due to its central role in many cellular pathways and its importance as a drug target (1). Activating mutations of EGFR are common in human malignancies (2), where they are implicated in the cellular survival, proliferation, and migration, and are the oncogenic drivers in glioblastoma and nonsmall-cell lung cancer (NSCLC) (3, 4).

Structurally, EGFR is composed of a cell-surface receptor, a transmembrane region, and an intracellular catalytic domain. The activation of the enzyme is triggered by binding of the EGF ligand on the extracellular domain, which induces dimerization of the receptors (5). During activation, an asymmetric dimer is formed in which the C lobe of the catalytic domain playing the "activator" interacts with the N lobe of the "receiver" domain, promoting and stabilizing the active conformation of the latter (6–8). The mechanism of activation bears some resemblance to that of cyclin-dependent kinase 2 (CDK2), which is activated by cyclin A (9). The catalytic domain (CD) of EGFR assumes a fold that is common to most protein kinases (10). It consists of a smaller N-terminal lobe, containing mostly β-sheets, and an α-helical C-terminal lobe. The ATP binding cleft is at the interface between the lobes. Kinase activation is mainly controlled by conformational changes in three conserved structural motifs close to the active site: the activation loop (A loop), the Asp-Phe-Gly (DFG) motif, and the αC-helix (Fig. 1). The active state conformation is characterized by an extended conformation of the A loop and the presence of a hairpin in its N-terminal region. The αC-helix assumes an "in" conformation, which is kept in place by a salt bridge formed by E762 and K745 (Fig. 1). Crystal structures of EGFR have revealed an inactive conformation similar to the one observed in Src tyrosine kinase ("Src-like" inactive conformation), in which a short α-helix is formed in the A loop, making an autoinhibitory interaction with a shifted αC-helix (6, 11).

Many somatic mutations of the CD of EGFR have clinical importance due to their activating or drug-resistance effects. The L858R mutation (also referred to as L834R; see *Materials and Methods* for residue numbering conventions) is the most common oncogenic mutation in the EGFR kinase domain (12, 13) and one of the most observed kinase mutations in human cancers (2, 14). It is found in a significant subset of patients with NSCLC (4, 15–17). L858R is activating; the mutant shows a 20–50-fold increase in the catalytic efficiency with respect to wild-type (WT) EGFR (6, 18). Two contrasting activation mechanisms have been proposed for this mutation. The first, based on the comparison of WT and mutant crystal structures, posits that the L858R mutation locks the kinase in the active state because it prevents residue 858, and flanking residues, from forming the helical conformation typical of the inactive state (18). The second mechanism, based on observations of microsecond-long molecular dynamics trajectories, proposes that L858R acts by reducing the intrinsic disorder in the αC-helix located in the N lobe of the "receiver" kinase. This in turn would favor dimerization and indirectly stabilize the active conformation (19).

Two L858R-selective, ATP-competitive tyrosine kinase inhibitors (TKIs), erlotinib and gefitinib (16, 20), are in clinical use (18). In the case of gefitinib, its selectivity for the mutant might be the consequence of increased affinity of the drug, which binds to an active-like conformation, combined with decreased affinity of ATP (21, 22). However, the conformational selectivity of EGFR inhibitors has been recently questioned, when erlotinib was found to also bind to the inactive EGFR CD conformation (23).

An additional problem is the emergence of drug-resistance mutations. One of the most frequent resistance mutations is the T790M, in which the so-called "gatekeeper" threonine at position 790 is mutated to methionine (17, 24). This mutation was first thought to reduce the affinity to the drugs due to a sterical clash in the drug binding site (17). However, Yun et al. showed that both the single-mutant T790M and the double-mutant L858R–T790M maintain the same low nanomolar affinity to gefitinib as the L858R mutant does (21). In contrast, the T790M mutation confers a higher affinity to ATP than L858R mutant so that the combined mutation L858R–T790M results in an activated enzyme resistant to ATP-competitive TKIs (25). What is more, there are several lines of evidence showing that L858R and T790M have a positive epistasis effect (14).

Although the biological effects of these important mutations are clear, a mechanistic explanation linking the mutations to changes in the conformational free-energy landscape is still missing. Thanks to the advances in force fields (26) and the use of specialized computer architectures (27) or enhanced sampling methods (28, 29), it is now possible to use all-atom molecular dynamics (MD) simulations to accurately describe the complex conformational transitions involved

Author contributions: F.L.G. designed research. L.S. performed research; L.S. and F.L.G. analyzed data; and L.S. and F.L.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: f.l.gervasio@ucl.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1221953110/-/DCSupplemental.

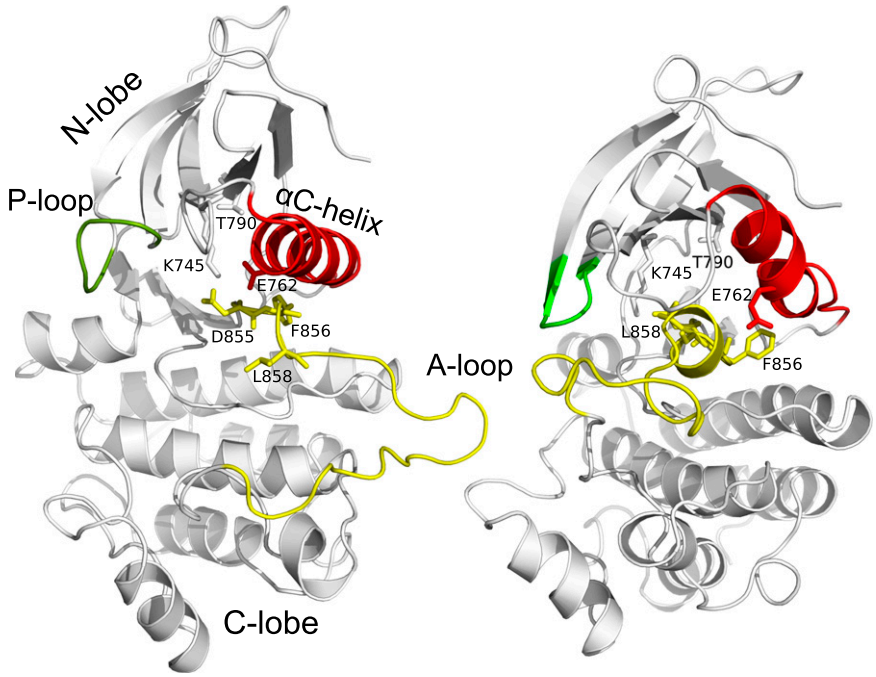

Fig. 1. Comparison of active (*Left*) and Src-like inactive (*Right*) structures of EGFR CD. Key structural elements are colored in red (αC helix), green (P-loop) and yellow (A-loop). Also shown as sticks is the DFG motif and the residues either mutated (L858, T790) or participating in salt-bridge interactions and used to define a collective variable (K745, D855, E762).

in kinase activation. With standard MD and specialized computer hardware, events on the microsecond (at most millisecond) timescale can be sampled, without the need of any a priori knowledge of the relevant reaction coordinates (19). Thus, they are useful to observe unexpected events, but cannot be used to obtain statistically converged populations of high-in-energy secondary conformations. With enhanced sampling methods, such as metadynamics (29, 30), there is no limit on the timescales that can be sampled, but some a priori knowledge of the reaction coordinates is usually needed (30). This need is less stringent when enhanced sampling methods are combined with a multiple-replica approach, as in the case of the parallel-tempering metadynamics simulations approach (PT-metaD) (28, 31). Indeed, it has been shown that PT-metaD is able to converge very complex conformational free-energy surfaces of kinases as a function of a few relevant variables (32, 33).

To elucidate the structural and dynamical consequences of the L858R and T790M mutations on the catalytic domain of EGFR and their synergistic interaction, here we performed extensive PT-metaD simulations of WT-EGFR, three oncogenic mutants: L858R, T790M, and T790M–L858R, and a control mutant, L858I. The total sampling time for each system was at least 3.7 μs, leading to well-converged free-energy surfaces.

Results

We have obtained converged free-energy surfaces (FESs) of the activation of the EGFR catalytic domain and of three EGFR oncogenic mutants (Fig. 2 and Fig. S1). Converged conformational FESs have been obtained for these systems with a fully atomistic explicit-solvent force field. In Fig. 2 we show the FES projected along the first two collective variables (CVs), CV1 and CV2, for the four systems. CV1 is defined as the difference between the distances of the two salt-bridge–forming couples: K745:E762 and K745:D855. When both salt bridges are formed, CV1 is close to 0. This CV is able to characterize the displacement of the functionally important glutamic acid E762 located in the αC-helix of the N lobe in its transition from the so-called "αC-in" to the "αC-out" conformation. The use of CV2 allows for a broad characterization of the A-loop conformations (*Materials and Methods*). It measures the displacement of the heavy atoms of the A loop from the Src-like inactive conformation (Fig. 1, *Right*). When its value is close to 0, the A loop assumes a conformation similar to that found in the Src-like inactive structure. Values larger than 0.7 correspond to open A-loop conformations as in the active state. In-between values of 0.4–0.5 are also observed and will be referred to as "semiclosed" conformations.

WT. The deepest free-energy minimum of WT-EGFR (Fig. 2) corresponds to the autoinhibited conformation where the A loop is folded and forms a short helix at its N-terminal end and the E762 residue of the αC-helix is rotated out of the catalytic site (αC-out) as in the inactive Src-like state (Fig. 1). This inactive conformation is found both in Src-family kinases (34) and cyclin-dependent kinases (CDKs) (35). Notably, there are other less-populated states, with higher free energy (2–4 kcal/mol), corresponding to inactive structures (αC-out) where the αC-helix is deformed and the A loop is semiclosed. The disordered αC-helix basin is characterized by a large deformation of the αC-helix in which only two turns are maintained, whereas the others are disordered (Fig. S2). In the semiclosed basin the αG-helix is surprisingly far from its usual position and interacts both with the P-loop and the C-terminal region of the A loop (Fig. S2). The ensemble of states found in this basin seems to correspond to the intermediate "extended" structures described in ref. 36. The transition from the semiclosed to the inactive state proceeds by local unfolding or "cracking" (37). Finally, in agreement with previous proposals (19, 36), we find that the active conformation of WT-EGFR, in the absence of ATP, is energetically unfavorable and infrequently sampled.

In contrast to the WT, all of the mutants populate to some extent the active conformation. Moreover, the investigated oncogenic mutations alter the free-energy landscape in a global way. Surprisingly, they generate new minima in different locations of the FES and alter the barriers between the minima.

L858R Mutant. The L858R mutation changes the conformational free-energy landscape. The disordered states seen in the case of

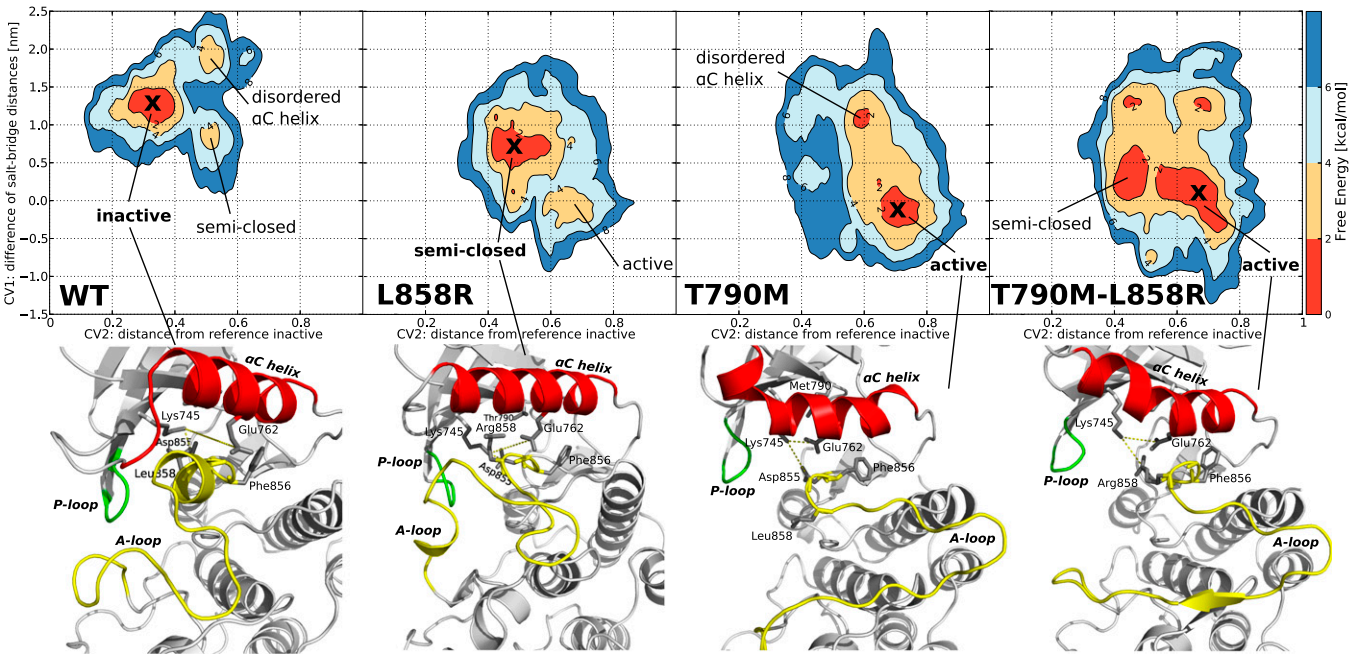

Fig. 2. Free energy surface of wild-type EGFR and the three mutants as a function of CV1 (*y* axis) in nm and CV2 (*x* axis) in arbitrary units. A cross indicates the global free energy minimum for which a representative structure is shown below. Each minimum is also named according to the corresponding main feature. The contour lines are drawn every 2 kcal/mol. In the structures below the free energy surfaces the αC-helix is shown in red, the A-loop in yellow and the P-loop in green. The distances between K745:E762 and K745:D855, whose difference constitute CV1, are displayed with a dashed yellow line.

WT-EGFR have disappeared. The most stable state of L858R corresponds to an ensemble of structures which are in between the active and inactive conformation (Fig. 2). The glutamic acid 762 is on average 0.5 nm closer to the catalytic site compared with WT-EGFR, but the salt bridge with lysine 745 is never formed as in the fully active conformation. The A loop assumes a semiclosed conformation where neither the short helix nor the hairpin at the N terminal end is stable. The observed changes are mainly due to the interactions that the positively charged R858 engages with the three negatively charged residues of the αC-helix. As shown in Fig. 3, R858 makes a salt bridge with the negatively charged glutamic acids 758 and 762 or with the aspartic acid 761, strongly stabilizing the αC-helix. At the same time, these favorable interactions prevent the formation of the short helical turn at the N-terminal region of the A loop (residues 858–862) that is typical of the inactive conformations (18).

In contrast with WT-EGFR, the fully active state is populated (L858R, Fig. 2), albeit it has a thermodynamic penalty (2–4 kcal/mol) with respect to the inactive state. At variance with the structures of the active L858R mutant crystallized in the presence of different inhibitors (18), the side chain of arginine 858 is found either completely exposed to the solvent or interacting with the aspartate 831 of the DFG motif. The electrostatic interactions involving R858 have also been observed in microsecond-long MD simulations (19).

Altogether, the effect of the L858R mutation is twofold: on one hand it destabilizes the inactive conformation; on the other it stabilizes the αC-helix structure. The stabilization of the α- helix

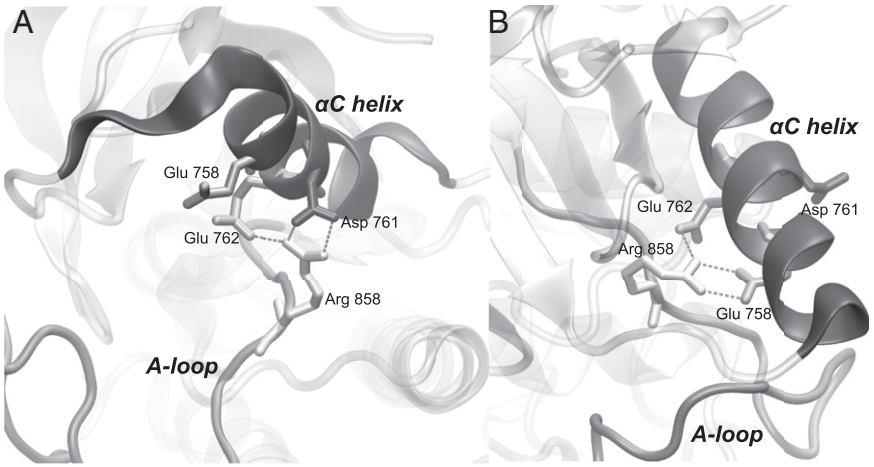

Fig. 3. Close-up of two possible conformations of the L858R mutant in its main free energy minimum. The R858 forms salt bridge interactions with either *A*) D761 and E762 or *B*) E758 and E762.

structure is expected to favor the formation of the dimer by decreasing the entropic penalty associated with the ordering of the disordered interface observed in the WT-EGFR monomer (7, 19).

L858I Mutant. Given the significant effect of the L858R mutation on the conformational free-energy landscape, we wondered what would be the effect of a "neutral" mutation at the same position. We have chosen L858I for two reasons. First, from a physicochemical viewpoint the properties of isoleucine are similar to those of leucine. Second, this specific mutation is not reported among the cancer samples in the Catalogue of Somatic Mutations (COSMIC) database (13). Of the 5,556 EGFR L858 mutations currently reported in the database, L858R, with 5,541 instances, is the most common. However, many other mutations such as L858M, L858P, L858A, L858Q, L858K, L858W, L858V, and L858G are found in at least one case. Because L858I is not found in the database, we expect it either to be neutral or inactivating. The L858I FES is reported in Fig. S3. As expected, it is very similar to the WT FES. The most populated energy basin corresponds to the inactive structure, in agreement with the expected inactivating effect.

T790M Mutant. In the case of the gatekeeper T790M mutant, the change in the FES is more dramatic and the equilibrium is shifted toward the active conformation (Fig. 2). The substitution of the threonine with a methionine stabilizes the active conformation by favoring the formation of a hydrophobic cluster around the phenylalanine F856 of the DFG motif. The cluster involves M790, F856, the methionine M766 in the αC-helix, and other residues of the regulatory spine (38). The M790–M766 distance is on average much shorter in the mutant than the corresponding T790–M766 distance in the WT-EGFR (0.5 ± 0.1 nm vs. 0.9 ± 0.2 nm, respectively). Interestingly, a basin where the αC-helix is disordered is also populated. Conformations in this basin show a deformed αC-helix and the broken salt bridge E762:K745 (Fig. S2). We do not observe the aspartate D858 of the DFG motif pointing outside of the catalytic site (DFG-out) as proposed in ref. 21. However, we have not explored the effect of protonation of the Asp of the DFG, which might stabilize the DFG-out form (36, 39).

T790M-L858R Mutant. The double-mutant T790M–L858R retains some of the features of the single mutants L858R and T790M, but shows also some characteristics hinting at a synergistic effect of the combined mutations. In particular, the αC-in conformation is very stable (CV1 in the interval $[-0.25; 0.25]$ nm) and only a small barrier (<2 kcal/mol) is present for the opening of the A loop, which thus exhibits a greater flexibility populating both the active and the semiclosed basins (Fig. 2). The global effect is to significantly stabilize the fully active state and to lower the free-energy barriers for the A-loop transition between the semiclosed and the fully extended structure. Some inactive conformations are still sampled and correspond to structures in which the αC-helix is distorted (Fig. S2).

Discussion

In contrast with other tyrosine kinases, the active state of the WT-EGFR CD monomer is rarely visited (40). This is in agreement with both its low catalytic activity (6) and the behavior observed in very long MD simulations (19, 36). Conformations in which the αC-helix is partially disrupted correspond to local minima. Because the αC-helix is at the dimer interface, its disorder could affect the stability of the dimer (19). In the L858R mutant, the substitution of the hydrophobic leucine residue in the A loop, with the positively charged arginine, creates new possibilities for the charged residues of the αC helix to interact. This mutation was first proposed to act by destabilizing the inactive conformation of the A loop, favoring the active conformation (6, 18). More recently, Shan et al. proposed an indirect mechanism whereby the mutation reduces the disorder in the αC-helix region, thus favoring dimerization (19). Our observations reconcile both views. We observe a dual effect. On one hand, the inactive state is disfavored due to the destabilization of the inactive-like short α-helix of the A loop. The fully active conformation that, in the absence of ATP, is unfavorable for the WT EGFR, appears as a stable secondary minimum in L858R. On the other hand, the αC-helix is stabilized by a cluster of salt bridges, suppressing the disordered states of WT-EGFR. Both the change in the relative stability of the active vs. the inactive state and the suppression of disorder at the interface, which favors the formation of the dimer, seem to play a role in the activation of L858R. The global effect is to shift the balance toward the active state and favor the formation of the asymmetric dimer.

In agreement with recent proposals (21), the gatekeeper T790M mutant does not seem to act by steric hindrance with the ATP-competitive inhibitors, but rather by stabilizing the active conformation. In this case the methionine participates in the hydrophobic core surrounding the active site. In particular, interacting with M766 of the αC-helix it promotes a more compact active site, with a closer helix, favoring the αC-in conformation. These results agree with the enhanced stabilization of the catalytic site observed when comparing the collective motions of the WT and mutant kinase domain (41). However, the T790M mutant does not suppress the disorder at the dimerization interface, and it is not expected to favor formation of the dimer. This observation might explain the greater frequency of L858R with respect to the T790M mutation in cancer patients.

In light of the present results, we can fully rationalize and quantify the epistatic effect due to the combination of the two mutations. As proposed in ref. 14, the total stabilization of the active state by the double mutations is more than that expected from a simple combination of the stabilization by the two single mutations. However, the effect is not limited to the relative stability of the active and inactive states, but it reshapes the free-energy landscape in a complex manner. Indeed, the combination of the two mutations in the T790M-L858R double mutant not only has the activating and rigidifying effects of the two single mutations, but it also lowers the free-energy barrier of the active–inactive transition of the A loop and stabilizes the correct helical structure of the αC-helix. In the T790M–L858R mutant the αC-helix and the E762:K745 salt bridge are stable. Thus, the conformation adopted by the A loop determines the balance between the active and the inactive form.

The complex changes seen in the conformational free-energy landscapes of the mutants are due both to local changes in the network of electrostatic and hydrophobic interactions and to long-range allosteric couplings. This nontrivial knowledge not only reconciles the observed differences of the mutants with the WT, but can also be used to devise new strategies for the development of TKIs less susceptible to resistance.

Materials and Methods

Protein Structures and Protein Data Bank. The Protein Data Bank (PDB) codes used for the active and inactive WT-EGFR structures are 2GS2 and 2GS7, respectively, and the used sequence comprises the interval L703–Q976 in our notation, L679–Q952 in the PDB notation. The amino acid numbering convention adopted here includes the 24-residue membrane targeting signaling sequence. Mutants and missing residues were modeled with MODELER (42).

Simulation Details. The MD simulations were performed using GROMACS 4.5 (43) with the PLUMED plug-in (44) for Metadynamics calculations. Each system is described by the Amber99SB*-ILDN force field (45) with the dihedral corrections of Best and Hummer (46), solvated with tip3p water molecules (47) and enclosed in a dodecahedron box with periodic boundary conditions. The van der Waals interactions were smoothly shifted to zero between 0.8 and 1.0 nm; the long-range electrostatic interactions were calculated by the particle mesh Ewald algorithm (48), with mesh spaced 0.12 nm, combined with a switch function for the direct space between 0.8 and 1.0 nm for better energy conservation. The system evolves in the canonical ensemble, coupled

with a velocity-rescale thermostat (49) and a time step of 2 fs. Each solvated system was prepared as detailed in *SI Text*.

Enhanced Sampling. PT-metaD in the Well-Tempered Ensemble has been performed for each system using six replicas at increasing temperatures (295, 300, 310, 325, 341, and 358 K) (see *SI Text* and Fig. S4 for further details). All six replicas are subject to the well-tempered metadynamics prescription in which a Gaussian is deposited in the collective variable space every 2 ps with height $W = W_0 e^{-V(s,t)/(f-1)T}$, where W_0=5 kJ/mol is the initial height, T is the temperature of the replica, $f = 10$ is the bias factor, and $V(s,t)$ is the bias potential at time t and CV value s. The following three collective variables are used: the difference between atomic distances $CV1 = d(N_Z^{K745}, C_\delta^{E762}) - d(N_Z^{K745}, C_\gamma^{D855})$; the distance in contact map space to the inactive A-loop conformation $CV2(R) = 1/N\sum_{\gamma\in\Gamma}(D_\gamma(R) - D_\gamma(R_{close}))^2$; the distance in contact map space to the active A-loop conformation $CV3(R) = 1/N\sum_{\gamma\in\Gamma}(D_\gamma(R) - D_\gamma(R_{active}))^2$. $D_\gamma(R)$ is a sigmoidal function that measures the degree of formation of the contact γ in the structure R and is defined as $D_\gamma(R) = w_\gamma \frac{1-(r_\gamma/r_\gamma^0)^n}{1-(r_\gamma/r_\gamma^0)^m}$, where r_γ is the contact distance in the structure R, r_γ^0 is the contact distance in either the reference inactive or active conformation depending if γ is a contact specific to the former or the latter conformation, w_γ is the weight of the contact and is set to 1 for regular contacts and 3 for salt bridges, N is a normalization constant, $n = 6$, and $m = 10$. The set of contacts Γ defining the contact map is defined in *SI Text*. The widths of the Gaussians in the three CV dimensions are $\sigma_1 = 0.015$ nm and $\sigma_{2,3} = 0.0083$. Note that CV2 and CV3 are adimensional.

Convergence and Self-Correcting Properties of PT-metaD. The simulations are run until the free energy in the bidimensional projections and in the monodimensional projections does not change more than 2 kcal/mol in the last 100 ns. This convergence criterion led to 680 ns long simulations for WT-EGFR and T790M-L858R and to 610 ns long simulations for L858R and T790M. To further check the convergence of the reconstructed free energy landscapes and to illustrate the self-correcting properties of PT-metaD calculations we performed two sets of calculations. Starting with a "wrong" free energy surface (the unconverged FES obtained for the WT kinase after 300ns) we run further 500ns of PT-metaD on the L858R and on the new L858I mutant (the total sampling time for each simulation was 3 μs). As shown in Fig. S5, the two free energy surfaces rapidly converge to the original FES, in the case of the L858R mutant and to a FES similar to that of the WT, in the case of the L858I mutant.

Analysis. The free energy surfaces are obtained by integrating the deposited bias during the simulation, as required by the metadynamics algorithm. For convenience, they are shown as a function of two CVs at a time (CV1, CV2) and (CV2, CV3) in the Supplementary Materials. To obtain a representative structure of each free energy basin, a clustering of the set of structures falling within a small (CV1, CV2) patch surrounding each basin has been performed. The single-linkage clustering algorithm of g_cluster program from the GROMACS package has been used with the RMSD on the CA atoms as the distance with a cutoff of 0.2 nm. In the most populated cluster of each basin, the central structure, i.e., the structure with smallest distance to all of the other members of the cluster, has been picked as representative of the basin.

ACKNOWLEDGMENTS. We thank M. Eck for useful discussions and suggestions concerning this work and K. Marino for proofreading the manuscript. We also acknowledge partial support from the Spanish Ministry of Science and Innovation (MICINN, Grant BIO2010-20166); the High End Computing Terascale Resource (HECToR) of the Partnership for Advanced Computing in Europe (PRACE) research infrastructure (Tier-1); (Tier-0) SuperMUC at the Leibniz-Rechenzentrum (LRZ), Germany (Seventh Framework Programme 2007-2013, Grant Agreement RI-283493); and the Emerald supercomputing facility at e-Infrastructure South for the computational resources.

1. Yarden Y, Sliwkowski MX (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2(2):127–137.
2. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* 458(7239): 719–724.
3. Cancer Genome Atlas Research Network (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455(7216): 1061–1068.
4. Hynes NE, Lane HA (2005) ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer* 5(5):341–354.
5. Schreiber AB, Libermann TA, Lax I, Yarden Y, Schlessinger J (1983) Biological role of epidermal growth factor-receptor clustering. Investigation with monoclonal anti-receptor antibodies. *J Biol Chem* 258(2):846–853.
6. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125(6): 1137–1149.
7. Arkhipov A, et al. (2013) Architecture and membrane interactions of the EGF receptor. *Cell* 152(3):557–569.
8. Endres NF, et al. (2013) Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* 152(3):543–556.
9. Jeffrey PD, et al. (1995) Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376(6538):313–320.
10. Noble MEM, Endicott JA, Johnson LN (2004) Protein kinase inhibitors: insights into drug design from structure. *Science* 303(5665):1800–1805.
11. Wood ER, et al. (2004) 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* 64(18):6652–6659.
12. Shigematsu H, Gazdar AF (2006) Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *Int J Cancer* 118(2):257–262.
13. Forbes SA, et al. (2011) COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* 39(Database issue):D945–D950.
14. Hashimoto K, Rogozin IB, Panchenko AR (2012) Oncogenic potential is related to activating effect of cancer single and double somatic mutations in receptor tyrosine kinases. *Hum Mutat* 33(11):1566–1575.
15. Lynch TJ, et al. (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350(21):2129–2139.
16. Pao W, et al. (2004) EGF 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 USA* 101(36):13306–13311.
17. Kobayashi S, et al. (2005) EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 352(8):786–792.
18. Yun CH, et al. (2007) Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. *Cancer Cell* 11(3):217–227.
19. Shan Y, et al. (2012) Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell* 149(4):860–870.
20. Paez JG, et al. (2004) EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304(5676):1497–1500.
21. Yun CH, et al. (2008) The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci USA* 105(6):2070–2075.
22. Wan S, Wright DW, Coveney PV (2012) Mechanism of drug efficacy within the EGF receptor revealed by microsecond molecular dynamics simulation. *Mol Cancer Ther* 11(11):2394–2400.
23. Park JH, Liu Y, Lemmon MA, Radhakrishnan R (2012) Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain. *Biochem J* 448(3):417–423.
24. Pao W, et al. (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med* 2(3):e73.
25. Tam IYS, et al. (2009) Double EGFR mutants containing rare EGFR mutant types show reduced in vitro response to gefitinib compared with common activating missense mutations. *Mol Cancer Ther* 8(8):2142–2151.
26. Lindorff-Larsen K, et al. (2012) Systematic validation of protein force fields against experimental data. *PLoS ONE* 7(2):e32131.
27. Shaw DE, et al. (2009) *Millisecond-scale molecular dynamics simulations on Anton* (ACM Press, New York), pp 1–11.
28. Bussi G, Gervasio FL, Laio A, Parrinello M (2006) Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J Am Chem Soc* 128(41):13435–13441.
29. Sutto L, Marsili S, Gervasio FL (2012) New advances in metadynamics. *WIREs Comput Mol Sci* 2:771–779.
30. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99(20):12562–12566.
31. Bonomi M, Parrinello M (2010) Enhanced sampling in the well-tempered ensemble. *Phys Rev Lett* 104(19):190601–190604.
32. Berteotti A, et al. (2009) Protein conformational transitions: the closure mechanism of a kinase explored by atomistic simulations. *J Am Chem Soc* 131(1):244–250.
33. Lovera S, et al. (2012) The different flexibility of c-Src and c-Abl kinases regulates the accessibility of a druggable inactive conformation. *J Am Chem Soc* 134(5):2496–2499.
34. Xu W, Doshi A, Lei M, Eck MJ, Harrison SC (1999) Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol Cell* 3(5):629–638.
35. De Bondt HL, et al. (1993) Crystal structure of cyclin-dependent kinase 2. *Nature* 363(6430):595–602.
36. Shan Y, Arkhipov A, Kim ET, Pan AC, Shaw DE (2013) Transitions to catalytically inactive conformations in EGFR kinase. *Proc Natl Acad Sci USA* 110(18):7270–7275.
37. Whitford PC, Miyashita O, Levy Y, Onuchic JN (2007) Conformational transitions of adenylate kinase: switching by cracking. *J Mol Biol* 366(5):1661–1671.
38. Kornev AP, Haste NM, Taylor SS, Eyck LF (2006) Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci USA* 103(47):17783–17788.
39. Shan Y, et al. (2009) A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proc Natl Acad Sci USA* 106(1):139–144.
40. Jura N, et al. (2011) Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Mol Cell* 42(1):9–22.
41. Dixit A, Verkhivker GM (2011) Computational modeling of allosteric communication reveals organizing principles of mutation-induced signaling in ABL and EGFR kinases. *PLOS Comput Biol* 7(10):e1002179.

42. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234(3):779–815.
43. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* 4:435–447.
44. Bonomi M, et al. (2009) PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun* 180:1961–1972.
45. Lindorff-Larsen K, et al. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78(8):1950–1958.
46. Best RB, Hummer G (2009) Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J Phys Chem B* 113(26):9004–9015.
47. Mahoney MW, Jorgensen WL (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* 112:8910.
48. Essmann U, et al. (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103: 8577.
49. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126(1):014101.