Bioorganic & Medicinal Chemistry xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Identification of novel drug-resistant EGFR mutant inhibitors by in silico screening using comprehensive assessments of protein structures

Tomohiro Sato [a,b], Hisami Watanabe [b], Keiko Tsuganezawa [b], Hitomi Yuki [b], Junko Mikuni [b], Seiko Yoshikawa [b], Mutsuko Kukimoto-Niino [b], Takako Fujimoto [b], Yumiko Terazawa [b], Motoaki Wakiyama [b], Hirotatsu Kojima [c], Takayoshi Okabe [c], Tetsuo Nagano [c,d], Mikako Shirouzu [b], Shigeyuki Yokoyama [a,b], Akiko Tanaka [b,c], Teruki Honma [b,*]

[a] Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[b] RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[c] Chemical Biology Research Initiative, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[d] Department of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:
Received 14 March 2012
Revised 20 April 2012
Accepted 21 April 2012
Available online xxxx

Keywords:
Epidermal growth factor receptor
Drug resistant mutant
In silico screening
Docking parameter optimization

ABSTRACT

EGFR is a target protein for the treatment of non small cell lung cancer (NSCLC). The mutations associated with the activation of EGFR kinase activity, such as L858R and G719S, destabilize the inactive conformation of EGFR and are closely linked with the development of NSCLC. The additional T790M mutation reportedly causes drug resistance against the commercially available EGFR inhibitors, gefitinib and erlotinib. In this study, we searched for novel G719S/T790M EGFR inhibitors by a new in silico screening strategy, using two datasets. The results of in silico screening using protein-ligand docking are affected by the selection of 3D structure of the target protein. As the first strategy, we chose the 3D structures for in silico screening by test dockings using the G719S/T790M crystal structure, its molecular dynamics snapshots, and known inhibitors of the drug-resistant EGFR. In the second strategy, we selected the 3D structures by test dockings using all of the EGFR structures, regardless of the mutations, and all of the known EGFR inhibitors. Using each of the 3D structures selected by the strategies, 1000 compounds were chosen from the 71,588 compounds. Kinase assays identified 15 G719S/T790M EGFR inhibitors, including two compounds with novel scaffolds. Analyses of their structure–activity relationships revealed that interactions with the mutated Met790 residue specifically increase the inhibitory activity against G719S/T790M EGFR.

1. Introduction

EGFR regulates various cellular events and plays an important role in the development of NSCLC.[1,2] NSCLC represents about 75% of all lung cancers, and the risk of NSCLC increases with mutations in EGFR. EGFR mutations were found in 40% of East Asian patients with NSCLC.[3] The crystal structures of the EGFR kinase domain with the L858R or G719S mutation were reported by Yun et al.[4] We recently determined the kinase-domain structure with the G719S/T790M double mutation and a bound AMPPNP.[5] The L858R and G719S mutations destabilize the inactive conformation of EGFR and increase the kinase activity. Figure 1 shows the kinase-domain structures of the wild-type (WT) and mutant EGFRs. The structures revealed that Ser719 in the G719S and G719S/T790M EGFRs is located at the base of the P-loop (Fig. 1c). The P-loop covers the ATP binding site in the inactive conformation, and shifts upward with the activation of EGFR. The G719S mutation decreases the flexibility of the dihedral angle of the amino acid residue. The movement of the P-loop is associated with the activation of the kinase activity of EGFR. The G719S mutation is considered to fix the loop in the position of the activating form and to destabilize the inactive conformation of EGFR.

Gefitinib (Fig. 2a), launched as Iressa by AstraZeneca[6] in 2002, and Erlotinib (Fig. 2b), launched as Tarceva by Roche[7] in 2004, are quinazoline-based ATP-competitive inhibitors of EGFR. These drugs are effective for NSCLC patients with EGFR activating mutations such as L858R and G719S.[1,2] However, although these drugs are beneficial in the early stage of treatment, drug resistant mutations are often found within one year after starting the treatments.[8–10] To address the problem of the drug resistant mutations, potent inhibitors for the drug resistant EGFRs must be developed.

The resistance to gefitinib is often caused by the T790M mutation. Schiffer et al. reported that the pIC_{50} of gefitinib is 6.6 (IC_{50} = 250 nM) for wild type EGFR (WT EGFR), 7.6 (IC_{50} = 25 nM) for L858R EGFR, and 5.3 for T790M EGFR.[11] The Thr790 residue is

* Corresponding author.
E-mail address: mshonma@ssbc.riken.jp (T. Honma).

0968-0896/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2012.04.042

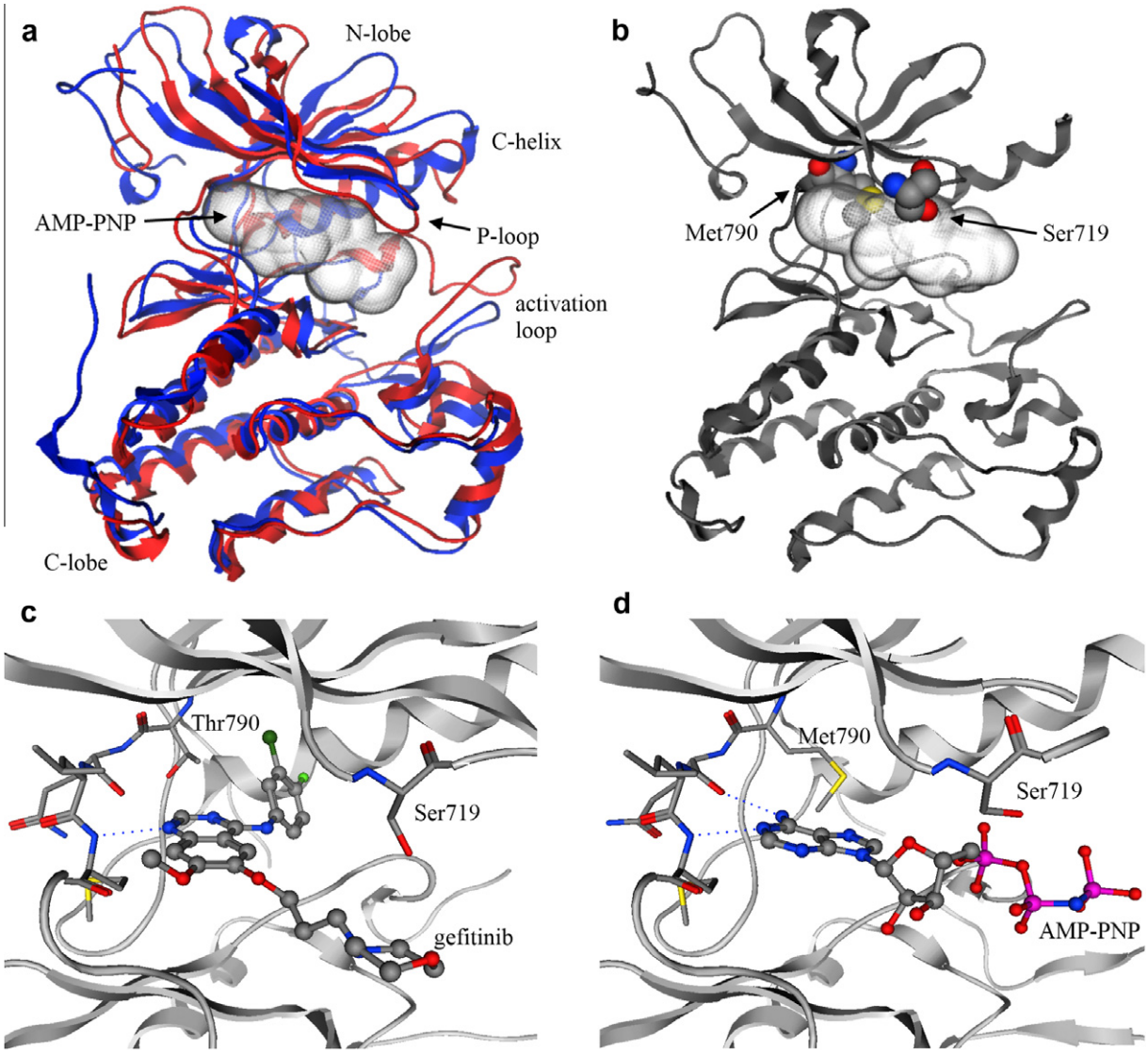


Figure 1. The 3D structures of EGFR. (a) Inactive conformation (2GS7; red) and active conformation (2GS6; blue) of EGFR. AMP-PNP in 2GS7 is shown in the ATP-binding site. (b) The crystal structure of G719S/T790M EGFR. (c) The ATP-binding site of G719S EGFR and gefitinib (2ITO) 4. (d) The ATP-binding site of G719S/T790M and AMP-PNP (deposited in PDB).

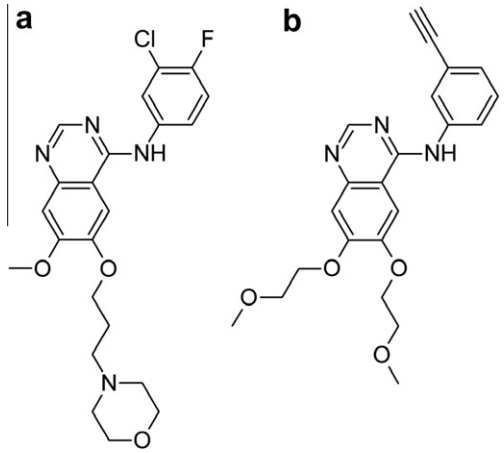


Figure 2. Structures of marketed NSCLC drugs targeting EGFR. (a) Gefitinib. (b) Erlotinib.

located deep within the ATP binding site, and is called a 'gatekeeper' residue (Fig. 1d). The T790M mutation increases the volume of the side chain, and thus the hydrophobic region of the ATP binding site becomes smaller. The shrinkage of binding site is considered to cause steric repulsion with drugs occupying the hydrophobic region and to decrease the affinities of the drugs. In 2008, Yun et al. reported a new mechanism for the drug resistance of L858R/T790M EGFR.[12] They reported the crystal structure of the complex between T790M EGFR and AEE788, and revealed that some quinazoline-based EGFR inhibitors can bind to T790M EGFR. Analyses of the binding affinities showed that the K_m value of EGFR and ATP was decreased by the L858R mutation at first, and then was restored when the T790M mutation occurred in L858R EGFR, while the K_i values of gefitinib for WT and L858R/T790M EGFRs were almost the same. Thus, the drug resistance of the T790M mutation may be caused by the increased affinity between EGFR and ATP. We also investigated whether this mechanism was applicable to the drug resistance of the G719S/T790M double mutant.

In this study, we searched for novel ATP-competitive G719S/T790M EGFR inhibitors by rational drug discovery. Most of the

known EGFR inhibitors share the quinazoline-based scaffolds, such as in gefitinib. As inhibitors for drug resistant EGFRs, several quinazoline-based irreversible inhibitors that form a covalent bond with Cys797 have been developed.[13,14] Recently, Zhou et al. reported a novel covalent pyrimidine-based inhibitor that specifically inhibits T790M EGFR.[15] The discovery of novel scaffolds that are more potent and selective to the drug resistant EGFRs is quite important and reversible inhibitors are desired, in terms of the toxicity.

In silico screening using protein-ligand docking was employed to identify novel G719S/T790M EGFR inhibitors. The accuracy of the docking poses and the efficiency of the in silico screening are often affected by the selection of the 3D structure of the target protein.[16] However, the selection of the proper structure has not been sufficiently investigated in the conventional in silico screening.

The 3D structures and the parameters of docking programs should be determined according to the discrimination performance of the known EGFR inhibitors from decoy compounds in test dockings prior to in silico screening. In this study, we created two datasets to investigate the screening efficiency and obtained the two docking conditions, respectively. As the first strategy, we only used 3D structures originating from the G719S/T790M EGFR crystal structure. To create alternative conformers of G719S/T790M EGFR, we modified the crystal structure using molecular dynamics (MD) simulation, and evaluated the screening efficiencies of both the crystal structure and MD snapshots. As the compound set for test dockings, T790M drug-resistant EGFR inhibitors were used. As the alternative strategy, we used all of the EGFR structures regardless of the mutations and all of the EGFR inhibitors, in order to maximize the diversity of the structural information. To select the best 3D structures for in silico screening, k-PALLAS, developed by Honma and co-workers, was used. k-PALLAS automatically performs test dockings using various protein structures, and selects the best docking conditions by considering the docking accuracy and the screening efficiency. Using the docking conditions determined by the two strategies, in silico screening of G719S/T790M EGFR inhibitors was conducted and the kinase inhibitory activities of the selected compounds were assayed.

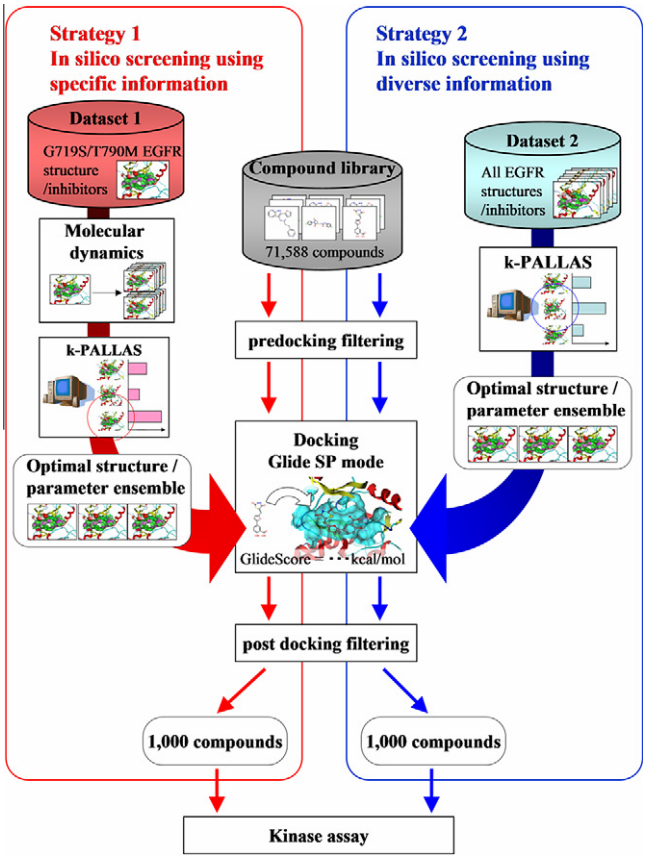


Figure 3. The screening procedure of G719S/T790M EGFR inhibitors, using two optimization strategies for molecular docking.

2. Results and discussion

2.1. Overview of the screening strategy

An overview of the in silico screening of chemical library for G719S/T790M EGFR is shown in Figure 3. At first, the EGFR structures and the parameters for efficient in silico screening were determined, in preparation for the screening. Two datasets containing EGFR structures, known inhibitors, and decoy compounds were created for the protein structure selection and the docking parameter optimization. The first dataset consisted of drug-resistant EGFR specific information, including the crystal structure of G719S/T790M EGFR, its MD snapshots, T790M drug resistant EGFR inhibitors, and decoy compounds. The second dataset consisted of more diverse information about EGFR, including all of the available EGFR crystal structures including any mutants, known EGFR inhibitors, and decoy compounds. In this study, Glide[17] (Schrodinger, Inc.[18]) was used for the protein-ligand docking. In Glide program, the manipulation of a parameter for van der Waals radii scaling, the scaling factor, can modify the distance of the Lennard–Jones potential function and avoid overly strict steric repulsion, by loosening the potential. We tested four settings of the scaling factor, 0.7, 0.8, 0.9, and 1.0 (default), in the test dockings using each EGFR structure.

The in silico screening of library compounds was performed using the two docking condition ensembles obtained from the two datasets, respectively. Totrov et al. reported the usefulness of combined ranking from the docking results of several carefully selected structures.[19] In this study, we tried to identify the optimal ensembles of three docking conditions from all possible combinations. After the preparation, 71,588 compounds were docked using the structures and the parameters. In the post-docking filtering, improper complex structures, such as a complex with a ligand located outside the ATP-binding site, were excluded. The top 1000 compounds were selected from the two docking condition ensembles, according to their docking score values. Their inhibitory activities against WT and G719S/T790M EGFRs were then assayed in vitro.

2.2. Selection of 3D structures for docking from the G719S/T790M EGFR structure and snapshots by molecular dynamics simulation (the first strategy to optimize the docking conditions)

The crystal structure of the G719S/T790M EGFR complex with AMP-PNP was solved by Yoshikawa et al. In this structure, the side chain of Met790 protruded into the binding site (Fig. 1d). In the superposed model with the complex structure of WT EGFR and gefitinib, steric clash of the side chain atoms of Met790 with gefitinib was observed. In the complex structure of T790M EGFR and AEE788 (PDB ID: 2JIU[11]), the side chain atoms of Met790 were shifted by the contact with AEE788. In addition, the B-factors of the side chain atoms of Met790 were higher than those of the neighboring atoms. These observations suggested that the side chain is flexible, and that its conformation in the G719S/T790M EGFR crystal structure is not suitable for the binding of diverse inhibitors. In this study, we employed molecular dynamics (MD) simulations to obtain suitable conformations of G719S/T790M EGFR for the in silico

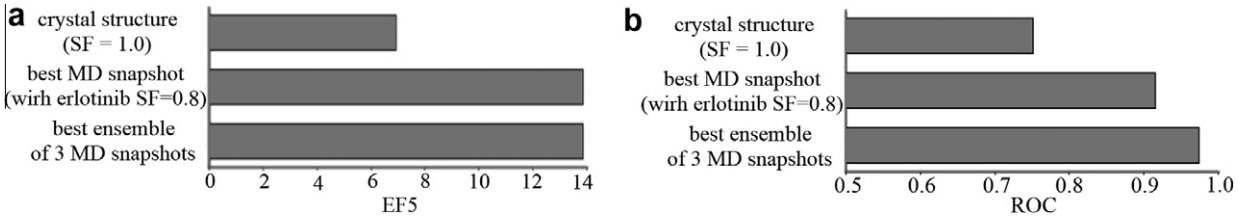


Figure 4. The results of the parameter optimization using the G719S/T790M EGFR crystal structure and the MD snapshots (SF: scaling factor). (a) EF5, (b) ROC score.

screening. The MD simulations were started from superimposed models between G719S/T790M EGFR and the known inhibitors, gefitinib, erlotinib, AEE788, and lapatinib. The known T790M EGFR inhibitors were superposed on G719S/T790M EGFR according to their positions in the complex structures with WT EGFR or other EGFR mutants. After the structural optimizations using the MMFF94x force field,[20–26] 500 ps MD simulations were performed and snapshots were obtained every 1 ps. As the candidates of 3D structures for docking, 100 snapshots were selected using hierarchical clustering, in order to represent diverse conformations. After the preparation of the protein structures, using Protein Preparation Wizard (Schrodinger, Inc.), 404 grid files were created using all possible combinations of the 101 EGFR structures (1 crystal structure and 100 snapshots) and scaling factors ranging from 0.7 to 1.0 in increments of 0.1.

To evaluate the screening efficiency of each structure to detect drug-resistant mutant inhibitors, 15 inhibitors and AMP-PNP, which are known to bind to drug resistant EGFR mutants, were used as positive examples. As decoy compounds, 36 compounds for each inhibitor were selected from ZINC drug-like subset,[27] by the method used in DUD database.[28] In total, 576 compounds were selected as decoys. Using the dataset, the screening efficiencies of the crystal structure and the 100 MD snapshots were calculated by the following procedure.

The 15 known inhibitors, AMP-PNP, and the corresponding 576 decoy compounds were docked into the ATP binding site of EGFR by Glide standard precision (SP) mode, using the 404 grid files created from the EGFR structures with four levels of scaling factors, comprehensively. The screening efficiencies of the docking conditions were evaluated by an enrichment factor[29] at the 5% level (EF5). The assessment of the screening efficiencies of the 404 docking conditions is summarized in Figure 4. By the modification of the G719S/T790M EGFR crystal structure and the selection of the optimal scaling factor, the EF5 improved to 13.9, which is approximately double the EF5 value (6.9) obtained using the intact crystal structure and the default scaling factor value. The structure with the best EF5 was an MD snapshot (392 psec at 300 K) of the erlotinib complex. The scaling factor was 0.8. A comparison of this snapshot to the crystal structure revealed that the side chain of Met790 was retracted from the binding site, which thus seemed more suitable for ligand binding. These results support the usefulness of the MD simulation to modify the structure of the target protein for in silico screening. By combining the three compound rankings using different docking conditions, the EF5 of the best ensemble showed the same value (13.9) as the single condition. Although the EF5 value was not improved by the combination, the ROC score, another metric of screening efficiency,[29] increased from 0.915 (the best single condition) to 0.973 (the best ensemble of three docking conditions).

2.3. Selection of 3D structures for docking from all EGFR structures (the second strategy to optimize the docking conditions)

To investigate the docking conditions using more diverse information, 31 EGFR structures registered in PDB, including two

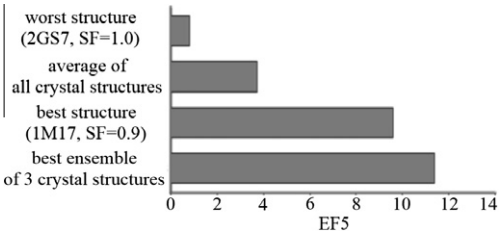


Figure 5. The resulting EF5 values of parameter optimization using all EGFR information (SF: scaling factor).

deposited EGFR structures solved by Yoshikawa et al., were used for test dockings. After the protein preparations, 124 grid files were created using all possible combinations of the 31 EGFR structures and the 4 scaling factors. To evaluate these structures, a dataset consisting of known inhibitors and decoy compounds was created. StARLITe database[30] and a literature search generated 1801 known inhibitors. Currently, StARLITe is available as ChEMBL database from EBI[31] To exclude structurally redundant compounds, the inhibitors were divided into 100 clusters according to their public MACCS keys fingerprint, using the Ward method on Pipeline Pilot.[32] From each cluster, the compound that was closest to the cluster center was selected for the dataset. For each known inhibitor, 36 decoy compounds were selected from ZINC drug-like compounds, using the method employed in DUD database. The 100 known inhibitors and the corresponding 3600 decoy compounds were docked to the EGFR grid files by Glide SP mode.

Figure 5 summarizes the assessment of the screening efficiencies of the 124 docking conditions, using k-PALLAS system. In terms of the EF5, the best docking condition was the EGFR structure of PDB ID: 1M17[33] (scaling factor = 0.9), with a score of 9.6 (EF5). 1M17 is a complex structure of WT EGFR and erlotinib. A comparison of 1M17 to 2GS7, an inactive conformation of WT EGFR with the worst EF5 (0.8), at the default scaling factor (1.0), revealed that 1M17 forms a larger ATP-binding site, with the P-loop moved upward and the DFG-in conformation. Combining the results of the three docking conditions (2ITT (scaling factor: 1.0), 2JIU (1.0), and 2EB3 (1.0)), the EF5 improved from 9.6 to 11.4. All three of the structures were active conformations. 2ITT is a complex of WT EGFR and the ATP-competitive inhibitor, AEE788. 2JIU is a complex of T790M EGFR and AEE788. 2EB3 is a complex of L858R EGFR and AMP-PNP. The average EF5 of the 31 EGFR structures with the default scaling factor setting (1.0) was 3.7. The screening efficiency was successfully improved by approximately threefold from the random choice of a single structure, by obtaining the best ensemble of three EGFR structures and scaling factors.

2.4. In silico screening of library compounds

Using the two optimized docking condition ensembles, 71,588 compounds from the chemical biology research initiative (CBRI)[34] were screened. As the predocking filtering, compounds with molecular weights greater than 600 were excluded. The remaining

70,660 compounds were docked to EGFR by Glide SP mode, using the docking conditions obtained by the two optimization strategies. To evaluate and compare the screening efficiencies of the two optimization strategies using the mutant-specific information and all EGFR information, we did not exclude the known EGFR inhibitors from the CBRI library.

As the post docking filtering, the improper docking poses that were not located in the ATP-binding pocket were eliminated from the results. Using superposed models of the docking poses and the X-ray ligands, the docking poses with a shared volume with the X-ray ligands of more than 220 Å^3 were chosen. Since this criterion was hard to satisfy for small compounds, we also calculated the 'shape Tanimoto similarity' between the docking poses and the X-ray ligands. The similarity indices were calculated for each docking pose and X-ray ligand pair, as the ratio of the shared volume and the summation of their volumes. The compounds with high shape Tanimoto similarity with any of the X-ray ligands (more than 0.45) were also selected for further analysis. For the screening using only G719S/T790M-specific information, the docking poses were filtered by the existence of at least one hydrogen bond between the compounds and the hinge region of EGFR (mainchain oxygen of Gln791, or mainchain oxygen or nitrogen of Met793) because all the known inhibitors of G719S/T790M EGFR form one or more hydrogen bonds at the hinge region.

After the post docking filtering, the compounds were ranked according to the best GlideScore among their docking poses. As a result, the 1000 top-ranked compounds were selected from each of the two in silico screenings. Among them, 169 compounds were redundantly selected in both screenings. By removing the duplicates, we requested CBRI to provide samples of the 1831 compounds and received 1828 compounds, excluding three unavailable compounds.

2.5. Kinase assays

From the results of the in vitro enzymatic assay, 21 compounds were found to inhibit more than 50% of the enzyme activity of WT or G719S/T790M EGFR at 10 μM, and their IC_{50} values for both WT and G719S/T790M were determined; among them, 15 compounds inhibited G719S/T790M EGFR. Both docking condition ensembles, the one optimized by all EGFR information and the other optimized by the drug-resistant mutant EGFR, identified 16 inhibitors. As the two in silico screening strategies detected comparable numbers of inhibitors, a statistically significant difference in the screening efficiencies was not found between the strategies employed in this study. Eleven inhibitors were contained in the 169 compounds selected by both strategies. The hit rate of redundantly selected compounds (11/169 = 7%) exceeded those of the individual methods, suggesting that combining several docking conditions would be useful, especially when only a limited number of compounds could be assayed. Among the 21 compounds, gefitinib, erlotinib, and quercetin, compounds **1**, **4**, and **12**, respectively, are registered in ChEMBL database (version 13) as known EGFR inhibitors. All of the IC_{50} values and the ranks of the docking scores of the 21 compounds are available in the Supplementary data.

The largest group in the 21 inhibitors obtained in this study was the quinazoline-based inhibitors sharing a 4-amino quinazoline scaffold (gefitinib, erlotinib, and compounds **1**–**8** in Fig. 6). Among

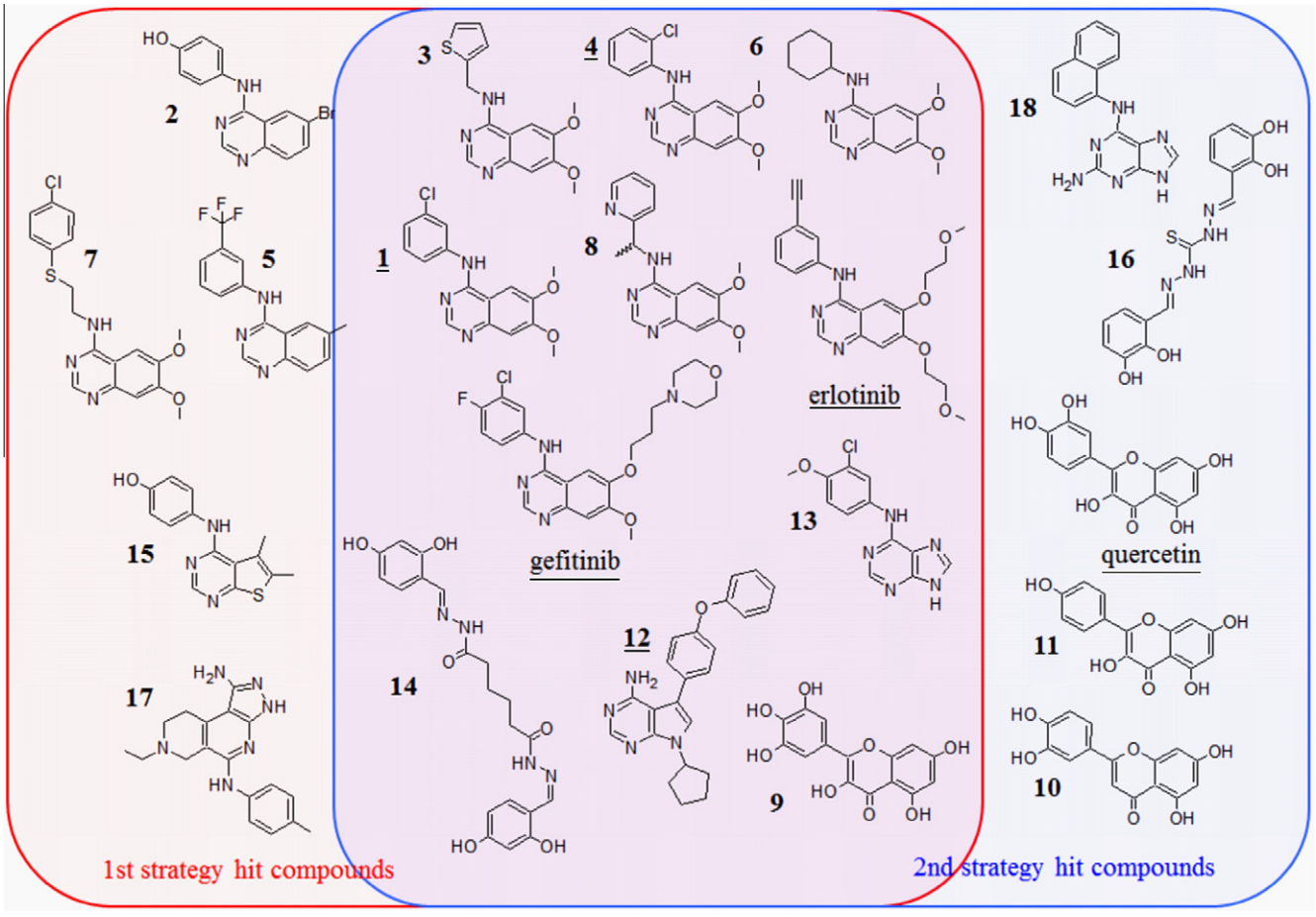


Figure 6. List of 21 candidate compounds detected as EGFR inhibitors. 1st strategy: docking condition obtained from G719S/T790M EGFR-specific information; 2nd strategy: docking condition obtained from all EGFR information. The names of the known inhibitors are underlined.

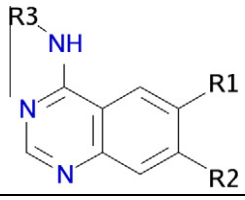
the 10 quinazoline-based inhibitors, seven compounds were detected by both screening strategies, and remaining three compounds were detected only by the docking conditions optimized to the drug-resistant mutant EGFR information (compounds **2**, **5**, and **7** in Fig. 6). Comparing the inhibitory activities between WT and G719S/T790M EGFR, 9 of the 10 quinazoline inhibitors showed selectivities for WT EGFR against G719S/T790M EGFR. Only compound **2** inhibited G719S/T790M EGFR more strongly than WT EGFR (Table 1). On the other hand, out of the 11 remaining non-quinazoline inhibitors, 4 inhibitors (compounds **14**, **16**, **17**, and **18**) showed selectivity for WT EGFR, and seven compounds (quercetin, compounds **9**, **10**, **11**, **12**, **13**, and **15**) showed selectivity for G719S/T790M EGFR. These results suggested that the T790M drug-resistant mutation particularly affects the potency of quinazoline inhibitors.

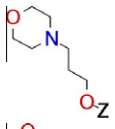
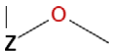
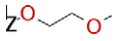
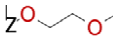
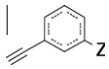
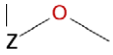
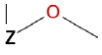
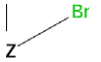
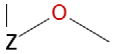
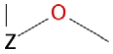
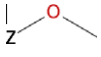
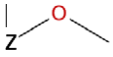
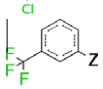
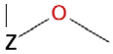
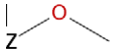
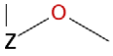
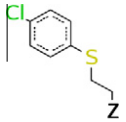
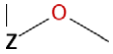
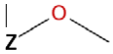
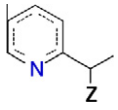
Other than the quinazoline inhibitors, four compounds were found to be quercetin and its analogs (compounds **9**–**11**). These compounds only inhibited G719S/T790M EGFR, with IC_{50} values less than 10 μM. Quercetin and two of its derivatives were only detected by the docking condition ensemble using all EGFR information. Among the remaining seven compounds, **12**–**18**, four compounds showed stronger inhibitory activities against WT EGFR, and three compounds showed stronger inhibitory activities against G719S/T790M EGFR. Compounds **14** and **16** have novel scaffolds, and do not share 2D structural similarities with known EGFR inhibitors, according to a similarity search using the Tanimoto coefficient of the public MDL MACCS keys fingerprint.

While most of the quinazoline inhibitors inhibited both WT and G719S/T790M EGFRs, compound **6** inhibited only WT EGFR (IC_{50}: 0.251 μM) and did not inhibit G719S/T790M EGFR (IC_{50}: >10 μM) (Table 2). Compound **6** has cyclohexane at the deep position of the ATP-binding site, while all of the other quinazoline inhibitors have a phenyl group at this position (Fig. 6d), which forms CH-π and S-π[35] interactions with Met790. The low inhibitory activity of compound **6** against G719S/T790M EGFR, as compared to the other quinazoline inhibitors with a phenyl group, considered to be a consequence of the loss of the CH-π and S-π interactions.

Seven of the eight inhibitors that inhibited G719S/T790M EGFR more strongly than WT EGFR (quercetin, compounds **2**, **9**, **10, 11**, **13**, and **15**) have either a hydroxyl group or a methoxy group at the *para*-position of the phenyl group, which is located in the hydrophobic pocket around Met790. Yoshikawa et al. reported that the phenyl group of gefitinib in the double mutant rotates upwards, as compared to its position in the wild-type structure, and the shift would be an adaptation to the modified binding site of the G719S/T790M double mutant.[5] The shift bring the phenyl

Table 1
Quinazoline-based inhibitors and their IC_{50} values against WT and G719S/T790M EGFRs

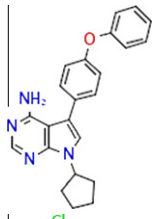
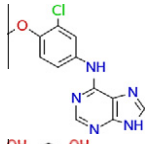
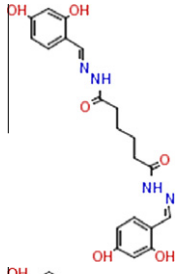
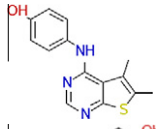
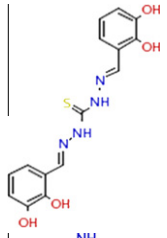
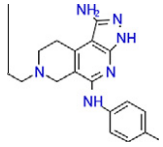
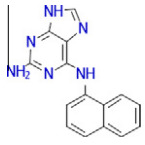


	R1	R2	R3	IC_{50} (μm)		Selectivity (WT/DM)
				WT EGFR	DM EGFR	
Gefitinib				0.006	0.114	0.053
Erlotinib				0.006	0.144	0.042
1				0.004	0.183	0.022
2				1.226	0.602	2.037
3				0.178	1.094	0.163
4				0.077	2.508	0.031
5				0.306	2.585	0.118
6				0.251	>10	<0.025
7				3.524	>10	<0.352
8				5.954	>10	>0.595

DM EGFR: G719S/T790M double mutant EGFR. The compounds investigated using induced fit docking are denoted with bold font.

Please cite this article in press as: Sato, T.; et al. *Bioorg. Med. Chem.* (2012), http://dx.doi.org/10.1016/j.bmc.2012.04.042

Table 2
Other classes of inhibitors and their IC_{50} values against WT and G719S/T790M EGFRs

Compound	Structure	IC50 (μM)		Selectivity (WT/DM)
		WT EGFR	DM EGFR	
12		1.533	0.46	3.333
13		>50	1.186	>42.159
14		0.102	1.357	0.075
15		7.439	1.821	4.085
16		0.139	3.623	0.038
17		1.994	>10	<0.199
18		6.397	>10	<0.640

DM EGFR: G719S/T790M double mutant EGFR. The compound investigated using induced fit docking is denoted with bold font.

group close to Lys745, thus, the hydrogen bond between Lys745 and hydroxyl group, or methoxy group on the phenyl ring could stabilize this conformation. To investigate the interactions between G719S/T790M EGFR and the above compounds in detail, we employed Induced Fit Docking in Schrodinger Maestro suite[36] for the prediction of the complex structures considering the protein flexibility. Figure 7 shows the resulting complex structures of G719S/T790M EGFR with compounds **13**, **2**, and gefitinib, as well as compound **6**. Apart from compound **6**, which does not inhibit G719S/T790M EGFR at all, the distances between the carbon atoms of the phenyl groups at the hydrophobic pocket and the nearest hydrogen atom at the methyl group of Met790 ranged from 2.83 to 3.07 Å, and the distances between the carbon atoms of the aromatic rings and the sulfur atom of Met790 ranged from 3.66 to 4.03 Å, indicating that CH-π and S-π interactions would contribute to the binding of G719S/T790M EGFR and the inhibitors. Compound **2** was the only quinazoline-based inhibitor with G719S/T790M selectivity over WT, showing more than twofold selectivity (Table 2). The unique structural feature of compound **2** is the hydroxyl group at the para-position of the phenyl group in the hydrophobic pocket, while the other quinazoline-based inhibitors have either a halogen or no functional group at this position. The CH-π interactions between the phenyl group and the hydrogen bonds between the hydroxyl group and Glu762 were predicted by the docking model. These T790M mutant specific interactions would contribute to the selective inhibition of G719S/T790M EGFR. Compound **13** was the most selective G719S/T790M EGFR inhibitor detected in this study. The IC_{50} values of compound **13** against WT

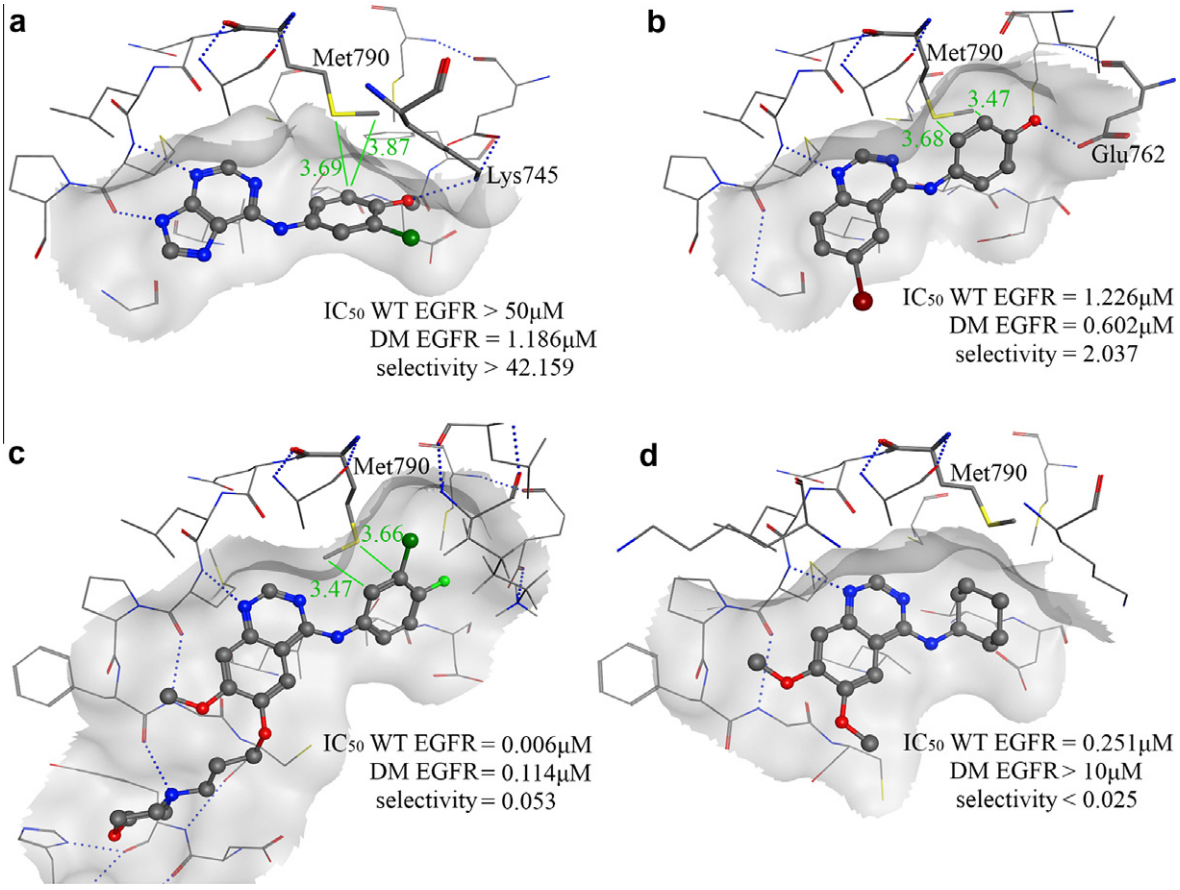


Figure 7. (a) The docking pose of **13** and G719S/T790M EGFR. (b) The docking pose of **2** and G719S/T790M EGFR. (c) The docking pose of gefitinib and G719S/T790M EGFR. (d) The docking pose of **6** and G719S/T790M EGFR.

EGFR and G719S/T790M EGFR were >50 μM (12.7% inhibition at 50 μM) and 1.2 μM, respectively, showing more than 42-fold selectivity (Table 2). Compound **13** has a methoxy group at the para-position of the phenyl group. At the *para*-position of the phenyl groups of compounds **2** and **13**, the oxygen atoms formed hydrogen bonds with either Glu762 or Lys745, and retain the CH-π and S-π interactions with Met790. These predicted complex structures suggested that the selectivity of the drug-resistant EGFR against WT EGFR could be improved by modifying the substituents at the para-position of the phenyl group in the hydrophobic pocket. The crystal structures of the complexes between G719S/T790M EGFR and the inhibitors detected in this study are currently analyzed by Yokoyama and co-workers.

3. Conclusion

In this paper, we identified 21 EGFR inhibitors including 15 G719S/T790M EGFR inhibitors from about 2000 candidate compounds selected by in silico screening. To optimize the docking conditions for the in silico screening, we employed k-PALLAS system to select the proper protein structures and parameters for the protein-ligand docking. As datasets for the optimization, two different datasets, consisting of drug resistant EGFR-specific data (strategy 1) and all EGFR data (strategy 2), were used. From the 1000 compounds selected by each of the two strategies, we identified 16 inhibitors by measuring the EGFR kinase activity. The two optimization strategies did not greatly differ in their screening efficiencies; however, more diverse scaffolds were detected by strategy 2. In terms of the screening efficiency, 169 compounds, which were redundantly selected by both strategies, showed a much higher hit rate (7%), as compared to the hit rates of the individual strategies.

We analyzed the structure–activity relationship between the docking models of the identified inhibitors and the G719S/T790M EGFR selectivity. For the G719S/T790M EGFR inhibitory activity, the contacts of the phenyl groups with the side chain of Met790, which are suitable for CH-π and/or S-π interactions, were found to be important.

Among the inhibitors that had higher G719S/T790M EGFR inhibitory activities than those against WT EGFR, the *para*-position substituents of the aromatic ring on Met790 were found to increase the inhibitory activity against G719S/T790M EGFR. In particular, compound **13**, which has a methoxy group at the *para*-position of the aromatic ring, showed 42-fold higher G719S/T790M EGFR inhibitory activity than that against WT EGFR. Among the scaffolds investigated in this study, the 4-anilino quinazoline scaffold seemed to be strongly influenced by the T790M mutation. Currently, among the drug resistant EGFR inhibitors, irreversible inhibitors have mainly reported.[13] The structure–activity relationships for specific interaction with the G719S/T790M mutant revealed in this study may contribute to the development of more potent and reversible drug resistant EGFR inhibitors.

4. Materials and methods

4.1. G719S/T790M EGFR information for docking optimization

The following data were used as the G719S/T790M EGFR dataset. As the EGFR structures for the docking, the crystal structure of G719S/T790M EGFR and 100 MD snapshots obtained by the pro-

Table 3
AMP-PNP and G719S/T790M EGFR inhibitors used as positive examples of k-PALLAS in the second optimization strategy

Compound	Reference
AEE788	Yun et al., 2008[11]
AMP-PNP	Yokoyama et al.,[5]
AV-412	Suzuki et al., 2007[36]
Bionet[38] 11N-067	In house preliminary assay
CI-1033	Carter et al., 2005[39]
CL-387	Carter et al., 2005,[39] Kobayashi et al. 2005[40]
EKB-569	Carter et al., 2005[39]
Erlotinib	Piepp et al., 2008[41]
Gefitinib	Yun et al., 2008[11]
HKI-272	Kwak et al., 2005[12]
Lapatinib	Glimer et al., 2008[42]
SU-11464	Mitsudomi et al. (2006)[43]
Unnamed compound **1**	Michalczyk et al., 2008[13]
Unnamed compound **2**	Michalczyk et al., 2008[13]
Unnamed compound **3**	Michalczyk et al., 2008[13]
ZD6474	Sequist 2007[44]

cedure described in the next section were used. Protein Preparation Wizard (Schrodinger, Inc.) was then used to add the hydrogens to the structure and to protonate/deprotonate at pH 7.4. Each structure was combined with four scaling factors for van der Waals radii scaling from 0.7 to 1.0, and the grid files were created for each structure/parameter combination. In total, 404 grid files were prepared for test dockings. As the compound dataset for test docking, 16 compounds were used as positive examples (Table 3). These compounds were obtained from literatures[12–14,37,39–44] and our preliminary assay. As decoys, 576 compounds were selected from ZINC drug-like subset, by the selection method used in DUD database.[28] The only modification of the selection method for decoys was that we excluded the duplicated compounds to select non-redundant decoys, while the original DUD method can select the same compound corresponding to multiple positive compounds.

4.2. Molecular dynamics simulation of G719S/T790M EGFR

The crystal structure of G719S/T790M EGFR was modified by the following procedure. At first, known inhibitors of drug-resistant EGFR were superposed on the ATP-binding site. Gefitinib, erlotinib, lapatinib, and AEE788 were reported to bind to T790M EGFR, and were used in this study. The crystal structures of the inhibitors complexed with EGFR were superposed on the G719S/T790M crystal structure to replace ANP-PNP by the inhibitors.

After the ligand replacement, the complex structures of G719S/T790M EGFR were optimized by molecular mechanics (MM), using AMBER99 force field[45–48] in the program MOE.[49] The amino acid residues with atoms that were more than 8 Å away from each ligand were fixed. The atoms of the mainchain and ligands were tethered. The hydrogen bonds between the ligands and EGFR were conserved, by restricting the distance between the hydrogen-bond donors and acceptors to less than 3.5 Å. After the optimization, the tethering of the ligand atoms was canceled, and the complex structures were optimized again. The eight structures both before and after the second optimization were used as the initial structures for MD simulations.

MOE Bulk Modeling & Simulations module was used for the MD simulation with AMBER99 forcefield. The temperature was gradually increased (25 K per 8 ps simulations) before the simulation at 300 K for 500 ps. By sampling snapshots every 1 psec from the simulations of eight structures, a total of 4000 structures were obtained.

To exclude similar structures from the snapshots, hierarchical clustering by Ward method, according to the 3D coordinate data, was used. At first, the snapshots were split into 10 clusters according to the X/Y/Z coordinates of the Met790 side chain atoms. In addition, each cluster was divided into 10 clusters, according to the coordinates of the alpha-carbon atoms of the residues within 4.5 Å from Met790. From each of the 100 clusters, the most energetically stable snapshot was selected as a candidate for the docking.

Table 4
MD snapshots used for in silico screening

Ligand	Scaling factor	Structural optimization before MD	Simulation time
Lapatinib	1.0	MM with ligand tether	314 psec
Erlotinib	0.8	MM with ligand tether	392 psec
Gefitinib	0.9	MM with ligand tether → MM without ligand tether	438 psec

After the assessment of the in silico screening efficiencies of the docking conditions using k-PALLAS, the three snapshots shown in Table 4 were selected as the docking conditions for the screening.

4.3. Dataset of all EGFR data for docking optimization

The following data were used as the dataset of all EGFR data including WT and any mutants. As EGFR structures for docking, 31 crystal structures of the EGFR kinase domain registered in PDB as of August, 2009, including two deposited EGFR structures, were used. The tested scaling factors ranged from 0.7 to 1.0, and were the same as in the case of the G719S/T790M EGFR-specific optimization. As a compound dataset for test docking, published EGFR inhibitors were collected from StARLITe database and by literature mining. StARLITe was previously a medicinal chemistry database containing biological activity and/or binding affinity data between various compounds and proteins. In total, 1801 EGFR inhibitors were chosen. As positive examples for test dockings, 100 representative inhibitors were selected by hierarchical clustering of the 1801 inhibitors, using Ward method according to their public MACCS keys fingerprint on Pipeline Pilot. The selected inhibitors were the cluster centers of the 100 clusters. As decoy compounds, 3600 compounds were selected from ZINC drug-like subset as negative examples in the same manner as for the preparation of the G719S/T790M EGFR-specific dataset. These 3700 compounds were used for the assessment of screening efficiencies.

4.4. k-PALLAS

k-PALLAS was developed by Honma and co-workers as a semi-automatic system to optimize the docking conditions for in silico screening. k-PALLAS can select one best docking condition or docking condition ensemble for efficient in silico screening. As datasets, the 3D structures of the target protein, the known inhibitors, and some decoy compounds are needed for k-PALLAS. Using all of the structures and docking parameters, comprehensive test dockings of the compound dataset were automatically performed. In this study, the optimal structures and parameters were selected according to the EF5 and ROC scores. EF (enrichment factor) is one of the most famous measures for assessing the efficiency of in silico screening, and it indicates the ratio of the number of active compounds obtained by in silico screening against that generated by random selection at the pre-defined sampling percentage. EF is often used to evaluate an early recognition property, such as EF5 (5% enrichment factor), of screening methods. On the one hand, ROC score evaluates the entire range (0–100% sampling) of screening efficiency, while EF evaluates the screening efficiency of only a particular sampling percentage. ROC score is defined as the ratio of the area under the receiver operation characteristic (ROC) curve of the in silico screening and the ideal selection. ROC curve plots the ratio of true positive samples (detected active

compounds) on the axis of false positive fractions and ranges from 0 (0%) to 1 (100%). k-PALLAS also automatically calculates the screening efficiency measures for ensembles of multiple docking conditions. To combine multiple docking results, compounds in the datasets were ranked by the best docking score (GlideScore) among all of the docking poses from the employed docking conditions. In this study, the combinations of up to three docking conditions were calculated. To assess all possible combinations of the three docking conditions, results of millions of patterns must be calculated. To save calculation time, we evaluated all combinations of two docking conditions at first, and selected the top 200 combinations according to their ROC scores. Then, we added the third docking conditions to the 200 docking condition pairs.

4.5. Protein-ligand docking

Standard precision (SP) mode of Schrodinger Glide[17] was used to dock the compounds to EGFR. In the optimization of the docking conditions using k-PALLAS, the scaling factors representing the cofactor to manipulate the influence of the distance in the Lennard–Jones potential function were investigated along with the EGFR structures. The scaling factors tested in this study ranged from 0.7 to 1.0, in increments of 0.1. As the preparation for docking, hydrogen atoms were added to the EGFR structures at pH 7.4, using Protein Preparation Wizard (Schrodinger, Inc.). The grid files were then created. The parameters for the grid generation, other than the scaling factor, were set by the default settings. For the preparation of the compound data for docking, tautomers, stereoisomers and protonation/deprotonation states at pH7.4 of the active and decoy compounds were enumerated, using Pipeline Pilot[30] of SciTegic,. The additional ring conformations of the compounds were then generated by LigPrep (Schrodinger, Inc.).

4.6. Chemical library used for in silico screening

The compounds stored in CBRI[32] are used for the screening. The CBRI library has been constructed in the University of Tokyo by Nagano et al. for the purpose of chemical biology research and drug discovery on the support of Targeted Proteins Research Program, which is a Japanese national project. The CBRI library is designed for rapid identification of biologically active compounds for arbitrary protein or phenotype. As of early 2008, the library contained 71,588 compounds, and these compounds were used for the in silico screening as a screening source.

4.7. Expression and purification of G719S/T790M EGFR samples

The DNA fragment encoding residues 695-1022 of human EGFR was cloned into pFastBacHT-C (Invitrogen[50]). The resulting plasmid was used to generate the G719S/T790M mutant plasmid, using a QuikChange Site-directed mutagenesis Kit (Stratagene[51]). The baculovirus for the His-tagged G719S/T790M mutant EGFR tyrosine kinase was created with the Bac-to-Bac Baculovirus Expression System (Invitrogen), and was used to infect Spodoptera frugiperda Sf9 cells. Cells were harvested 48 h after infection, resuspended in buffer A (20 mM Tris–HCl (pH 8.0), 300 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, 10 mM imidazole) with Complete EDTA-free protease inhibitor cocktail (Roche[6]), and then disrupted by nitrogen cavitation. After centrifugation, the supernatant was incubated with Ni-NTA agarose beads (QIAGEN[52]), and then the protein was eluted with buffer A containing 250 mM imidazole. The His-tag was cleaved by TEV protease, and then the protease and the cleaved His-tag were removed by passages through Ni-NTA agarose beads. The protein sample was further purified by Mono Q and Superdex 200 column chromatography (GE Healthcare[53]), in a final buffer containing 20 mM Tris–HCl (pH 8.0), 150 mM NaCl, 2 mM DTT, and 10% glycerol.

4.8. Compounds

The compound dataset used for the in silico screening was provided by CBRI. All of the compounds tested in this study were purchased from commercial sources and their purities were validated by CBRI. The suppliers and the purities of the 21 WT and G719S/T790M EGFR inhibitors detected in this study are as follows.

Gefitinib was purchased from Pharmeks.[54] The supplier code is P2001S-336532. The purity was 94.2%, as measured using diode array: absorption at 214 nm. Erlotinib was purchased from Pharmeks. The supplier code is P2001S-339165. The purity was 95.2%, as measured using diode array: absorption at 214 nm. Quercetin was purchased from Maybridge.[55] The supplier code is XBX 00307. The purity was 83.8%, as measured using evaporative light scattering. Compound **1** was purchased from Sigma–Aldrich.[56] The supplier code is T4182. The purity was 95.5%, as measured by diode array: absorption at 214 nm. Compound **2** was purchased from ChemBridge.[57] The supplier code is 6947187. The purity was 92.1%, as measured using diode array: absorption at 214 nm. Compound **3** was purchased from Bionet. The supplier code is 12N-055. The purity was 87.7%, as measured using diode array: absorption at 214 nm. Compound **4** was purchased from Bionet. The supplier code is 11N-059. The purity was 99.1%, as measured using evaporative light scattering. Compound **5** was purchased from ChemBridge. The supplier code is 6133504. The purity was 95.1%, as measured using diode array: absorption at 214 nm. Compound **6** was purchased from Maybridge. The supplier code is GK 03357. The purity was 99.2%, as measured using evaporative light scattering. Compound **7** was purchased from Bionet. The supplier code is 12N-063. The purity was 91.4%, as measured using diode array: absorption at 214 nm. Compound **8** was purchased from Enamine.[58] The supplier code is T5641886. The purity was 90.5%, as measured using diode array: absorption at 214 nm. Compound **9** was purchased from Sigma–Aldrich. The supplier code is M6760. The purity was 95.5%, as measured using diode array: absorption at 214 nm. Compound **10** was purchased from Prestwick.[59] The supplier code is Prestw-870. The purity was 87.4%, as measured using diode array: absorption at 214 nm. Compound **11** was purchased from Prestwick. The supplier code is Prestw-1098. The purity was 86.4%, as measured using evaporative light scattering. Compound **12** was purchased from Sigma–Aldrich. The supplier code is C8863. The purity was 90.5%, as measured using diode array: absorption at 214 nm. Compound **13** was purchased from InterBioScreen.[60] The supplier code is STOCK6S-27559. The purity was 92.6%, as measured using diode array: absorption at 214 nm. Compound **14** was purchased from ChemBridge. The supplier code is 5474286. The purity was 81.9%, as measured using diode array: absorption at 214 nm. Compound **15** was purchased from InterBioScreen. The supplier code is STOCK3S-63907. The purity was 97.3%, as measured using evaporative light scattering. Compound **16** was purchased from Enamine. The supplier code is T0517-6476. The purity was 91.4%, as measured using evaporative light scattering. Compound **17** was purchased from InterBioScreen. The supplier code is STOCK5S-63648. The purity was 90.8%, as measured using diode array: absorption at 214 nm. Compound **18** was purchased from InterBioScreen. The supplier code is STOCK6S-38272. The purity was 97.0%, as measured using diode array: absorption at 214 nm.

4.9. Kinase assays for WT and G719S/T790M EGFRs

We first measured the Km values of ATP with both WT and G719S.T790M EGFRs. The effect of the ATP concentration on the

kinase activities were measured with 20–2000 μM ATP. The K_m value for this double mutant was determined 126.6 μM, and that for wild type was 25.3 μM. To avoid the effect of the affinity and to reflect the cellular ATP condition, we added a final concentration of 500 μM ATP to our kinase assays.

Kinase assays were performed using a Labchip EZ Reader II system (Caliper Life Sciences[61]) in a 96-well plate (Greiner[62]) at room temperature. The assay was performed in a total volume of 60 μl in a buffer containing 50 mM HEPES, pH 7.4, 10 mM MgCl2, 0.01% Brij-35, 1 mM DTT, 1% protease inhibitor (Calbiochem[63]), and 1% phosphatase inhibitor (Calbiochem). The reaction was started by the addition of ATP and a fluorescein-labeled substrate peptide, FL-Peptide 22 (5-FAM-EEPLYWSFPAKKK-$CONH_2$) (Caliper Life Sciences), after 30 min preincubation of the reaction mixture. The concentrations of WT EFGR (Carna Biosciences[64]) and purified G719S/T790M EGFR were 5 nM. The substrate peptide concentration was 1.5 μM. The ATP concentration was 500 μM, which was in excess for the assays. The compounds, dissolved in DMSO were added at the beginning of the preincubation period. The concentrations of the tested compounds were 10 μM for the screening, and varied from 0.004 to higher than 10 μM for the IC_{50} determinations. The final DMSO concentration was 0.6% in all assays, which represented 0.5% from the compounds and 0.1% from the substrate peptide. The reaction was terminated by the addition of 10 μl termination buffer (50 mM HEPES, pH 7.4, 140 mM EDTA, 0.01% Brij-35) after a 120 min incubation of the reaction mixture for WT or a 60 min incubation for G719S/T790M EGFR. A 60 μl aliquot of each reaction mixture was transferred to a 384-well plate (Coster[65]), and the substrate phosphorylation was measured by the mobility shift assay method.

4.10. Induced fit docking

The induced fit docking was performed using Schrodinger Maestro Suite. The procedure was described by Sherman et al.[34] The crystal structure of G719S/T790M EGFR was used as the protein structure. As the initial docking, gefitinib and compounds 2, 6, and 13 were docked into the structure, using a relatively low scaling factor (0.4) for the both protein and the ligand nonpolar atoms. After the initial docking, side chain and backbone refinement by the homology modeling program Prime in Schrodinger Maestro Suite, and minimization of the docking pose were performed. Redocking of each ligand by Glide was performed for the refined protein structures within 30.0 kcal/mol of the best structure. The scoring function of Glide XP (extra precision) mode was used to evaluate the docking pose.

Acknowledgments

The authors thank Dr. Furuya and Dr. Inoue of PharmaDesign, Inc. for providing the StARLITe data about EGFR inhibitors, and Mr. Tsuruse and Dr. Ishida of Nittetsu Hitachi Systems Engineering, Inc. for the implementation of k-PALLAS. This work was supported in part by the Targeted Proteins Research Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2012.04.042.

References and notes

1. Paez, J. G.; Janne, P. A.; Lee, J. C.; Tracy, S.; Greulich, H.; Gabriel, S.; Herman, P.; Kaye, F. J.; Lindeman, N.; Boggon, T. J.; Naoki, K.; Sasaki, H.; Fujii, Y.; Eck, M. J.; Sellers, W. R.; Johnson, B. E.; Meyerson, M. *Science* **2004**, *304*, 1497.
2. Lynch, T. J.; Bell, D. W.; Sordella, R.; Gurubhagavatula, S.; Okimoto, R. A.; Brannigan, B. W.; Harris, P. L.; Haserlat, S. M.; Supko, J. G.; Haluska, F. G.; Louis, D. N.; Christiani, D. C.; Settleman, J.; Haber, D. A. *N. Eng. J. Med.* **2004**, *350*, 2129.
3. Thomas, R. K.; Weir, B.; Meyerson, M. *Clin. Cancer Res.* **2006**, *12*, 4384s.
4. Yun, C. H.; Boggon, T. J.; Li, Y.; Woo, M. S.; Greulich, H.; Meyerson, M.; Eck, M. J. *Cancer Cell* **2007**, *11*, 217.
5. Yoshikawa, S.; Kukimoto-Niino. M.; Parker, L.; Handa, N.; Terada, T.; Fujimoto, T.; Terazawa, Y.; Wakiyama, M.; Sato, M.; Sano, S.; Kobayashi, T.; Tanaka, T.; Chen, L.; Liu, Z. J.; Wang, B. C.; Shirouzu, M.; Kawa, S.; Semba, K.; Yamamoto, T.; Yokoyama, S. *Oncogene* **2012**, http://dx.doi.org/10.1038/onc.2012.21. [Epub ahead of print].
6. AstraZeneca, London, U.K. http://www.astrazeneca.com/.
7. Roche, Ltd, Basel, Switzerland http://www.roche.com/.
8. Pao, W.; Miller, V. A.; Politi, K. A.; Riely, G. J.; Somwar, R.; Zakowski, M. F.; Kris, M. G.; Varmus, H. *PLoS Med.* **2005**, *2*, e73.
9. Kobayashi, S.; Boggon, T. J.; Dayaram, T.; Janne, P. A.; Kocher, O.; Meyerson, M.; Johnson, B. E.; Eck, M. J.; Tenen, D. G.; Halmos, B. *N. Eng. J. Med.* **2005**, *352*, 786.
10. Kosaka, T.; Yatabe, Y.; Endoh, H.; Yoshida, K.; Hida, T.; Tsuboi, M.; Tada, H.; Kuwano, H.; Mitsudomi, T. *Clin. Cancer Res.* **2006**, *12*, 5764.
11. Schiffer, H. H.; Reding, E. C.; Fuhs, S. R.; Lu, Q.; Piu, F.; Wong, S.; Littler, P. L.; Weiner, D. M.; Keefe, W.; Tan, P. K.; Nash, N. R.; Knapp, A. E.; Olsson, R.; Brann, M. R. *Mol. Pharmacol.* **2007**, *71*, 508.
12. Yun, C. H.; Mengwasser, K. E.; Toms, A. V.; Woo, M. S.; Greulich, H.; Wong, K. K.; Meyerson, M.; Eck, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2070.
13. Kwak, E. L.; Sordella, R.; Bell, D. W.; Godin-Heymann, N.; Okimoto, R. A.; Brannigan, B. W.; Harris, P. L.; Driscoll, D. R.; Fidias, P.; Lynch, T. J.; Rabindran, S. K.; McGinnis, J. P.; Wissner, A.; Sharma, S. V.; Isselbacher, K. J.; Settleman, J.; Haber, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7665.
14. Michalczyk, A.; Kluter, S.; Rode, H. B.; Simard, J. R.; Grutter, C.; Rabiller, M.; Rauh, D. *Bioorg. Med. Chem.* **2008**, *16*, 3482.
15. Zhou, W.; Ercan, D.; Chen, L.; Yun, C. H.; Li, D.; Capelletti, M.; Cortot, A. B.; Chirieac, L.; Iacob, R. E.; Padera, R.; Engen, J. R.; Wong, K. K.; Eck, M. J.; Gray, N. S.; Janne, P. A. *Nature* **2009**, *462*, 1070.
16. Rueda, M.; Bottegoni, G.; Abagyan, R. *J. Chem. Inf. Model.* **2010**, *50*, 186.
17. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739.
18. Schrödinger Inc., New York, NY http://www.schrodinger.com/.
19. Totrov, M.; Abagyan, R. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178.
20. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.
21. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520.
22. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553.
23. Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587.
24. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616.
25. Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 720.
26. Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730.
27. Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model* **2005**, *45*, 177.
28. Huang, N.; Shoichet, B. K.; Irwin, J. J. *J. Med. Chem.* **2006**, *49*, 6789.
29. Truchon, J. F.; Bayly, C. I. *J. Chem. Inf. Model* **2007**, *47*, 488.
30. StARLITe; Inpharmatica Ltd: London, UK, 2007.
31. EBI (European Bioinformatics Institute) http://www.ebi.ac.uk/.
32. Pipeline Pilot; Accelrys Software Inc.: San Diego California, USA, 2007.
33. Stamos, J.; Sliwkowski, M. X.; Eigenbrot, C. *J. Biol. Chem.* **2002**, *277*, 46265.
34. Chemical Biology Research Initiative http://www.cbri.u-tokyo.ac.jp/.
35. Ringer, A. L.; Senenko, A.; Sherrill, C. D. *Protein Sci.* **2007**, *16*, 2216.
36. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. *J. Med. Chem.* **2006**, *49*, 534.
37. Suzuki, T.; Fujii, A.; Ohya, J.; Amano, Y.; Kitano, Y.; Abe, D.; Nakamura, H. *Cancer Sci.* **2007**, *98*, 1977.
38. Bionet, London, U.K. http://www.keyorganics.co.uk/.
39. Carter, T. A.; Wodicka, L. M.; Shah, N. P.; Velasco, A. M.; Fabian, M. A.; Treiber, D. K.; Milanov, Z. V.; Atteridge, C. E.; Biggs, W. H., 3rd; Edeen, P. T.; Floyd, M.; Ford, J. M.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Mehta, S. A.; Patel, H. K.; Pao, W.; Sawyers, C. L.; Varmus, H.; Zarrinkar, P. P.; Lockhart, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11011.
40. Kobayashi, S.; Ji, H.; Yuza, Y.; Meyerson, M.; Wong, K. K.; Tenen, D. G.; Halmos, B. *Cancer Res.* **2005**, *65*, 7096.
41. Peipp, M.; Schneider-Merck, T.; Dechant, M.; Beyer, T.; van Bueren, J. J.; Bleeker, W. K.; Parren, P. W.; van de Winkel, J. G.; Valerius, T. *J. Immunol.* **2008**, *180*, 4338.
42. Gilmer, T. M.; Cable, L.; Alligood, K.; Rusnak, D.; Spehar, G.; Gallagher, K. T.; Woldu, E.; Carter, H. L.; Truesdale, A. T.; Shewchuk, L.; Wood, E. R. *Cancer Res.* **2008**, *68*, 571.
43. Mitsudomi, T.; Kosaka, T.; Yatabe, Y. *Int. J. Clin. Oncol.* **2006**, *11*, 190.
44. Sequist, L. V. *Oncologist* **2007**, *12*, 5.
45. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
46. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
47. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
48. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
49. MOE (Molecular Operating Environment), version 2007.09; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2009.
50. Invitrogen, Karlsruhe, Germany http://www.invitrogen.com/.

51. Stratagene, La Jolla, CA http://www.stratagene.com/.
52. Qiagen Inc., Valencia, CA http://www.qiagen.com/.
53. GE Healthcare Ltd, Buckinghamshire, U.K. http://www.gelifesciences.com/.
54. Pharmeks Ltd., Moscow, Russia http://www.pharmeks.com/.
55. Maybridge Chemical Co., Ltd, Cornwall U.K. http://www.maybridge.com/.
56. Sigma–Aldrich Co., St. Louis, MO http://www.sigmaaldrich.com/.
57. ChemBridge Co., San Diego, CA http://www.chembridge.com/.
58. Enamine Ltd, Kiev, Ukraine http://www.enamine.net/.
59. Prestwick Chemical Inc., Illkirch, France http://www.prestwickchemical.com/.
60. InterBioScreen Ltd, Moscow, Russia http://www.ibscreen.com/.
61. Caliper Life Sciences Inc., Hopkinton, MA http://www.caliperls.com/.
62. Greiner, Frickenhausen, Germany http://www.greinerbioone.com/.
63. Calbiochem, San Diego, CA http://www.calbiochem.com/.
64. Carna Biosciences Inc., Kobe, Japan http://www.carnabio.com/.
65. Coster, Cambridge, MA http://www.corning.com/lifesciences.

Please cite this article in press as: Sato, T.; et al. *Bioorg. Med. Chem.* (2012), http://dx.doi.org/10.1016/j.bmc.2012.04.042