

Structural basis of mutant-selectivity and drug-resistance related to CO-1686

Xiao-E Yan^{1,2,*}, Su-Jie Zhu^{1,2,*}, Ling Liang^{1,2}, Peng Zhao^{1,2}, Hwan Geun Choi^{3,4} and Cai-Hong Yun^{1,2}

¹Institute of Systems Biomedicine, School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China

²Department of Biophysics, School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China

³Department of Biological Chemistry & Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

⁴Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

*These authors have contributed equally to this work

Correspondence to: Cai-Hong Yun, **email:** yunch@hsc.pku.edu.cn

Keywords: NSCLC, EGFR kinase, T790M, CO-1686, structural pharmacology

Received: March 31, 2017

Accepted: May 08, 2017

Published: June 21, 2017

Copyright: Yan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Non-small-cell lung cancers (NSCLCs) caused by activating mutations in the kinase domain of epidermal growth factor receptor (EGFR) initially respond to first-generation reversible drugs gefitinib and erlotinib. However, clinical efficacy is limited due to the development of drug-resistance that in more than half of the cases are driven by the secondary T790M mutation. CO-1686 is one of the third generation irreversible inhibitors that inhibits EGFR activating mutants, including those with concurrent T790M, while avoiding the off-target toxicity owing to inhibition of wild-type EGFR in treating EGFR mutation-positive NSCLCs. Despite the remarkable success, the experimentally determined structure of this agent in complex with EGFR T790M remains unknown. In this study, we determined crystal structures of EGFR T790M or L858R mutants covalently bound by CO-1686. Based on these structural data, we can explain why CO-1686 irreversibly inhibits EGFR and selectively prefers T790M, which may help improving this or similar compounds, and explain why EGFR L718Q and L844V mutations incur resistance to this agent.

INTRODUCTION

Lung cancer is the leading cause of cancer related deaths, which accounts for nearly one third of all cancer deaths worldwide [1]. Despite prolonged research and clinical prevention strategies, the 5-year survival rate of lung cancer is less than 20% in patients in the United States (<http://seer.cancer.gov>). The two major types of lung cancers are non-small-cell lung cancer (NSCLC) and small-cell lung cancer, which account for about 85% and 15% of all lung cancers, respectively [2].

Activating mutations in the kinase domain of epidermal growth factor receptor (EGFR) are one of the major causes of NSCLCs, among which the most frequently seen mutations are the single site mutation

leading to a leucine-to-arginine substitution at residue 858 (L858R) in exon 21, and the deletion mutation in exon 19 resulting in loss of the pentapeptide ELREA (delE746-A750). Patients harboring these typical activating mutations respond very well to therapy with the 4-anilinoquinazoline based reversible EGFR tyrosine kinase inhibitors (TKIs), such as gefitinib and erlotinib (the first-generation drugs) [3–6].

Unfortunately, despite dramatic initial response to the targeted therapy, EGFR mutation-positive patients usually suffer from the development of drug resistance and tumor progression after 9 to 14 months of treatment. In approximately 50-60% of the relapsed cases the drug-resistance are driven by a secondary point mutation in EGFR that leads to a threonine to methionine substitution

at residue 790, the gatekeeper residue (T790M) [7–9]. Previous structural and mechanistic studies showed that EGFR T790M enhances the affinity of the kinase for ATP, leading to the reduction of the efficacy of any ATP-competitive inhibitors. In consequence new agents to counter EGFR T790M drug-resistance mutation must overcome the reinforced ATP binding affinity conferred by T790M [10]. The second-generation irreversible EGFR TKIs such as neratinib (HKI-272) had been expected to inhibit T790M-mutant EGFR with the capacity of competing off ATP by covalent binding to EGFR. However, these TKIs were effective only in *in vitro* or preclinical studies but not in clinical trials, probably due to the potent inhibition of wild-type EGFR accounting for adverse toxic effects [11, 12].

To resolve this issue, novel EGFR TKIs have been developed to selectively inhibit EGFR activating mutations with concomitant T790M while sparing the wild-type EGFR. The third generation pyrimidine-based irreversible EGFR TKIs (Figure 1), such as WZ4002, AZD9291 (osimertinib) and CO-1686 (rociletinib), not only effectively inhibit EGFR T790M, but also are much less potent in inhibiting wild-type EGFR and other kinases, are therefore expected to reduce side effects compared to afatinib [13–15]. The recent studies showed that both AZD9291 and CO-1686 exhibited excellent clinical efficacy in NSCLC patients harboring EGFR T790M, with more than 50% response rates, meanwhile skin and gastrointestinal toxicities are also less than those typically observed for the first generation EGFR TKIs [16–18]. In order to elucidate the binding mode of the compound to EGFR, to understand the structural basis of its specificity toward the T790M mutation, and to learn the drug-resistance mechanism conferred by several newly identified CO-1686-resistant mutations, we conducted the structural pharmacological studies on EGFR T790M and L858R mutants with CO-1686.

RESULTS

Overall crystal structures of EGFR T790M or L858R in complex with CO-1686

CO-1686 is currently in phase I/II clinical trials in NSCLCs harboring EGFR activating mutations [14]. Despite the exciting safety and efficacy of CO-1686 in human clinical trials, the experimentally determined structure of this agent in complex with EGFR kinase remains unclear. Walter, A.O. *et al.* proposed a theoretical model of CO-1686 in complex with EGFR T790M based on the previously reported crystal structure of EGFR T790M/WZ4002 complex [13, 14]. However, it still remains doubtful if this theoretical model really illustrated the true structure of the T790M/CO-1686 complex. We therefore determined the complex crystal structures of EGFR T790M/CO-1686 and EGFR L858R/CO-1686 (Figure 2).

The overall structures of these two complexes are highly similar to each other. One EGFR protein molecule is observed in the asymmetric unit of both crystal structures. The binding modes of CO-1686 to EGFR T790M and L858R are essentially the same and EGFR mutant proteins adopt the “DFG-in/C-helix in” active conformation (Figure 2A). As expected, CO-1686 covalently binds to EGFR through Cys797 in both structures (Figure 2B, 2C). The solution exposed N-acetylpiperazine moiety of the compound is coplanar to the methoxybenzene moiety and it bends towards Leu718 and wraps its side-chain in the structures of T790M/CO-1686 and L858R/CO-1686, which would facilitate the binding of the compound to EGFR (Figure 3A). It was predicted to be perpendicular to the methoxybenzene moiety in the theoretical model [14], which may not be reasonable since this conformation would incur too close contact between the piperazine and the Leu718 side-chain.

Structural basis for CO-1686 specificity

In the T790M/CO-1686 and L858R/CO-1686 complex crystal structures, the anilinopyrimidine core of CO-1686 form two hydrogen bonds with Met793 amide and carbonyl in the hinge of the kinase (Figure 3A). The methoxyl substituent extends towards Leu792 side-chain. This methoxyl substituent in CO-1686 likely plays a key role in the superior selectivity of CO-1686 towards EGFR since other kinases with a cysteine residue equivalent to Cys797 in EGFR (such as Jak3 and the TEC-family kinases) usually harbor a bulkier residue (tyrosine and phenylalanine, respectively) at the position equivalent to Leu792 that would hinder the binding of the compound due to steric hindrance with the methoxyl.

In the structure of T790M/CO-1686, the trifluoromethyl (-CF₃) substituent attached to the pyrimidine ring contacts the mutant gatekeeper residue Met790 by hydrophobic interaction, which may be the only difference between the T790M/CO-1686 and L858R/CO-1686 structures. The hydrophobicity afforded by the T790M gatekeeper mutation is beneficial to the potency of this drug towards EGFR. A wild-type gatekeeper residue (Thr790) would not afford this beneficial interaction with the compound (such as that observed in the L858R/CO-1686 structure), which explains why this agent prefers binding to the T790M mutant.

Structural basis for drug-resistance conferred by L718Q and L844V

Despite the excellent efficacy of CO-1686 in clinic treatment, it is inevitable that a number of patients will eventually develop acquired drug resistance after long-term drug administration. Recently, the data showed that Cys797 was the most common site of secondary mutations (C797S and C797G) mediating resistance to WZ4002,

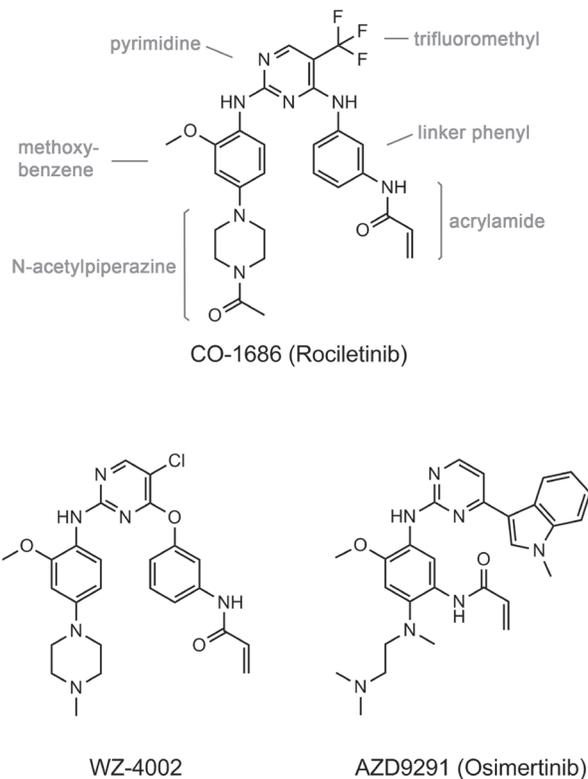


Figure 1: Chemical structures of the third generation EGFR TKIs discussed in this report.

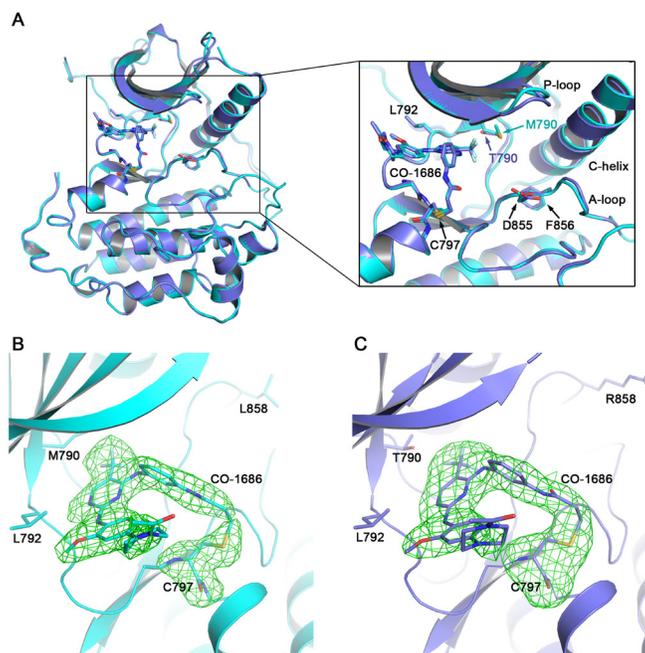


Figure 2: Overall EGFR/CO-1686 complex structure and covalent linkage between the compound and the kinase. (A) Superimposition of T790M/CO-1686 and L858R/CO-1686 complex crystal structures. (B) The $F_o - F_c$ omit map of CO-1686/Cys797 in the T790M/CO-1686 complex crystal structure. (C) The $F_o - F_c$ omit map of CO-1686/Cys797 in the L858R/CO-1686 complex crystal structure. The EGFR T790M and L858R mutant proteins are shown as cyan and slate cartoons, respectively. CO-1686 and the key amino acid residues discussed in this report are shown as sticks with their carbon atoms colored in the same way as the protein. The $F_o - F_c$ omit maps are contoured at 2.5σ and shown as green meshes.

CO-1686, and AZD9291; and secondary mutations L718Q and L844V appeared frequently in WZ4002 and CO-1686 resistant models [19].

In the determined EGFR T790M/CO-1686 and L858R/CO-1686 complex crystal structures, multiple hydrophobic interactions were found to facilitate the binding of the compound to EGFR. The pyrimidine core of the compound forms a hydrophobic interaction with Leu844 primarily; the methoxybenzene moiety forms a hydrophobic interaction with Leu718; while the linker phenyl ring carrying the acrylamide warhead of the compound into proximity to the thiol of Cys797 forms

hydrophobic interactions with both Leu718 and Val726 side-chains (Figure 3A). Therefore the hydrophobic residues Leu718, Val726 and Leu844 all play essential roles in the binding of the compound to EGFR. This observation can explain the recent finding that L718Q and L844V cause resistance to CO-1686 [19]. The L718Q mutation would destroy the beneficial hydrophobic interaction with the methoxybenzene of the compound due to steric hindrance and/or loss of hydrophobicity. The L844V mutation though retains hydrophobicity, the side-chain of Val is too short to make significant hydrophobic interaction with the pyrimidine core of the compound (Figure 3B).

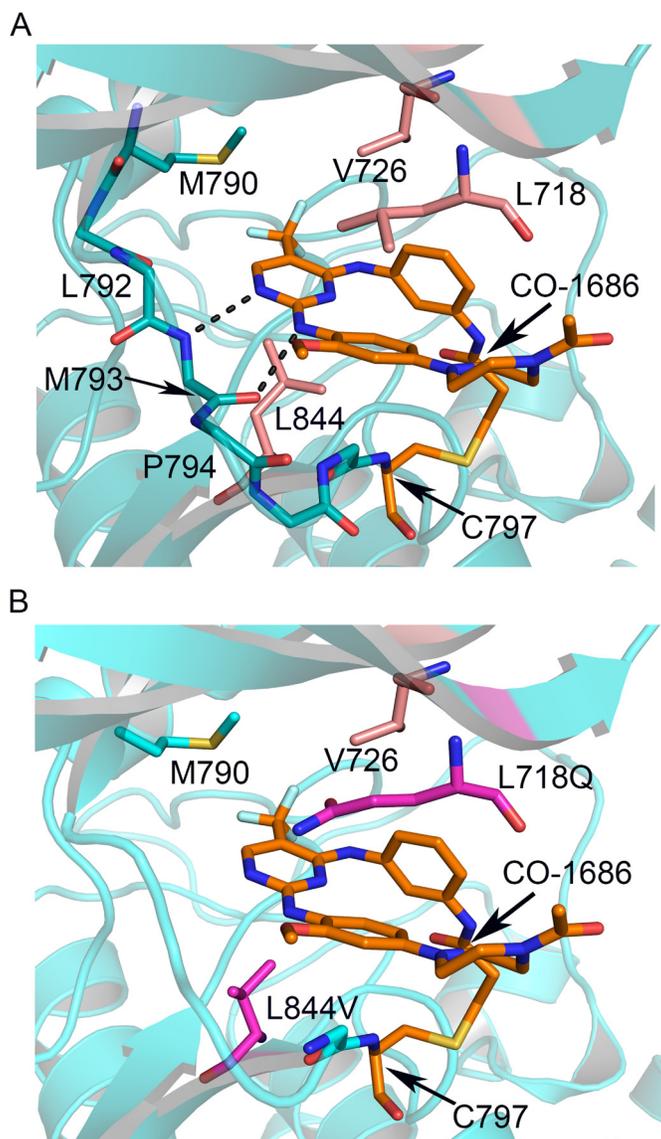


Figure 3: Interactions between CO-1686 and EGFR and the structural basis of drug-resistance conferred by L718Q and L844V. (A) Crystal structure of CO-1686 in complex with EGFR T790M. The EGFR kinase is shown as cartoons in cyan, and the bound CO-1686 is shown as sticks in orange. The amide and carbonyl atoms of Met793 interact with the aminopyrimidine of CO-1686 through hydrogen bonds shown by dashed lines. Residues contacting CO-1686 are shown as sticks. **(B)** Structural modeling illustrating the influences of EGFR L718Q and L844V mutations to the interactions with CO-1686. The L718Q mutation (carmine) is predicted to hinder the binding of CO-1686 owing to steric hindrance and/or abolishment of hydrophobic interaction, while the shorter side chain of L844V (carmine) will weaken the hydrophobic interaction with the pyrimidine core of CO-1686.

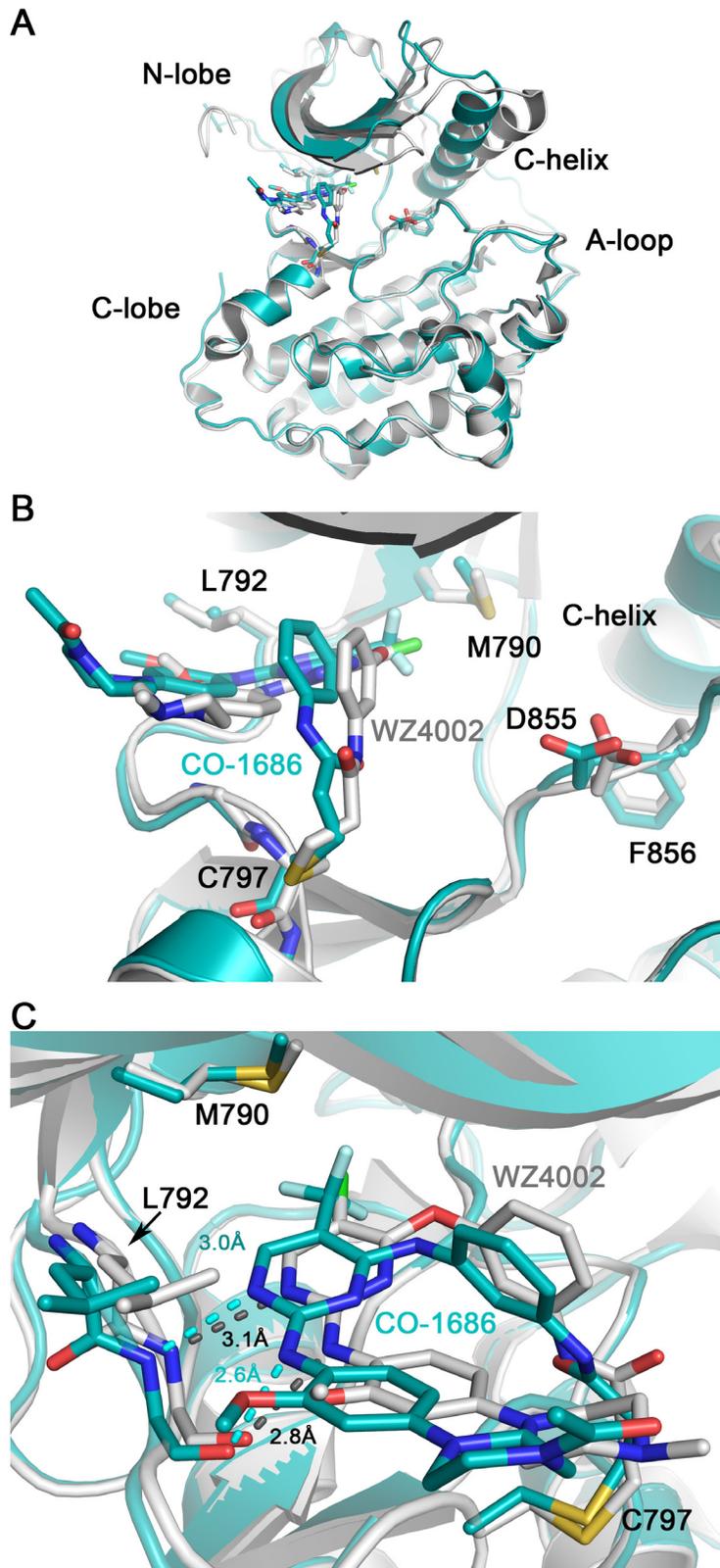


Figure 4: Comparison of T790M/CO-1686 and T790M/WZ4002 complex crystal structures. Superimposition of the T790M/CO-1686 and T790M/WZ4002 overall structures is shown in panel (A), while the side-view and top-view of the ATP binding pocket are shown in panel (B) and (C), respectively. The EGFR kinase in the CO-1686 and WZ4002 complex crystal structures are shown as cyan and gray cartoons, respectively. The compound and key residues are shown as sticks colored in the same way as the protein. The hydrogen bonds between the anilino-pyrimidine core and Met793 main-chain amide and carbonyl are indicated by dashed lines. The lengths of these hydrogen bonds are labeled.

Table 1: Data collection and refinement statistics of the EGFR/CO-1686 complex crystal structures

	CO-1686/EGFR T790M	CO-1686/EGFR L858R
Data collection		
Space group	<i>I</i> 23	<i>I</i> 23
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	145.5, 145.5, 145.5	143.5, 143.5, 143.5
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50-2.35(2.43-2.35)	50-2.70(2.91-2.70)
<i>R</i> _{p.i.m.} ^a (%)	3.3(54)	3.5(3.49)
<i>I</i> /σ <i>I</i>	27.6(2)	25.98(2.19)
Completeness (%)	100(100)	98.3(99.5)
Redundancy	16.7(16.3)	7.3(7.4)
Refinement		
Resolution (Å)	41.99-2.35	38.36-2.7
No. reflections	20894	13462
<i>R</i> _{work} / <i>R</i> _{free} ^b	0.217/0.250	0.208/0.252
No. atoms		
Protein	2456	2348
Water	108	19
<i>B</i> -factors		
Protein	40.86	75.72
Water	42.12	73.59
R.M.S. deviations		
Bond lengths (Å)	0.016	0.011
Bond angles (°)	1.307	1.122
Ramachandran plot		
Favored, %	98.67	97.89
Allowed, %	1.33	2.11
Disallowed, %	0	0

*Values in parentheses are for highest-resolution shell. One crystal was used for each data set.

^a $R_{p.i.m.} = \sum_h \left(\frac{1}{n_h - 1} \right) \sum_l |I_{hl} - \langle I_h \rangle| / \sum_h \sum_l \langle I_h \rangle$, where *I*_{hl} is the *l*th observation of reflection *h* and $\langle I_h \rangle$ is the average intensity for all observations *l* of reflection *h*. *R*_{p.i.m.} is a multiplicity-independent R factor to evaluate diffraction data quality [31].

^b The R factor for refinement is defined as:

$$R = \frac{\sum |F_{obs}| - |F_{cal}|}{\sum |F_{obs}|}$$

where *F*_{obs} and *F*_{cal} are observed and calculated structure factor amplitudes, respectively. *R*_{work} is calculated using reflections included in the refinement, while *R*_{free} is calculated using reflections excluded from the refinement.

Comparison between EGFR T790M/CO-1686 and T790M/WZ4002 complex crystal structures

WZ4002 is another third-generation mutant-selective covalent inhibitors targeting EGFR T790M [13]. Although the binding modes of CO-1686 and WZ4002 to EGFR are almost the same (Figure 4A), there remain subtle differences between them. Probably because the trifluoromethyl substituent on the pyrimidine ring contacting the mutant gatekeeper residue Met790 is much bulkier than the chlorine substituent in WZ4002, the CO-1686 core scaffold is “pushed out” a bit compared to WZ4002 (Figure 4B), and the distance of the bidentate hydrogen bonds between the anilinopyrimidine core and the ‘hinge’ residue Met793 are closer in the T790M/CO-1686 structure than those observed in the T790M/WZ4002 structure (Figure 4C). What’s more, the C-helix in T790M/CO-1686 structure rotates about 20° inwards when compared with the T790M/WZ4002 complex (Figure 4A). However, since the two complex crystal structures, T790M/CO-1686 and T790M/WZ4002, were made in different ways (co-crystallization versus soaking) and the crystallization conditions were different, these differences might simply be caused by the methodology and/or crystallization condition differences.

DISCUSSION

In this study we determined the crystal structures of EGFR T790M or L858R in complex with the third-generation drug CO-1686, which provide insights into why CO-1686 displays preference towards the EGFR T790M drug-resistant mutation while sparing the wild-type EGFR. In the T790M/CO-1686 complex structure, the anilinopyrimidine scaffold of CO-1686 fits well to the methionine gatekeeper. The hydrophobic interaction between the trifluoromethyl moiety and the Met790 side-chain enhances the affinity of CO-1686 towards T790M. The methoxyl group is necessary for the favorable selectivity of the compound to EGFR against other kinases bearing bulkier residues in the position equivalent to Leu792 in EGFR. Previously data showed that selectivity of CO-1686 to EGFR T790M/L858R versus wild-type EGFR is >12-fold and versus other kinases is 20- to 2000-fold [20]. *In vivo*, studies with human tumor xenografts also confirmed remarkable antitumor activity as well as selectivity over wild-type EGFR [14, 20]. Significantly, irreversible kinase inhibitors, such as WZ4002, rely on covalent bond formation for potent inhibition [13].

Our crystal structures help illustrate the reasons why L718Q and L844V are resistant to CO-1686. Since the chemical structure of CO-1686 is highly similar to that of WZ4002, and their binding modes to EGFR are highly similar to each other, the same drug-resistance mechanism should apply to WZ4002, too. Since Leu718 and Leu844 both directly contribute to CO-1686/WZ4002

binding through hydrophobic interactions, any mutations to eliminate these hydrophobic interactions would interfere with the drug binding. It is well known that mutations are a common mechanism of drug resistance to kinase inhibitors. The L844V mutation has been previously detected in a NSCLC patient [21]. Moreover, EGFR L718P, L718V, and L718M mutations have been described in clinic [22–24]. According to our model, all these mutations may weaken the binding affinity of CO-1686/WZ4002. However, since CO-1686 and WZ4002 are ATP-competitive inhibitors, whether these mutations incur resistance to these agents needs to be further investigated. For example, if these mutations also weaken ATP binding, they may not result in drug-resistance.

In summary, our crystallographic data provided insights into the structural basis of the selectivity of CO-1686 towards EGFR T790M, which may be helpful for future improvement of this or similar compounds. We observed the hydrophobic interactions between Leu718/Val726/Leu844 and CO-1686, which can explain why L718Q and L844V are resistant to CO-1686/WZ4002, and help predicting other potential drug-resistance mutations.

MATERIALS AND METHODS

Cloning and expression of EGFR 696-1022 mutants

Construct spanning residues 696-1022 of the human EGFR and harboring the L858R mutation was generated as previously described [25]. Construct of the human EGFR kinase domain (residues 696-1022) harboring the T790M mutation was generated from the cDNA of wild-type EGFR by site-directed mutagenesis and cloned into the pFastBac HTA vector (Invitrogen). A 6x-His tag followed by a TEV protease cleavage site was fused to the N-terminus of the EGFR protein to facilitate later purification of the protein. Transfection, virus generation and amplification were carried out in sf9 insect cells according to the official protocol of the Bac-to-Bac Baculovirus expression system (Invitrogen). The EGFR mutant kinases were then expressed in sf9 insect cells.

Purification of EGFR 696-1022 mutants

Cell pellets were suspended in lysis buffer (20 mM Tris-HCl, 150 mM NaCl, 5 mM KCl, 20 mM imidazole, 1 mM tris(2-carboxyethyl) phosphine hydrochloride, pH 8.0) supplemented with protease inhibitor mixture (Complete EDTA-free, Roche) and lysed by sonication. The lysate was centrifuged at 20,000 rpm for one hour at 4°C, then the supernatant was incubated with Ni-NTA Sepharose beads (GE Healthcare). The beads were washed with wash buffer (20 mM Tris-HCl, 500 mM NaCl, 1% Glycerol, 1 mM TCEP, 20 mM imidazole, pH 8.0) and then the protein was eluted with the same buffer

supplemented with 300 mM imidazole. The eluted protein was concentrated to 1 mL and incubated with His-tagged TEV for 4 hours at 4°C. After that, the uncleaved fusion protein and the His-tagged TEV were removed by diluting the sample with wash buffer to 20 mM imidazole and passing it through the Ni-NTA Sepharose beads (GE) for the second time. The flow-through containing only the untagged EGFR proteins were then concentrated to 0.5 mL. The mutant proteins were further purified by size-exclusion chromatography (Superdex 200) in the wash buffer and was concentrated to 15 mg/mL. Aliquots were made and flash frozen in liquid nitrogen and stored in -80°C refrigerator.

Crystallization

The apo-L858R crystals were obtained by hanging drop vapor diffusion using 0.5 µl of protein (4 mg/ml in 20 mM Tris, 500 mM NaCl, 1% glycerol, 1 mM TCEP) and 0.5 µl of reservoir solution containing 0.1 M HEPES pH 7.8, 40% PEG400, 0.15 M NaCl, 5 mM TCEP. Crystals of CO-1686/L858R were made by soaking the apo-L858R crystals for 4 hours at 20°C in the reservoir solution supplemented with 1 mM CO-1686. The EGFR T790M/CO-1686 complex crystals were prepared by co-crystallization. The EGFR T790M protein (15 mg/mL in 20 mM Tris, 500 mM NaCl, 1% glycerol, 5 mM TCEP) was pre-incubated with 1 mM CO-1686 on ice for 4 hours before setting up the crystallization tray. The initial crystals were obtained by sitting drop vapor diffusion against the reservoir solution of 0.0665 M HEPES pH 7.5, 1.1 M tri-sodium citrate at 20°C. The best crystals for data collection were grown by hanging drop vapor diffusion using 0.5 µl of protein and 0.5 µl of reservoir solution containing 0.07 M HEPES pH 7.5, 0.9 M tri-sodium citrate, 5 mM TCEP at 20°C. For data collection, all crystals were rapidly dipped in reservoir solution supplemented with 25% ethylene glycol and flash frozen in liquid nitrogen.

Structure determination

X-ray diffraction data were collected on beamline ID-19 at Advanced Photo Source (APS) at 100K at Argonne National Laboratory or beamline BL19U1 at Shanghai Synchrotron Radiation Facility (SSRF). The diffraction data were processed using HKL3000 [26]. The structures were determined by molecular replacement with Phaser [27] utilizing the publicly available EGFR L858R/ANP structure (PDB ID 2ITV) as the searching model for the L858R/CO-1686 and EGFR T790M structure (PDB ID 2JIT) as the searching model for T790M/CO-1686. Repeated rounds of manual refitting and crystallographic refinement were then performed using COOT [28] and Phenix [29]. Topology and parameter files for the inhibitor were generated using PRODRG [30]. Diffraction data

and refinement statistics are summarized in Table 1. The atomic coordinates and structure factors for T790M/CO-1686 and L858R/CO-1686 have been deposited in the Protein Data Bank with entry IDs 5XDK and 5XDL, respectively.

Author contributions

X.E.Y. and S.J.Z. prepared the complex crystals and solved the crystal structures. P.Z. carried out the molecular cloning and prepared the baculovirus. H.G.C synthesized the inhibitor CO-1686. X.E.Y., L.L. and C.H.Y. analyzed the data and wrote the manuscript. C.H.Y. instructed the project.

ACKNOWLEDGMENTS

We thank the staffs of APS beamline ID-19 and SSRF Beamline BL19U1 for technical assistance in diffraction data collection. This work was supported by National Natural Science Foundation of China (No. 31270769), the National Basic Research Program of China (973 Program, No. 2012CB917202) and the Ministry of Science and Technology of China (NCET-12-0013).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin.* 2017; 67:7-30.
2. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med.* 2008; 359:1367-1380.
3. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science.* 2004; 304:1497-1500.
4. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med.* 2004; 350:2129-2139.
5. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L, Mardis E, Kupfer D, Wilson R, et al. 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 U S A.* 2004; 101:13306-13311.

6. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer*. 2007; 7:169-181.
7. Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, Bergethon K, Shaw AT, Gettinger S, Cosper AK, Akhavanfard S, Heist RS, Temel J, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci Transl Med*. 2011; 3:75ra26.
8. Kosaka T, Yatabe Y, Endoh H, Yoshida K, Hida T, Tsuboi M, Tada H, Kuwano H, Mitsudomi T. Analysis of epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer and acquired resistance to gefitinib. *Clin Cancer Res*. 2006; 12:5764-5769.
9. Balak MN, Gong Y, Riely GJ, Somwar R, Li AR, Zakowski MF, Chiang A, Yang G, Ouerfelli O, Kris MG, Ladanyi M, Miller VA, Pao W. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. *Clin Cancer Res*. 2006; 12:6494-6501.
10. Yun CH, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong KK, Meyerson M, Eck MJ. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci U S A*. 2008; 105:2070-2075.
11. Sequist LV, Besse B, Lynch TJ, Miller VA, Wong KK, Gitlitz B, Eaton K, Zacharchuk C, Freyman A, Powell C, Ananthakrishnan R, Quinn S, Soria JC. Neratinib, an irreversible pan-ErbB receptor tyrosine kinase inhibitor: results of a phase II trial in patients with advanced non-small-cell lung cancer. *J Clin Oncol*. 2010; 28:3076-3083.
12. Miller VA, Hirsh V, Cadranel J, Chen YM, Park K, Kim SW, Zhou C, Su WC, Wang M, Sun Y, Heo DS, Crino L, Tan EH, et al. Afatinib versus placebo for patients with advanced, metastatic non-small-cell lung cancer after failure of erlotinib, gefitinib, or both, and one or two lines of chemotherapy (LUX-Lung 1): a phase 2b/3 randomised trial. *Lancet Oncol*. 2012; 13:528-538.
13. Zhou W, Ercan D, Chen L, Yun CH, Li D, Capelletti M, Cortot AB, Chirieac L, Jacob RE, Padera R, Engen JR, Wong KK, Eck MJ, et al. Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature*. 2009; 462:1070-1074.
14. Walter AO, Sjin RT, Haringsma HJ, Ohashi K, Sun J, Lee K, Dubrovskiy A, Labenski M, Zhu Z, Wang Z, Sheets M, St Martin T, Karp R, et al. Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC. *Cancer Discov*. 2013; 3:1404-1415.
15. Cross DA, Ashton SE, Ghiorghiu S, Eberlein C, Nebhan CA, Spitzler PJ, Orme JP, Finlay MR, Ward RA, Mellor MJ, Hughes G, Rahi A, Jacobs VN, et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov*. 2014; 4:1046-1061.
16. Janne PA, Yang JC, Kim DW, Planchard D, Ohe Y, Ramalingam SS, Ahn MJ, Kim SW, Su WC, Horn L, Haggstrom D, Felip E, Kim JH, et al. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *N Engl J Med*. 2015; 372:1689-1699.
17. Sequist LV, Soria JC, Goldman JW, Wakelee HA, Gadgeel SM, Varga A, Papadimitrakopoulou V, Solomon BJ, Oxnard GR, Dziadziuszko R, Aisner DL, Doebele RC, Galasso C, et al. Rocicetinib in EGFR-mutated non-small-cell lung cancer. *N Engl J Med*. 2015; 372:1700-1709.
18. Niederst MJ, Hu H, Mulvey HE, Lockerman EL, Garcia AR, Piotrowska Z, Sequist LV, Engelman JA. The allelic context of the C797S mutation acquired upon treatment with third-generation EGFR inhibitors impacts sensitivity to subsequent treatment strategies. *Clin Cancer Res*. 2015; 21:3924-3933.
19. Ercan D, Choi HG, Yun CH, Capelletti M, Xie T, Eck MJ, Gray NS, Janne PA. EGFR mutations and resistance to Irreversible pyrimidine based EGFR inhibitors. *Clin Cancer Res*. 2015; 21:3913-3923.
20. Tjin Tham Sjin R, Lee K, Walter AO, Dubrovskiy A, Sheets M, Martin TS, Labenski MT, Zhu Z, Tester R, Karp R, Medikonda A, Chaturvedi P, Ren Y, et al. *In vitro* and *in vivo* characterization of irreversible mutant-selective EGFR inhibitors that are wild-type sparing. *Mol Cancer Ther*. 2014; 13:1468-1479.
21. Penzel R, Sers C, Chen Y, Lehmann-Muhlenhoff U, Merkelbach-Bruse S, Jung A, Kirchner T, Buttner R, Kreipe HH, Petersen I, Dietel M, Schirmacher P. EGFR mutation detection in NSCLC--assessment of diagnostic application and recommendations of the German Panel for Mutation Testing in NSCLC. *Virchows Arch*. 2011; 458:95-98.
22. Taron M, Ichinose Y, Rosell R, Mok T, Massuti B, Zamora L, Mate JL, Manegold C, Ono M, Queralt C, Jahan T, Sanchez JJ, Sanchez-Ronco M, et al. Activating mutations in the tyrosine kinase domain of the epidermal growth factor receptor are associated with improved survival in gefitinib-treated chemorefractory lung adenocarcinomas. *Clin Cancer Res*. 2005; 11:5878-5885.
23. Locatelli-Sanchez M, Couraud S, Arpin D, Riou R, Bringuier PP, Souquet PJ. Routine EGFR molecular analysis in non-small-cell lung cancer patients is feasible: exons 18-21 sequencing results of 753 patients and subsequent clinical outcomes. *Lung*. 2013; 191:491-499.
24. Liu Y, Wu BQ, Zhong HH, Hui P, Fang WG. Screening for EGFR and KRAS mutations in non-small cell lung carcinomas using DNA extraction by hydrothermal pressure coupled with PCR-based direct sequencing. *Int J Clin Exp Pathol*. 2013; 6:1880-1889.
25. Yun CH, Boggon TJ, Li Y, Woo MS, Greulich H, Meyerson M, Eck MJ. Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation

- and insights into differential inhibitor sensitivity. *Cancer Cell*. 2007; 11:217-227.
26. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr*. 2006; 62:859-866.
 27. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40:658-674.
 28. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:486-501.
 29. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:213-221.
 30. Schuttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:1355-1363.
 31. Evans P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr*. 2006; 62:72-82.