

Revealing acquired resistance mechanisms of kinase-targeted drugs using an on-the-fly, function-site interaction fingerprint approach

Zheng Zhao¹ and Philip E. Bourne^{*,1,2}

1. Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia 22904, United States of America

2. School of Data Science, University of Virginia, Charlottesville, Virginia 22904, United States of America

***Corresponding author**

Philip E. Bourne: peb6a@virginia.edu

Abstract

Although kinase-targeted drugs have achieved significant clinical success, they are frequently subject to the limitations of drug resistance, which has become a primary vulnerability to targeted drug therapy. Therefore, deciphering resistance mechanisms is an important step in designing more efficacious, anti-resistant, drugs. Here we studied two FDA-approved kinase drugs: Crizotinib and Ceritinib, which are first- and second-generation anaplastic lymphoma kinase (ALK) targeted inhibitors, to unravel drug-resistance mechanisms. We used an on-the-fly, function-site interaction fingerprint (on-the-fly Fs-IFP) approach by combining binding free energy surface calculations with the Fs-IFPs. Establishing the potentials of mean force and monitoring the atomic-scale protein-ligand interactions, before and after the L1196M-induced drug resistance, revealed insights into drug-resistance/anti-resistant mechanisms. Crizotinib prefers to bind the wild type ALK kinase domain, whereas Ceritinib binds more favorably to the mutated ALK kinase domain, in agreement with experimental results. We determined that ALK kinase-drug interactions in the region of the front pocket are associated with drug resistance. Additionally, we find that the L1196M mutation does not simply alter the binding modes of inhibitors, but also affects the flexibility of the entire ALK kinase domain. Our work provides an understanding of the mechanisms of ALK drug resistance, confirms the usefulness of the on-the-fly Fs-IFP approach and provides a practical paradigm to study drug-resistance mechanisms in prospective drug discovery.

Introduction

Targeted drug therapy has become a major cancer treatment beyond surgery, radiation therapy, chemotherapy and immunotherapy¹. Validating primary targets and developing targeted drugs has received increasing attention². In concert, over the last 20 years, the kinase family has been recognized as important drug targets³⁻⁴. As of August 2018, 42 kinase-targeted drugs have been approved by the U.S. Food and Drug Administration (FDA)⁵ and have revolutionized clinical therapy against multiple diseases, especially different cancer types, such as non-small cell lung cancer (NSCLC), melanoma and leukemia^{1, 6}. However, the efficacy of drugs is frequently subject to the limitation of drug-acquired resistance. For example, Crizotinib has been frequently found to be ineffective for the majority of patients after one to two years' treatment against ALK-positive NSCLC, due to the acquired mutations at the binding site⁷. As such, patients have a need for next-generation drugs effective against acquired drug resistance^{5, 8}. Thus, deciphering the mechanisms of acquired resistance is a step forward in designing novel and efficacious anti-resistant drugs⁹.

Exploring resistance mechanisms has been challenging due to the diversity of specific drug-binding mechanisms. For example, Yun et al.¹⁰ concluded that the resistance against Gefitinib and Erlotinib, which are two FDA-approved kinase drugs used to treat patients with EGFR-overexpression-induced NSCLC, is caused by the gatekeeper T790M mutation; this mutation increased the binding affinity of ATP in EGFR kinase. In another example, the resistance against Dasatinib, a tyrosine-kinase inhibitor employed to treat people with Ph+ chronic myeloid leukemia and acute lymphoblastic leukemia¹¹, is caused by weakening the drug's affinity to the binding pocket due to a T338M mutation in the cSrc and Abl kinase family¹². Worse still, the resistance mechanisms remain undetermined for most targeted drugs⁶. For instance, Crizotinib is the first FDA-approved drug against ALK-positive NSCLC, and its drug resistance frequently

occurs in patients, along with the development of mutations to the binding site¹³⁻¹⁴. On one hand, researchers have suggested that the mechanism involves acquired mutations blocking the binding of inhibitors, such as where the L1196M mutation confers ALK kinase high-level resistance to Crizotinib¹⁵ through decreased binding affinity¹⁶. In contradiction, the high-resolution crystal structures of the wild-type and L1196M ALK-Crizotinib complexes (PDB ids: 2xp2 and 2yfx) suggest the binding modes are similar¹⁷. Thus, it is essential to develop efficient strategies to reveal resistance mechanisms when designing future drugs. To this end, we here provide an effective method called the on-the-fly, function-site interaction fingerprint (Fs-IFP) approach to explore kinase drug-resistance mechanisms.

The on-the-fly Fs-IFP approach we propose is formed by combining binding free-energy calculations with the Fs-IFP encoding. The binding free-energy calculation is an important measure in understanding the ligand-binding recognition processes and differs from the binding affinity for ligands¹⁸. Theoretically, there are different methods to calculate binding free energy, such as thermodynamic integration, free-energy perturbation and probability distributions and histograms¹⁹. Here, we used a sophisticated, improved sampling strategy (umbrella sampling) with a WHAM analysis to obtain the binding free-energy surface (FES), which has been successfully applied in different complex systems²⁰. More importantly, in our proposed scheme, we monitored the binding process by using an Fs-IFP encoding strategy. This strategy has practical applications in analyzing a variety of binding pockets²¹⁻²⁴ to reveal details of binding, notably the differences before and after mutation(s).

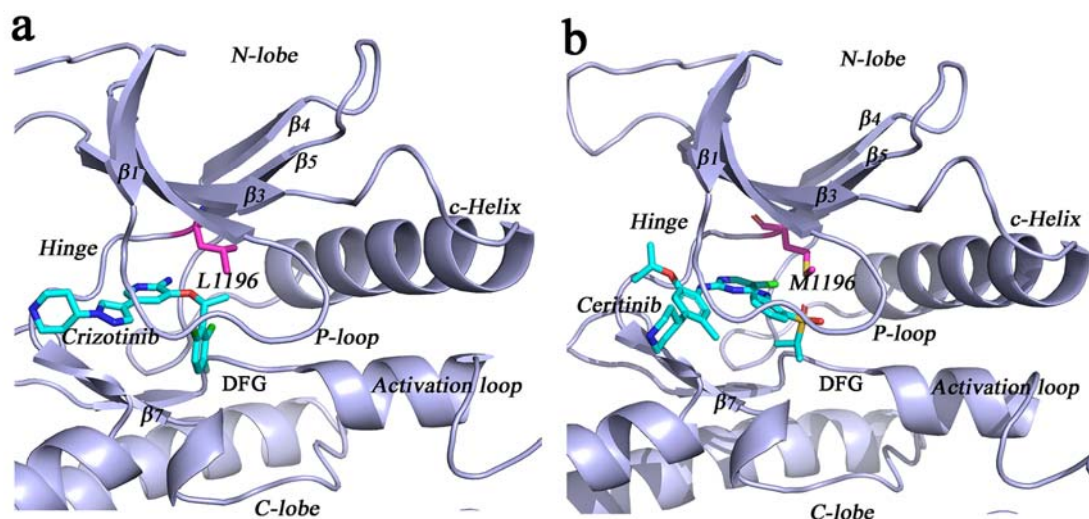


Figure 1. Representative ALK-drug interactions. (a) Crizotinib-bound ALK complex. (b) Ceritinib-bound ALK complex. Residue 1196 is the gatekeeper, and the L1196M mutation is shown when binding Ceritinib.

In this work, we explore the resistance/anti-resistant mechanisms of two drugs (Crizotinib and Ceritinib) against ALK in the case of wild type and L1196M mutation. Ceritinib is a second-generation ALK inhibitor which efficiently overcomes L1196M-induced drug resistance. Ceritinib has shown stronger inhibition, not only for the wild-type ALK, but also for the L1196M mutant²⁵. Using the on-the-fly Fs-IFP approach, we obtained the dynamic binding details of the two drugs by exploring the binding free energy surfaces (FESs). We revealed the details of a resistance mechanism for Crizotinib as well the anti-resistant mechanism of Ceritinib. This includes details of the flexibility of the apo and holo ALK kinase domains. The hope is that these insights contribute to an understanding of the drug resistance mechanisms and to next-generation drug design. The on-the-fly pipeline for studying drug resistance mechanisms can be applied more generally beyond ALK kinase.

Results

Drug binding profiles

We determined the binding processes of Crizotinib and Ceritinib by exploring the corresponding binding free-energy surfaces (Figure 2). We note that every binding free-energy profile for different drug-binding systems is an energy-decreasing procedure from the dissociated state to the bound state, along the predefined reaction coordinates (Supplemental Figure S1); that means the drug-binding process is spontaneous and thermodynamically favorable. In the bound state, the potentials of mean force (PMF) have a global minimum state, and the relative PMF well depths are -9.8 ± 0.7 , -11.0 ± 0.5 , -11.3 ± 0.7 , and -13.6 ± 0.8 kcal.mol⁻¹ for wild/mutated Crizotinib-bound and Ceritinib-bound systems, respectively, in agreement with experimental results^{16, 26}. Based on the comparisons of wild/mutated PMFs (Figure 2a), Crizotinib prefers to bind to the wild ALK kinase binding site rather than the L1196M-mutated binding site. Comparatively, Ceritinib binds more favorable to a L1196M-mutated ALK kinase (Figure 2b). Moreover, Ceritinib has slightly lower binding free-energy potential, which means that Ceritinib binds not only to wild-type but also to mutated ALK binding sites more closely than Crizotinib. Given this information, we explore the resistance mechanisms of Crizotinib, and the anti-resistant mechanisms of Ceritinib at the atomic level using the Fs-IFP encoding method.

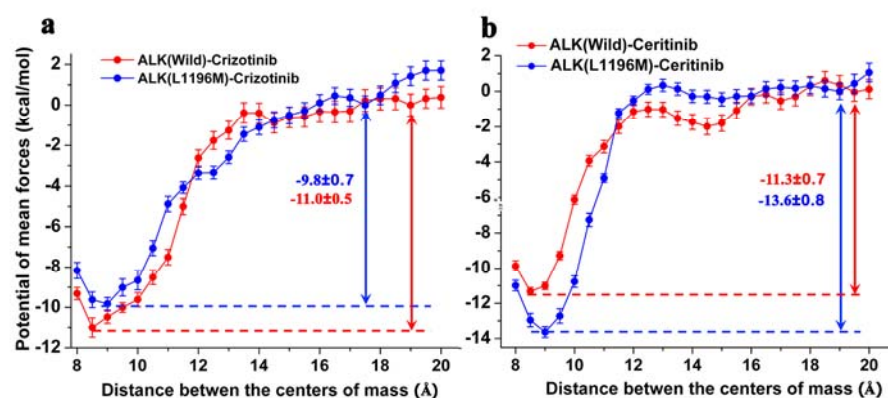


Figure 2. Potential of mean force profiles for wild-type or mutated drug-binding processes along with the predefined reaction coordinates. (a) Crizotinib binding profile; (b) Ceritinib binding profile.

Drug-binding features

The Fs-IFPs for the structures of the drug-bound complexes illustrate the binding characteristics and differences before and after mutation. We extracted Fs-IFPs for conformations of the sampling windows with the lowest PMF (i.e., the bound state). Significant differences are presented (Figure 3).

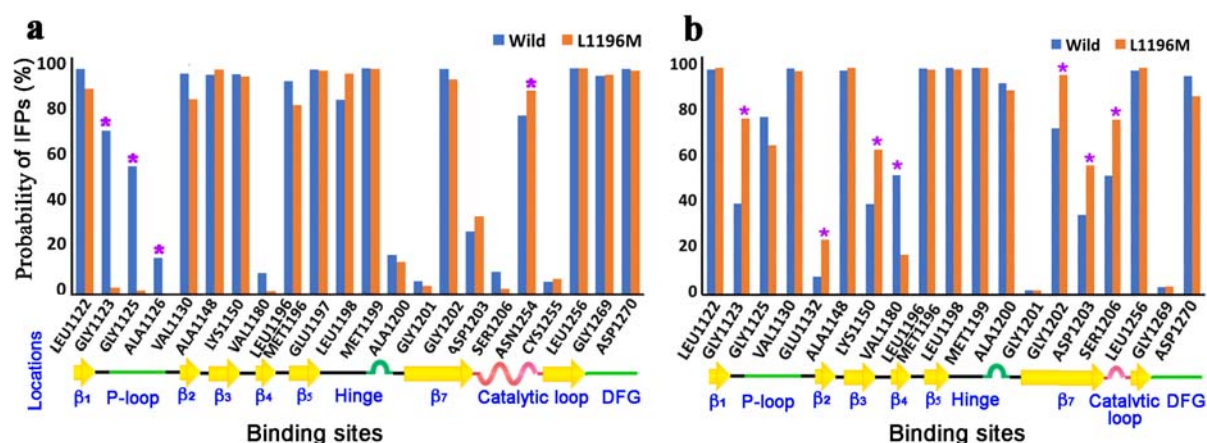


Figure 3. Comparisons of the Fs-IFPs between the wild-type and the L1196M mutated systems in the drug-bound states. (a) Crizotinib-bound systems. (b) Ceritinib-bound systems.

In total, the features of Crizotinib binding to the wild or mutated ALK kinase are very similar (Figure 3a). At the regions of β_1 , β_2 , β_3 , β_5 , Hinge, β_7 and DFG, there are conserved ALK-Crizotinib interactions (these interactions occur in more than 90% of conformations), which show no significant changes before or after mutation. Noteworthy is the lack of significant change in the hydrophobic interactions between residue 1196 and Crizotinib regardless of the mutation. Further, at the P-loop region consisting of G1123, G1125 and A1126, the interactions are significantly weakened in the mutated systems (Figure 3a). These decreased interactions will induce the P-loop to become more flexible and more sensitive to the selectivity of Crizotinib, thereby inferring drug resistance.

Similarly, the conserved ALK-Ceritinib interactions also exist in the regions of β_1 , β_3 , β_5 , Hinge, Catalytic loop, and DFG, regardless of the L1196M mutation (Figure 3b). However, in the mutated case, there are multiple residues including G1123, E1132, K1150, G1202, D1203, and S1206, presenting more conserved interactions than in wild type. These more conserved interactions contribute to the binding affinity of Ceritinib in the mutated system as shown in the corresponding free-energy profile (Figure 2b). It is worth noting that G1202, D1203, and S1206 are located in the region of β_7 , which constitutes the front pocket²¹. Meanwhile, G1123 is located in the region of the P-loop; the P-loop also constitutes part of the front pocket. Thus, these more conserved interactions at the front pocket contribute to overcome drug resistance. Compared to the change in ALK-Crizotinib interaction characteristics before and after mutation, the front pocket, including P-loop and β_7 , plays the key role in inducing or

overcoming the acquired L1196M drug resistance. This suggests an anti-resistant drug should be designed with more attention to achieving conserved interactions in the region of the front pocket.

Hydrophobic interactions at the gatekeeper

The mutation of the gatekeeper (residue 1196) is often attributed to drug resistance¹². Here we carefully checked the hydrophobic interactions between the gatekeeper and the ligand before and after mutation in the bound state. We present the distribution of the hydrophobic interactions between the gatekeeper and the ligand (Figure 4). For the wild-type ALK-Crizotinib system, the maximum probability distribution (MPD) is 3.9 Å, where there are stronger hydrophobic interactions compared to the L1196M-mutated system, which has an MPD of 4.2 Å (Figure 4a). In contrast, in the Ceritinib-bound case, there is a stronger hydrophobic interaction in the L1196M system, with an MPD of 3.55 Å, than in the wild-type with an MPD of 3.7 Å. These trends in hydrophobic interaction are in agreement with the drug-binding affinities implying that the hydrophobic interactions at the gatekeeper also play a role in ligand binding. Thus, along with the Fs-IFP features, the hydrophobic interactions result in slight drug resistance to Crizotinib and contribute to overcome the drug resistance to Ceritinib.

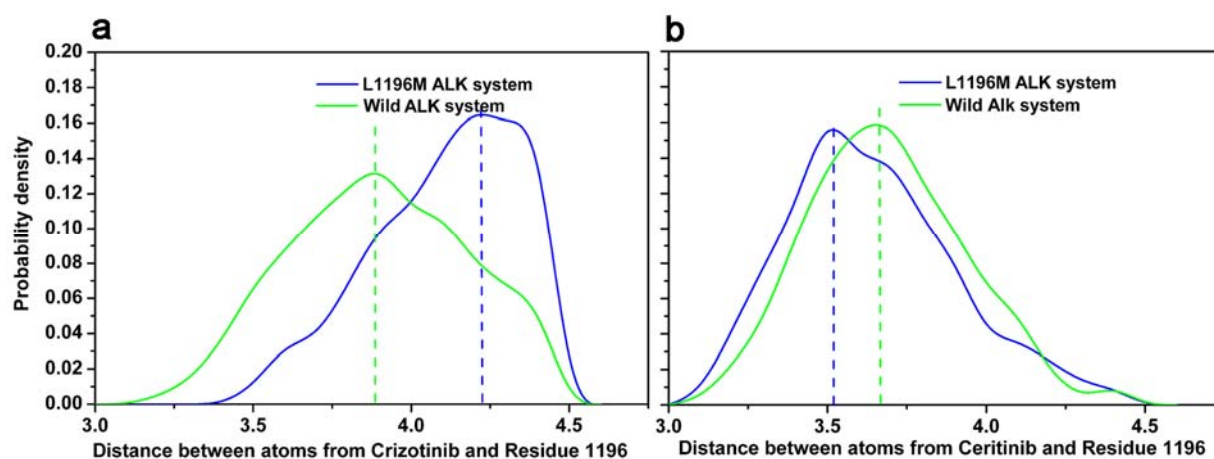


Figure 4. Probability distribution of interatomic distances between Residue 1196 and the corresponding ligand. (a) in the Crizotinib-bound system. (b) in the Ceritinib-bound system.

Collective motions of the ALK kinase domain

The collective change in conformation of target proteins always leads to drug-resistance mechanisms²⁷⁻²⁸. We analyzed the trajectories of both wild-type and mutant apo ALK kinase conformations using principal component analysis (PCA) (Figure 5a). Overall the dynamics trajectories show similar features and magnitudes in the core structures of both wild-type and mutated apo AKL kinase domains. However, in regions of flexibility, the collective motions of the apo AKL kinase domain indicate a significant difference before and after the L1196M mutation (Figure 5a). The catalytic loop presents large flexibility, with a strong dynamic magnitude in the wild type, whereas the N-lobe has increased flexibility in the mutated kinase domain (Figure 5a). The difference in mutation-induced flexibility would, therefore, seem a possible contributor to drug resistance.

Considering the ALK-drug complexes with the lowest PMF, in the wild-type ALK-Crizotinib complex (Figure 5b), Crizotinib is located at the binding pocket, competes with ATP and has a stable binding interaction with ALK. Compared to apo ALK, the essential dynamics of wild-type Crizotinib-bound ALK kinase shows a similar magnitude but moves in a different direction. However, in the mutated ALK-Crizotinib system, the conformation space appears to be similar to that in the L1196M apo ALK system (Figure 5a-b, L1196M). Clearly, the essential motion is different between the wild-type and mutated ALK-Crizotinib complexes, notably the flexibility of the P-loop. In the wild type, the P-loop is rigid, and interacts with Crizotinib, but appears to be

more flexible in the mutated case (Figure 5b). That behavior agrees with the observed differences in the interaction fingerprints (Figure 3).

To evaluate the population propensity of Ceritinib-bound ALK systems, we also analyzed the conformations from the trajectories of the ALK-Ceritinib complexes (Figure 5c). Here, conformational motions are distinctly different from the Crizotinib-bound wild-type system, but similar to the apo kinase, regardless of whether they are wild-type or mutated. There is a slight propensity for flexibility at the P-loop region in L1196M Ceritinib-bound complexes, which is in accordance with the dynamic modes of the apo ALK kinase, which is also more flexible at the N-lobe, due to mutations.

In summary, the PCA analysis supports the difference in Fs-IFP's before and after mutations, and demonstrates the dynamic motions of the apo and holo kinase domains. The Crizotinib-bound kinase domain has a different dynamic mode from the apo ALK conformations or the Ceritinib-bound ALK complexes.

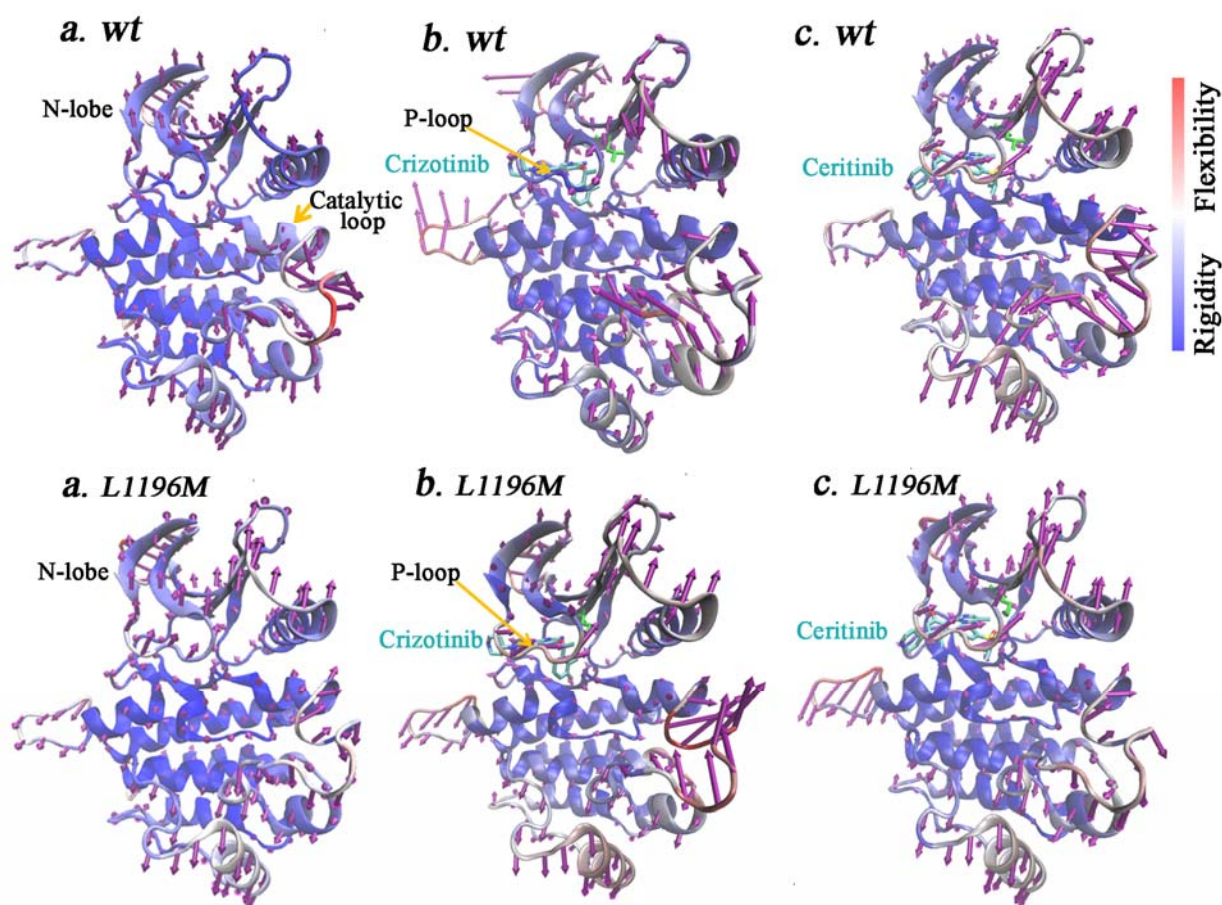


Figure 5. Essential dynamics analysis using PCA. (a) on the apo ALK kinase domain system; (b) on the Crizotinib-bound ALK system; (c) on the Ceritinib-bound ALK system. Arrows (purple) on every structure show the first principal direction for each residue with significant perturbation. The length of each arrow shows the magnitude of the motion for each residue. Different backbone colors show structural flexibility (red) or rigidity (blue).

Conclusions

We present an on-the-fly Fs-IFP approach by combining free-energy calculations and Fs-IFP encoding to reveal the resistance mechanism for two kinase-targeted drugs. Our calculations support experimental evidence that Crizotinib induces drug resistance, whereas Ceritinib can

overcome L1196M-induced drug resistance. Fs-IFP encoding reveals the details of the binding mode before and after mutation, which is useful in differentiating subtle changes in the target-drug interaction. This finding has broader implications across the kinase family given the high similarity of ATP binding. We also studied the dynamic modes for the ALK kinase domain and the drug-bound kinase complexes. The wild-type Crizotinib-bound ALK complex displays a different motion than the wild-type or mutated apo ALK kinase domain. In comparison, the Ceritinib-bound complex has a dynamic mode similar to apo ALK kinase, which is favorable for overcoming drug resistance.

Our computational analysis provides detailed insights into experimentally observed first-generation drug resistance and second-generation drug anti-resistance. It is challenging to highlight the key factors responsible for ALK-targeted drug resistance because there are conserved interactions between ALK kinase and the drug (both Crizotinib and Ceritinib), especially, in the regions of the Hinge and the gatekeeper. Using an on-the fly Fs-IFP scheme, we not only present the binding characteristics of every ligand at the binding sites but also exhibit the specific change in Fs-IFPs after mutation. The difference in Fs-IFPs shows that the ALK kinase-drug interactions in the region of the front pocket are related to drug resistance. Overall this methodology paves the way for designing next-generation anti-resistant drugs.

Methods

1. The on-the-fly function-site interaction fingerprint (on-the-fly Fs-IFP) approach

In this work, we provided an on-the-fly Fs-IFP approach by combining the binding free energy surface calculation and the function-site interaction fingerprint method (Fs-IFP).

1.1 Calculating the binding free energy surface

The binding free energy surface is calculated by using umbrella sampling (US) with the weighted histogram analysis method (WHAM). The US method represents significant progress in enhancing the sampling of the free energy surface²⁹. The theoretical basis has been well described²⁹⁻³⁰ and not repeated here. Following US, using WHAM for free energy calculations has emerged as an efficient scheme²⁰, especially as the number of dimension of the reaction coordinates and the complexity of free energy surface increases³¹. A brief overview of the WHAM extension is provided in the supporting information.

1.2 Encoding function-site interaction fingerprints (Fs-IFPs)

Fs-IFP is a method to determine protein-ligand interaction characteristics at the functional site and to do so on a proteome-wide scale as detailed in previous applications²¹⁻²³. The protein–ligand interfacial interaction is described using 1D fingerprints, which can discriminate between the large number of potential ligand binding modes.

The method has been described in detail elsewhere²¹⁻²³. Briefly, we first aligned all of binding sites with a sequence-independent alignment tool SMAP using default parameters³². Then, for every conformation, the protein-ligand interactions of every involved residue are encoded as a 7-bit fingerprint using the predefined geometric rules³² for seven types of interactions: (1) van der Waals; (2) aromatic face to face; (3) aromatic edge to face; (4) hydrogen bond (protein as hydrogen bond donor); (5) hydrogen bond (protein as hydrogen bond acceptor); (6) electrostatic interaction (protein positively charged); and (7) electrostatic interaction³³. Output includes the

name and index of the interacting atoms corresponding to the fingerprints, in other words, which atoms are contributing to specific interaction features.

In this way, we systematically obtained the interaction features by encoding the Fs-IFP of every snapshot through every umbrella sampling trajectory.

2. Computational details

Initial conformations of the wild-type and the L1196M mutated ALK-Crizotinib complex were taken from the Protein Data Bank³⁴ (PDB IDs: 2xp2³⁵ and 2yfx¹⁷, respectively). Moreover, the two PDB structures after removing the corresponding ligand were used as the initial structures for the wild-type and mutant apo ALK kinase systems, respectively. The missing loops were built using the Modeller software³⁶ using the corresponding ALK amino acid sequence³⁷. Similarly, the wildtype ALK-Ceritinib-binding complex was taken from the PDB (PDB ID: 4mkc). Because no L1196M ALK-Ceritinib-binding complex was available, we obtained the complex by docking Ceritinib into an ALK L1196M conformation (PDB ID: 2yfx) using Surflex v4.1³⁸. The force field parameters and the topology files of Crizotinib and Ceritinib were assignment using the ParamChem server³⁹⁻⁴⁰. The protonation states of the charged residues of all six systems were determined assuming a constant pH of 7.0. Then all the systems were solved in a rectangular water box with an 18Å buffer from any solute atom, respectively. Counter ions were added to ensure an ion concentration of 0.20 M and electroneutrality. The CHARMM36 force field⁴¹, CHARMM general force field⁴² and TIP3P force field⁴³ were used for the ALK, Crizotinib and Ceritinib and water molecules, respectively. All systems were optimized using an ACEMD MD setup protocol⁴⁴: 2ps minimization, 100ps for NVT, 1ns for NPT with heavy-atom constraints and 1ns for NPT without any constraints. Subsequently, 150ns MD simulations were

performed to equilibrate every system. In all MD simulations, all bonds were constrained using SHAKE and the integration time step was 4 fs. The temperature bath used the Langevin method, and 1 atm pressure was maintained using the Berendsen method⁴⁵. All MD simulations and later umbrella sampling were carried out using the ACEMD software⁴⁴.

The free energy simulation of ALK-drug system was carried out in a serial of discrete windows. For every ALK-drug system, we used 25 windows following the bias potential with the distance of centers of mass as the predefined reaction coordinates (supplemental information). In every window, 160ns MD simulations were carried out and the last 100ns MD trajectories were analyzed to obtain PMFs and encode the Fs-IFPs.

For the apo ALK kinase, a 1200ns MD simulation was carried out and the last 1000ns trajectory used as the essential dynamics analysis. All the essential dynamics analysis was performed using PCA and visualized using Normal mode Wizard in VMD⁴⁶.

The Fs-IFP encoding was run using IChem software⁴⁷ with the similarity of the pairwise Fs-IFPs calculated using the Tanimoto coefficient (TC)⁴⁸.

Acknowledgements

We thank Prof. Ajay N. Jain for providing the Surflex v4.1. We thank Prof. Haiyan Liu for his guidance in the FES analysis using WHAM. This work was partly supported by the University of Virginia (PEB).

References

1. Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; Overington, J. P., A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **2017**, *16* (1), 19-34.
2. Behan, F. M.; Iorio, F.; Picco, G.; Goncalves, E.; Beaver, C. M.; Migliardi, G.; Santos, R.; Rao, Y.; Sassi, F.; Pinnelli, M.; Ansari, R.; Harper, S.; Jackson, D. A.; McRae, R.; Pooley, R.; Wilkinson, P.; van der Meer, D.; Dow, D.; Buser-Doepner, C.; Bertotti, A.; Trusolino, L.; Stronach, E. A.; Saez-Rodriguez, J.; Yusa, K.; Garnett, M. J., Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* **2019**, *568* (7753), 511-516.
3. Ferguson, F. M.; Gray, N. S., Kinase inhibitors: the road ahead. *Nat. Rev. Drug Discov.* **2018**.
4. Knapp, S.; Sundstrom, M., Recently targeted kinases and their inhibitors-the path to clinical trials. *Curr. Opin. Pharmacol.* **2014**, *17*, 58-63.
5. Bhullar, K. S.; Lagaron, N. O.; McGowan, E. M.; Parmar, I.; Jha, A.; Hubbard, B. P.; Rupasinghe, H. P. V., Kinase-targeted cancer therapies: progress, challenges and future directions. *Mol. Cancer* **2018**, *17* (1), 48.
6. Gharwan, H.; Groninger, H., Kinase inhibitors and monoclonal antibodies in oncology: clinical implications. *Nat Rev Clin Oncol* **2016**, *13* (4), 209-227.
7. Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L., How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **2010**, *9* (3), 203-214.
8. Holohan, C.; Van Schaeybroeck, S.; Longley, D. B.; Johnston, P. G., Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer.* **2013**, *13* (10), 714-726.

9. Barouch-Bentov, R.; Sauer, K., Mechanisms of drug resistance in kinases. *Expert. Opin. Investig. Drugs* **2011**, *20* (2), 153-208.
10. Yun, C. H.; Mengwasser, K. E.; Toms, A. V.; Woo, M. S.; Greulich, H.; Wong, K. K.; Meyerson, M.; Eck, M. J., The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (6), 2070-2075.
11. Brave, M.; Goodman, V.; Kaminskas, E.; Farrell, A.; Timmer, W.; Pope, S.; Harapanhalli, R.; Saber, H.; Morse, D.; Bullock, J.; Men, A.; Noory, C.; Ramchandani, R.; Kenna, L.; Booth, B.; Gobburu, J.; Jiang, X.; Sridhara, R.; Justice, R.; Pazdur, R., Sprycel for chronic myeloid leukemia and Philadelphia chromosome-positive acute lymphoblastic leukemia resistant to or intolerant of imatinib mesylate. *Clin. Cancer Res.* **2008**, *14* (2), 352-359.
12. Mondal, J.; Tiwary, P.; Berne, B. J., How a Kinase Inhibitor Withstands Gatekeeper Residue Mutations. *J. Am. Chem. Soc.* **2016**, *138* (13), 4608-4615.
13. Choi, Y. L.; Soda, M.; Yamashita, Y.; Ueno, T.; Takashima, J.; Nakajima, T.; Yatabe, Y.; Takeuchi, K.; Hamada, T.; Haruta, H.; Ishikawa, Y.; Kimura, H.; Mitsudomi, T.; Tanio, Y.; Mano, H.; Group, A. L. K. L. C. S., EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors. *N. Engl. J. Med.* **2010**, *363* (18), 1734-1739.
14. Doebele, R. C.; Pilling, A. B.; Aisner, D. L.; Kutateladze, T. G.; Le, A. T.; Weickhardt, A. J.; Kondo, K. L.; Linderman, D. J.; Heasley, L. E.; Franklin, W. A.; Varella-Garcia, M.; Camidge, D. R., Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. *Clin. Cancer Res.* **2012**, *18* (5), 1472-1482.
15. Katayama, R.; Khan, T. M.; Benes, C.; Lifshits, E.; Ebi, H.; Rivera, V. M.; Shakespeare, W. C.; Iafrate, A. J.; Engelman, J. A.; Shaw, A. T., Therapeutic strategies to overcome crizotinib

resistance in non-small cell lung cancers harboring the fusion oncogene EML4-ALK. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (18), 7535-7540.

16. Katayama, R.; Shaw, A. T.; Khan, T. M.; Mino-Kenudson, M.; Solomon, B. J.; Halmos, B.; Jessop, N. A.; Wain, J. C.; Yeo, A. T.; Benes, C.; Drew, L.; Saeh, J. C.; Crosby, K.; Sequist, L. V.; Iafrate, A. J.; Engelman, J. A., Mechanisms of acquired crizotinib resistance in ALK-rearranged lung Cancers. *Sci. Transl. Med.* **2012**, *4* (120), 120ra17.

17. Huang, Q.; Johnson, T. W.; Bailey, S.; Brooun, A.; Bunker, K. D.; Burke, B. J.; Collins, M. R.; Cook, A. S.; Cui, J. J.; Dack, K. N.; Deal, J. G.; Deng, Y. L.; Dinh, D.; Engstrom, L. D.; He, M.; Hoffman, J.; Hoffman, R. L.; Johnson, P. S.; Kania, R. S.; Lam, H.; Lam, J. L.; Le, P. T.; Li, Q.; Lingardo, L.; Liu, W.; Lu, M. W.; McTigue, M.; Palmer, C. L.; Richardson, P. F.; Sach, N. W.; Shen, H.; Smeal, T.; Smith, G. L.; Stewart, A. E.; Timofeevski, S.; Tsaparikos, K.; Wang, H.; Zhu, H.; Zhu, J.; Zou, H. Y.; Edwards, M. P., Design of potent and selective inhibitors to overcome clinical anaplastic lymphoma kinase mutations resistant to crizotinib. *J. Med. Chem.* **2014**, *57* (4), 1170-1187.

18. Cournia, Z.; Allen, B.; Sherman, W., Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57* (12), 2911-2937.

19. Hansen, N.; van Gunsteren, W. F., Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10* (7), 2632-2647.

20. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A., THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13* (8), 1011-1021.

21. Zhao, Z.; Xie, L.; Xie, L.; Bourne, P. E., Delineation of Polypharmacology across the Human Structural Kinome Using a Functional Site Interaction Fingerprint Approach. *J. Med. Chem.* **2016**, *59* (9), 4326-4341.
22. Zhao, Z.; Xie, L.; Bourne, P. E., Insights into the binding mode of MEK type-III inhibitors. A step towards discovering and designing allosteric kinase inhibitors across the human kinome. *PLoS One* **2017**, *12* (6), e0179936.
23. Zhao, Z.; Liu, Q.; Bliven, S.; Xie, L.; Bourne, P. E., Determining Cysteines Available for Covalent Inhibition Across the Human Kinome. *J. Med. Chem.* **2017**, *60* (7), 2879-2889.
24. Zhao, Z.; Xie, L.; Bourne, P. E., Structural Insights into Characterizing Binding Sites in Epidermal Growth Factor Receptor Kinase Mutants. *J. Chem. Inf. Model.* **2019**, *59* (1), 453-462.
25. Friboulet, L.; Li, N.; Katayama, R.; Lee, C. C.; Gainor, J. F.; Crystal, A. S.; Michellys, P. Y.; Awad, M. M.; Yanagitani, N.; Kim, S.; Pferdekamper, A. C.; Li, J.; Kasibhatla, S.; Sun, F.; Sun, X.; Hua, S.; McNamara, P.; Mahmood, S.; Lockerman, E. L.; Fujita, N.; Nishio, M.; Harris, J. L.; Shaw, A. T.; Engelman, J. A., The ALK inhibitor ceritinib overcomes crizotinib resistance in non-small cell lung cancer. *Cancer Discov.* **2014**, *4* (6), 662-673.
26. Johnson, T. W.; Richardson, P. F.; Bailey, S.; Brooun, A.; Burke, B. J.; Collins, M. R.; Cui, J. J.; Deal, J. G.; Deng, Y. L.; Dinh, D.; Engstrom, L. D.; He, M.; Hoffman, J.; Hoffman, R. L.; Huang, Q.; Kania, R. S.; Kath, J. C.; Lam, H.; Lam, J. L.; Le, P. T.; Lingardo, L.; Liu, W.; McTigue, M.; Palmer, C. L.; Sach, N. W.; Smeal, T.; Smith, G. L.; Stewart, A. E.; Timofeevski, S.; Zhu, H.; Zhu, J.; Zou, H. Y.; Edwards, M. P., Discovery of (10R)-7-amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2H-8,4-(metheno)pyrazolo[4,3-h][2,5,11]-benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a macrocyclic inhibitor of anaplastic lymphoma kinase (ALK) and c-ros oncogene 1 (ROS1) with preclinical brain

exposure and broad-spectrum potency against ALK-resistant mutations. *J. Med. Chem.* **2014**, *57* (11), 4720-4744.

27. Lovera, S.; Morando, M.; Pucheta-Martinez, E.; Martinez-Torrecuadrada, J. L.; Saladino, G.; Gervasio, F. L., Towards a Molecular Understanding of the Link between Imatinib Resistance and Kinase Conformational Dynamics. *PLoS Comput. Biol.* **2015**, *11* (11), e1004578.

28. Noble, M. E.; Endicott, J. A.; Johnson, L. N., Protein kinase inhibitors: insights into drug design from structure. *Science* **2004**, *303* (5665), 1800-1805.

29. Torrie, G. M.; Valleau, J. P., Nonphysical Sampling Distributions in Monte Carlo Free Energy estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 13.

30. Mezei, M., Adaptive Umbrella Sampling - Self-Consistent Determination of the Non-Boltzmann Bias. *J. Comput. Phys.* **1987**, *68* (1), 237-248.

31. Souaille, M.; Roux, B., Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput Phys Commun* **2001**, *135* (1), 40-57.

32. Kellenberger, E.; Muller, P.; Schalon, C.; Bret, G.; Foata, N.; Rognan, D., sc-PDB: an annotated database of druggable binding sites from the Protein Data Bank. *J. Chem. Inf. Model.* **2006**, *46* (2), 717-727.

33. Marcou, G.; Rognan, D., Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model.* **2007**, *47* (1), 195-207.

34. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235-242.

35. Cui, J. J.; Tran-Dube, M.; Shen, H.; Nambu, M.; Kung, P. P.; Pairish, M.; Jia, L.; Meng, J.; Funk, L.; Botrous, I.; McTigue, M.; Grodsky, N.; Ryan, K.; Padrique, E.; Alton, G.;

- Timofeevski, S.; Yamazaki, S.; Li, Q.; Zou, H.; Christensen, J.; Mroczkowski, B.; Bender, S.; Kania, R. S.; Edwards, M. P., Structure based drug design of crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal-epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK). *J. Med. Chem.* **2011**, *54* (18), 6342-6363.
36. Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinformatics* **2016**, *54*, 5.6.1–5.6.37.
37. The UniProt, C., UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **2017**, *45* (D1), D158-D169.
38. Jain, A. N., Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46* (4), 499-511.
39. Vanommeslaeghe, K.; MacKerell, A. D., Jr., Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, *52* (12), 3144-3154.
40. Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr., Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, *52* (12), 3155-3168.
41. Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30* (10), 1545-1614.

42. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31* (4), 671-690.
43. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926-935.
44. Harvey, M. J.; Giupponi, G.; Fabritiis, G. D., ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* **2009**, *5* (6), 1632-1639.
45. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81* (8), 3684-3690.
46. Bakan, A.; Meireles, L. M.; Bahar, I., ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **2011**, *27* (11), 1575-1577.
47. Desaphy, J.; Raimbaud, E.; Ducrot, P.; Rognan, D., Encoding protein-ligand interaction patterns in fingerprints and graphs. *J. Chem. Inf. Model.* **2013**, *53* (3), 623-637.
48. Galili, T., dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **2015**, *31* (22), 3718-3720.

Table of Content

