Insights into the binding mode of MEK type-III inhibitors. A step towards discovering and designing allosteric kinase inhibitors across the human kinome

Zheng Zhao[1], Lei Xie*[2,3] and Philip E. Bourne*[4]

1. National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Bethesda, MD 20892, USA

2. Department of Computer Science, Hunter College, The City University of New York, NY 10065, USA

3. The Graduate Center, The City University of New York, NY 10016, USA

4. Office of the Director, National Institutes of Health, Bethesda, MD 20892, USA

***To whom correspondences should be addressed:**

Lei Xie: Tel: +1-212-396-6550; email: lei.xie@hunter.cuny.edu

Philip E. Bourne: Tel: +1-301-402-9818; email: philip.bourne@nih.gov

Abstract

Protein kinases are critical drug targets for treating a large variety of human diseases. Type-I and type-II kinase inhibitors frequently exhibit off-target toxicity or lead to mutation acquired resistance. Two highly specific allosteric type-III MEK-targeted drugs, Trametinib and Cobimetinib, offer a new direction. Thus, understanding the binding mechanism of existing type-III kinase inhibitors will provide insights into improving the efficiency of designing and discovering such drugs. In this work we have systematically studied the binding mode of MEK-targeted type-III inhibitors using structural systems pharmacology and molecular dynamics simulation. Our studies provide detailed sequence, structure, interaction-fingerprint, pharmacophore and binding-site information on the binding characteristics of MEK type-III kinase inhibitors. Moreover, we have screened for similar target binding sites across the human kinome suggesting other potential kinase targets for type-III inhibitors.

Introduction

Kinases are phosphorylation enzymes that catalyze the transfer of phosphate groups from ATP to specific substrates and are critical in most cellular life processes.[1, 2] Abnormal kinase regulation, which leads to signal disruption and cell deregulation, is implicated in many diseases, particularly cancers.[3] Thus a number of kinase-targeted small molecule inhibitors have been developed and are important in anti-cancer therapy.[4] Since 2001, 30 small molecule kinase inhibitors[5, 6] have been approved by the US Food and Drug Administration (FDA) for anti-cancer therapy and more inhibitors are undergoing clinical trials[7,8].

However, reported off-target toxicities and acquired-mutation resistance[9] require kinase-targeted inhibitors of lower dose and high specificity. Three types of kinase-targeted inhibitors, type-I, type-II and type-III, have been developed.[10, 11] Type-I inhibitors are ATP-competitive inhibitors and occupy the ATP-binding binding pocket, a highly conserved kinase catalytic scaffold with strong binding affinity for ATP. The difficulty of achieving selectivity combined with mutation-induced resistance has led to the development of type-II and type-III inhibitors. Type-II inhibitors bind to the inactive kinase and not only occupy the ATP-binding pocket, but extend to occupy the adjacent less-conserved allosteric site across the DFG segment, which is accessible in the DFG-out inactive kinase conformation.[12] Most type-I and type-II inhibitors bind to multiple targets.[9] Type-III inhibitors are unique in that they occupy highly specific allosteric sites as exemplified by the type-III MEK inhibitors.[13, 14] The reported evidence so far suggests that type-III inhibitors are high selective and effective even if the kinase or signaling pathway has undergone drug-resistance residue mutation.[15] For example, a FDA-approved type-III MEK kinase inhibitor, Cobimetinib, can overcome the resistance from the typical BRAF V600E mutation in melanoma.[14] However, thus far, there is no systematic means for identifying the preferred characteristics for specific type-III inhibitors.[8] Existing type-III kinase inhibitors mainly target

MEK.[15] Thus it is important to understand the molecular characteristics of the interaction between MEK kinase and type-III kinase inhibitors, so as to extend the design of type-III kinase inhibitors to other kinases.

In this work we have integrated our structural systems biology strategy and molecular dynamics simulation methods to gain insight into type-III kinase inhibitors and their interactions with kinases across the human kinome. Our structural system biology strategy harnesses different omics data to compare and discover the gene-level, protein-level and structure-level information on protein-ligand interactions.[16] We have successfully applied this strategy to drug design and discovery for different target-drug kinase systems and the Ebola virus proteome.[5, 17] With increased computing power and more efficient algorithms, molecular dynamics (MD) simulation is now becoming a routine tool for drug design accounting for the reality of a flexible target structure and flexible target-drug binding.[18] In this paper, we performed detailed MEK kinase-drug function-site interactional fingerprint analysis using our structural system biology strategy. We also performed two MD simulations up to 1.2 μs in an explicit water box to obtain insights into the behavior of flexible MEK kinases as targets, with and without a representative ligand, Cobimetinib.[19] By comparing the structural trajectories between MEK kinase with and without ligand, we determined the structural flexibility and interactional network for type-III inhibitor binding in MEK kinases. In addition, we delved into the details of genetic structural, pharmacophore' and mechanistic understanding of MEK kinase-drug binding. Finally, we explored the whole human kinome to identify potentially new opportunities for type-III inhibitor drug design.

Results

Binding characteristics of crystalized ligand-bound MEK complexes.

We obtained the binding characteristics of ligand-bound MEK complexes as shown in Figure 1. Figure 1a illustrates the alignment of the 28 ligand-bound MEK structures with the ligands shown in the same allosteric binding site. We calculated the detailed interactions between MEK and the ligand using the Fs-IFP method (Figure 1b). The highly conserved interactions between the respective ligands and MEK include K97, L115, L118, V127, M143, C^{207}DFGVS^{212}, I215 and M219 (Figure 1b, and 1c). These conserved interactions can be divided into three spatial regions.

The first region is the hydrophobic sub-pocket consisting of L115, L118, V127 and M143 as shown in purple in Figure 1c. All interactions are apolar (Figure 1b). Correspondingly, all ligands have hydrophobic groups that can be accommodated in the sub-pocket. For example, Comimetinib has a 2-fluoro-4-iodoanilino fragment, as shown in purple in Figure 2 (4an2), which is a well known hydrophobic pocket binder. Other ligands also have the same or similar groups (Figure 2, the group in purple) so as to achieve the high binding affinity of the conserved sub-pocket. Pharmacophore modeling, as shown in Figure 1d, also revealed similar patterns with three common hydrophobic groups, H3, H4 and R7.

The second region is K97, an important catalytic residue, is located at the roof of the binding pocket. K97 has a conserved interaction with the oxygen atom (marked O* in Fig. 2, 4an2). Moreover, the molecular moiety O* is conserved as shown in dark blue (Figure 2; 4an2) either as an oxygen or nitrogen atom. Pharmacophore modeling (Figure 1d) is consistent with one donor-type hydrogen bond, D2.

The third region consists of C^{207}DFGVS^{212}, I215 and M219, which form a loop and a helix which act like an arm to hug the inhibitor. DFG is directly involved in kinase catalytic activity across the family. I215 and M219 are located at the activation loop. Interestingly in MEK the activation loop is a conserved helix which forms the allosteric site. S212 of the loop forms a

conserved interaction (Fig. 1b) including an apolar and hydrogen-bond interaction. This is consistent with experiment which shows S212 plays a key role in phosphorylation and hence MEK activity.[14] In this third region all active ligands have one atom, F, N or O that interacts with S212 (shown as F* in red Figure 2, 4an2). Pharmacophore modeling (Fig. 1d, H5 and R6) illustrates that all ligands have common features in their interaction with S212. Taken together, the aforementioned three regions make major contributions to ligand binding in the allosteric pocket. Summarizing Figure 1d, the hydrophobic heads (H3, H4 and R7) accommodate the hydrophobic sub-pocket and D2 and H5 interact with the roof amino acid K97 and S212 of the loop, respectively. It is expected that an active MEK inhibitor would have these chemical functional groups or similar. Furthermore, in 3D space, these atoms are spatially conserved, as shown by the triangles in Figure 2 for D2, H4 and H5, suggesting that the design of MEK allosteric inhibitors should follow this pharmacophore. Similar conserved spatial requirements have been reported in the design of other allosteric inhibitors.[20] Of course, besides the conserved pharmacophores, different inhibitors may use specific interactions involving other amino acids to achieve selectivity, as shown in Figure 1b. Specifically, the tail of the inhibitor, shown in green in Figure 2, can be changed using different chemical group or atoms. In keeping, the pharmacophore does not reveal common features in the tail (Fig. 1d).

MEK structural flexibility and insights into the binding mechanism

Beyond the static PDB structures we considered how type-III inhibitors induce MEK structural flexibility. We performed two 0.6 μs MD simulations for MEK kinase with and without the inhibitor bound, respectively. The Cα-atom RMSD is shown in Figure S1. The Cα-RMSDs of the apo and holo structures are similar. However, their Cα-atom fluctuations show significant difference flexibility, especially in the ligand-bound regions as shown in Figure 3a, where we

compared MEK structural flexibility before and after binding using the two last 0.5 μs equilibrated MD trajectories. For apo MEK the flexibility change mainly comes from the P-loop, the activation loop and the C-terminal lobe as shown in Figure 3a. The corresponding collective motion as inferred from the first principal component of PCA is shown in Figure 3b. As a comparison, in the MEK-Cobimetinib complex, the main fluctuations come from the parts of the C-helix, the sequences following the activation loop and the C-terminal lobe (Fig. 3a). The obvious difference before and after binding inhibitor is that the collective motions of the P-loop and activation loop have undergone a substantial reduction in the MEK-Cobimetinib complex (Fig. 3c). However, the flexibility of the C-helix has significantly increased in apo MEK. Similar to other kinases, the P-loop contributes to conformational flexibility and plays an important role in binding and recognizing phosphoryl moieties.[21] Moreover, this flexible P-loop motif, along with other beta-sheets and helixes, generally form a pocket into which the phosphate groups can insert.[22, 23] Like the P-loop, the activation loop shows a similar change in flexibility, before and after Cobimetinib binding. Thus, MEK is induced to fit the inhibitor and the resulting increased rigidity in the activation-loop reduces MEK activity. Within the kinases the C-helix, Lys97 and DFG peptide contribute to forming the ATP binding site needed for enzyme catalysis, including a hydrogen bond between the C-helix and Lys97. Upon Cobimetinib binding MEK enters an inactive state and the hydrogen bond is broken. Subsequently, the C-helix is more flexible than in the active state.[8, 24] In addition, the activation loop is a helix in MEK. However, in most kinases, the activation loop is a flexible loop. This implies that the helix is responsible for forming the binding allosteric pocket that fits the type-III inhibitor. Thus consideration of the nature of the adopted helix as part of the activation loop is an important consideration when designing type-III inhibitors.

Conserved interactions with inhibitors from S212 and K97

As aforementioned, the inhibitors derived from PDB MEK structures have a similar core and common functional groups forming a conserved spatial triangular arrangement (Fig. 2). Correspondingly, in the MEK-Cobimetinib MD trajectories the conserved interactions between MEK and respective inhibitors were also evaluated. Two key interactions between S212 and K97 and the inhibitors were found (Fig. 4). The interaction between S212 and the F* atom of the inhibitor is shown in red; from the probability distribution (Fig. 4b) the center of the peak is at 3.2 Å, which suggests that a hydrogen bond interaction is conserved at all times to maintain the binding affinity and hinder MEK phosphorylation. This observed hydrogen bond agrees with experiment results.[14] The O* atom is another conserved polar atom contributing to the effective binding. As shown in green in Figure 4, the position of the peak in the probability distribution is at approximately 5.2 Å, which agrees with the distance found in PDB crystal structures.[13, 14] This distance guarantees that there is an interaction between the O* and the carboxy group of the aspartic acid in the DFG segment, which contributes to forming the catalytic center.[25] Thus the O* atom from the inhibitor, carboxy group of the aspartic acid of DFG and the amino group of the side chain of K97 form the pseudo catalytic center and subsequently result in deactivation of the kinase catalytic function.[26]

Ranking similar binding sites to MEK using SMAP across the structural kinome

To design type-III inhibitors for any given kinase target, establishing similar binding pockets to MEK is an important step. Using SMAP[27-29], we screened for similar binding sites across the human kinome (Table 1). There are 38 binding pockets that have more than 55% similarity to MEK. Not surprisingly, 35 of these structures are MEK kinases, two are P38a[30] and one is BRAF[2]. All have activation loops that contain a helix. We aligned the sequences of MEK, P38a and BRAF

with particular attention to the activation loop. The similarity between them is not high as shown in Figure S2, where the activation loops are marked with a rectangle. These results imply that kinases with similar secondary structure in their activation loops have the potential to be inhibited by type-III inhibitors even though their global sequences do not have high similarity with MEK. Stated another way, a helix in the activation loop provides structural insight into designing type-III inhibitors and caused us to investigate which kinases can potentially form a helix in the activation loop.

Predicting activation loops with helix character across the human kinome

We predicted the secondary structure from protein sequence of all kinases in the human kinome to explore the kinome for the potential inhibition pose suitable for type-III inhibitors.[8] The kinases with the potential adopted helix within the activation loop are shown in Figure 5 and detailed in Table S1. The top 24 kinases are mostly located within the STE group forming the MAPK cascade. Thus these 24 kinases are potential targets of type-III inhibitors which could be screened in vivo or in vitro.

Conclusions

Recently FDA approved type-III allosteric inhibitors provide an important opportunity to design and discover more efficient and lower-dose kinase inhibitors. Study of the characteristics of the type-III inhibitor-bound binding site provides structural insights into the design of new allosteric drugs. Here we study the characteristics of the MEK binding site, the chemical nature of the inhibitors that bind MEK, the characteristics of binding and the dynamic nature of the interaction. Further, based on all 3D kinase structures, we screened for potential allosteric inhibitor-bound binding sites include revealed BRAF and P38a in addition to MEK. Finally, based on the

distinctive helix characteristic of the MEK activation loop to adopt an inhibitory pose,[31] we predicted 24 kinases that can potentially form the allosteric site to bind type-III allosteric inhibitors. In summary, our in silico analysis and prediction contributes to our understanding of the structure and ligand requirement for designing allosteric inhibitors.

Methods

Function-site interaction fingerprint scheme

Function-site Interaction FingerPrint (Fs-IFP) is a method to reveal the functional site binding characteristics and to compare binding sites on a proteome scale.[5] Here we use Fs-IFP to reveal the binding characteristics of the MEK-inhibitor complex for all released MEK structures from the Protein Data Bank (PDB).[32] In brief, firstly we downloaded all released 35 MEK structures from the PDB. Herein the 28 ligand-bound structures formed the MEK structure dataset; then we aligned all the binding sites of these ligand-bound structures using SMAP[27-29] and encoded the Fs-IFP as published in an earlier paper[5] using the Pyplif software.[33] The end result for each structure is a representation of the interactional details between every involved amino acid and the ligand using one-dimension bits.

Pharmacophore modeling

Pharmacophore modeling was performed using Maestro in Schrodinger released 2016.02.[34] Based on the MEK structure dataset, we collected the activity data of all ligands for MEK, the binding affinity data are shown in Table S2. Among them, 19 entries with MEK IC50 value were used as input. The pharmacophore model was trained using the ligands with an IC50 of less than 10 nM.

MD simulation

Two MD simulations were performed with starting conformations taken from the PDB; pdb id 4an2 for the MEK-Cobimetinib complex, and the MEK apo structure from pdb id 3zls. Both initial conformations were prepared for MD simulations using the ACEMD protocol.[35] The protonation states of both systems were assigned as pH=7.0, similar to the cellular environment. Then every His state and every disulfide bond were checked to make sure they conformed to pH=7.0. The systems were solvated in a rectangular water box with at least a 12Å shell buffer away from any-solute atoms and charged ions were added to ensure an ionic strength of 0.20 M and electroneutrality of the systems. The CHARMM27 force field,[36, 37] CHARMM general force field[38] and TIP3P force field were used for the kinase, ligand, and water molecules, respectively. Both systems were relaxed with a conversional MD protocol including 2ps minimization, 100ps for NVT, 1ns for NPT with heavy-atom constraints and 1ns for NPT without any constraints. Subsequently, 0.6 μs MD simulations were performed for every system. In both MD simulations all bonds were constrained using SHAKE and the integration time step was 4 fs. The temperature bath used the Langevin method, and 1atm pressure was maintained using the Berendsen method. Both simulations were carried out using the ACEMD software[35] on the NIH high-performance Biowulf cluster (https://hpc.nih.gov/). The MD results were analyzed using the conformations during the last 0.5 μs MD trajectory. The analysis of MD trajectories, including Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF) and Principal Component Analysis (PCA), were performed with the Wordom tool.[39]

Screening for similar binding pockets across the human structural kinome

Approximately three thousands protein kinase structures have been solved by X-ray and NMR methods. We assembled 2797 kinase structures from the PDB which included the catalytic domain

as the human structural kinome.[5] We then used a MEK-Cobimetinib complex (PDB id 4LMN) as a template to rank similar binding sites from the human structural kinome by performing a one-to-all comparison using SMAP.[27-29] The top-ranked binding sites with $p\text{-values} < 0.05$ were chosen for further analysis.

Secondary structure prediction

Determining type-III inhibitors needs to be validated, not only by biological activity assay, but also through analysis of an appropriate crystal structure.[8] Unfortunately, the type-III binding pocket is often not present in the crystal structure when there is no specific type-III inhibitor cocrystalized.[8] To obtain the MEK-similar allosteric site in such kinases, we predicted the MEK-like secondary structure of the activation loop to establish the binding pose of the type-III inhbitor.[31] We used a protein secondary structure prediction server Jpred4[40] for the task. First, the 516 kinase domain sequences for the eukaryotic protein kinase superfamily were downloaded at kinase.com and were aligned using the cluster omega software.[41] From the alignment the amino acid sequences of the activation loop for every kinase were extracted with additional 11 N-terminal and 11 C-terminal residues adjoining the DFG segments. Finally, the activation loop structure was predicted using the JPred RESTful API(v1.5)[40] with default parameters.

Acknowledgment

This research was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health (Z.Z. and P.E.B.), the National Library of Medicine, National Institutes of Health under Award R01LM011986 (L.X.), and the National Institute on

Minority Health and Health Disparities, National Institutes of Health, under Award

G12MD007599 (L.X.).

References

1. Klebl, B., Müller, G., and Hamacher, M., (Eds.) (2011) *Protein Kinase as Drug Targets*, Vol. 49, WILEY-VCH Verlag GambH & Co. KGaA, Weinheim.
2. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science 298*, 1912-1934.
3. Lahiry, P., Torkamani, A., Schork, N. J., and Hegele, R. A. (2010) Kinase mutations in human disease: interpreting genotype-phenotype relationships, *Nat. Rev. Genet. 11*, 60-74.
4. Abramson, R. (2016) Overview of Targeted Therapies for Cancer. My Cancer Genome https://www.mycancergenome.org/content/molecular-medicine/overview-of-targeted-therapies-for-cancer/ (Updated April 26).
5. Zhao, Z., Xie, L., Xie, L., and Bourne, P. E. (2016) Delineation of Polypharmacology across the Human Structural Kinome Using a Functional Site Interaction Fingerprint Approach, *J. Med. Chem. 59*, 4326-4341.
6. Wu, P., Nielsen, T. E., and Clausen, M. H. (2016) Small-molecule kinase inhibitors: an analysis of FDA-approved drugs, *Drug Discov. Today 21*, 5-10.
7. Rask-Andersen, M., Zhang, J., Fabbro, D., and Schioth, H. B. (2014) Advances in kinase targeting: current clinical use and clinical trials, *Trends Pharmacol. Sci. 35*, 604-620.
8. Muller, S., Chaikuad, A., Gray, N. S., and Knapp, S. (2015) The ins and outs of selective kinase inhibitor development, *Nat. Chem. Biol. 11*, 818-821.
9. Gharwan, H., and Groninger, H. (2016) Kinase inhibitors and monoclonal antibodies in oncology: clinical implications, *Nat. Rev. Clin. Oncol. 13*, 209-227.
10. Kooistra, A. J., Kanev, G. K., van Linden, O. P., Leurs, R., de Esch, I. J., and de Graaf, C. (2016) KLIFS: a structural kinase-ligand interaction database, *Nucleic Acids Res. 44*, D365-371.
11. Roskoski, R., Jr. (2016) Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes, *Pharmacol Res. 103*, 26-48.
12. Zhao, Z., Wu, H., Wang, L., Liu, Y., Knapp, S., Liu, Q., and Gray, N. S. (2014) Exploration of type II binding mode: A privileged approach for kinase inhibitor focused drug discovery?, *ACS Chem. Biol. 9*, 1230-1241.
13. Rice, K. D., Aay, N., Anand, N. K., Blazey, C. M., Bowles, O. J., Bussenius, J., Costanzo, S., Curtis, J. K., Defina, S. C., Dubenko, L., Engst, S., Joshi, A. A., Kennedy, A. R., Kim, A. I., Koltun, E. S., Lougheed, J. C., Manalo, J. C., Martini, J. F., Nuss, J. M., Peto, C. J., Tsang, T. H., Yu, P., and Johnston, S. (2012) Novel Carboxamide-Based Allosteric MEK Inhibitors: Discovery and Optimization Efforts toward XL518 (GDC-0973), *ACS Med. Chem. Lett. 3*, 416-421.
14. Hatzivassiliou, G., Haling, J. R., Chen, H., Song, K., Price, S., Heald, R., Hewitt, J. F., Zak, M., Peck, A., Orr, C., Merchant, M., Hoeflich, K. P., Chan, J., Luoh, S. M., Anderson, D. J., Ludlam, M. J., Wiesmann, C., Ultsch, M., Friedman, L. S., Malek, S., and Belvin, M. (2013) Mechanism of MEK inhibition determines efficacy in mutant KRAS- versus BRAF-driven cancers, *Nature 501*, 232-236.
15. Wu, P., Clausen, M. H., and Nielsen, T. E. (2015) Allosteric small-molecule kinase inhibitors, *Pharmacol Ther. 156*, 59-68.
16. Bourne, P. E., and Xie, L. (2017) Harnessing 'Big Data' in Systems Pharmacology, *Annu. Rev. Pharmacol. Toxicol. 57*, doi:10.1146/annurev-pharmtox-010716-104659.

17. Zhao, Z., Martin, C., Fan, R., Bourne, P. E., and Xie, L. (2016) Drug repurposing to target Ebola virus replication and virulence using structural systems pharmacology, *BMC Bioinformatics 17*, 90.
18. De Vivo, M., Masetti, M., Bottegoni, G., and Cavalli, A. (2016) Role of Molecular Dynamics and Related Methods in Drug Discovery, *J. Med. Chem. 59*, 4035-4061.
19. Garnock-Jones, K. P. (2015) Cobimetinib: First Global Approval, *Drugs 75*, 1823-1830.
20. Jia, Y., Yun, C. H., Park, E., Ercan, D., Manuia, M., Juarez, J., Xu, C., Rhee, K., Chen, T., Zhang, H., Palakurthi, S., Jang, J., Lelais, G., DiDonato, M., Bursulaya, B., Michellys, P. Y., Epple, R., Marsilje, T. H., McNeill, M., Lu, W., Harris, J., Bender, S., Wong, K. K., Janne, P. A., and Eck, M. J. (2016) Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors, *Nature 534*, 129-132.
21. Smith, C. A., and Rayment, I. (1996) Active site comparisons highlight structural similarities between myosin and other P-loop proteins, *Biophys. J. 70*, 1590-1602.
22. Deyrup, A. T. (1998) Deletion and Site-directed Mutagenesis of the ATP-binding Motif (P-loop) in the Bifunctional Murine Atp-Sulfurylase/Adenosine 5'-Phosphosulfate Kinase Enzyme, *J. Biol. Chem. 273*, 9450-9456.
23. McClendon, C. L., Kornev, A. P., Gilson, M. K., and Taylor, S. S. (2014) Dynamic architecture of a protein kinase, *Proc. Natl. Acad. Sci. U.S.A. 111*, E4623-4631.
24. Mazanetz, M. P., Laughton, C. A., and Fischer, P. M. (2014) Investigation of the flexibility of protein kinases implicated in the pathology of Alzheimer's disease, *Molecules 19*, 9134-9159.
25. Adams, J. A. (2001) Kinetic and Catalytic Mechanisms of Protein Kinases, *Chem. Rev. 101*, 2271-2290.
26. Kornev, A. P., Haste, N. M., Taylor, S. S., and Eyck, L. F. (2006) Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism, *Proc. Natl. Acad. Sci. U.S.A. 103*, 17783-17788.
27. Xie, L., and Bourne, P. E. (2007) A robust and efficient algorithm for the shape description of protein structures and its application in predicting ligand binding sites, *BMC Bioinformatics 8 Suppl 4*, S9.
28. Xie, L., and Bourne, P. E. (2008) Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments, *Proc. Natl. Acad. Sci. U.S.A. 105*, 5441-5446.
29. Xie, L., Xie, L., and Bourne, P. E. (2009) A unified statistical model to support local sequence order independent similarity searching for ligand-binding sites and its application to genome-based drug discovery, *Bioinformatics 25*, i305-312.
30. Adachi-Yamada, T., Nakamura, M., Irie, K., Tomoyasu, Y., Sano, Y., Mori, E., Goto, S., Ueno, N., Nishida, Y., and Matsumoto, K. (1999) p38 mitogen-activated protein kinase can be involved in transforming growth factor beta superfamily signal transduction in Drosophila wing morphogenesis, *Mol. Cell Biol. 19*, 2322-2329.
31. Lee, C. C., Jia, Y., Li, N., Sun, X., Ng, K., Ambing, E., Gao, M. Y., Hua, S., Chen, C., Kim, S., Michellys, P. Y., Lesley, S. A., Harris, J. L., and Spraggon, G. (2010) Crystal structure of the ALK (anaplastic lymphoma kinase) catalytic domain, *Biochem. J. 430*, 425-437.
32. Berman, H. M., Bhat, T. N., Bourne, P. E., Feng, Z., Gilliland, G., Weissig, H., and Westbrook, J. (2000) The Protein Data Bank and the challenge of structural genomics, *Nat. Struct. Biol. 7 Suppl*, 957-959.
33. Radifar, M., Yuniarti, N., and Istyastono, E. P. (2013) PyPLIF: python-based protein-ligand fnteraction fingerprinting, *Bioinformatics 9*, 325-328.
34. Schrödinger Release 2016-2: Maestro, version 10.6, Schrödinger, LLC, New York, NY, 2016.
35. Harvey, M. J., Giupponi, G., and Fabritiis, G. D. (2009) ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale, *J. Chem. Theory Comput. 5*, 1632-1639.

36. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem. 4*, 187-217.
37. 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., and Karplus, M. (2009) CHARMM: the biomolecular simulation program, *J. Comput. Chem. 30*, 1545-1614.
38. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and Mackerell, A. D., Jr. (2010) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J. Comput. Chem. 31*, 671-690.
39. Seeber, M., Cecchini, M., Rao, F., Settanni, G., and Caflisch, A. (2007) Wordom: a program for efficient analysis of molecular dynamics simulations, *Bioinformatics 23*, 2625-2627.
40. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015) JPred4: a protein secondary structure prediction server, *Nucleic Acids Res. 43*, W389-394.
41. Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y. M., Buso, N., and Lopez, R. (2015) The EMBL-EBI bioinformatics web and programmatic tools framework, *Nucleic Acids Res. 43*, W580-584.

Figure 1. Binding characters of MEK-ligand complexes. **a)** All MEK structures aligned using SMAP. **b)** Encoding all MEK-ligand interactions. Every line represents the MEK-ligand interaction of one structure, and every column represents the interaction between specific amino acids and the bound ligand. Different colors represent the different types of fingerprint interactions: yellow, no interaction; blue, apolar interaction; red, apolar interaction and hydrogen bond interaction (protein as donor); deep red, hydrogen bond interaction (protein as donor); pink, polar interaction and aromatic interaction; and grey, apolar interaction and hydrogen bond interaction (protein as acceptor). **c)** Spatial representation of MEK-ligand interactions. **d)** Pharmacophore modeling: H, hydrophobic group; R, aromatic ring; D, hydrogen-bond donor.

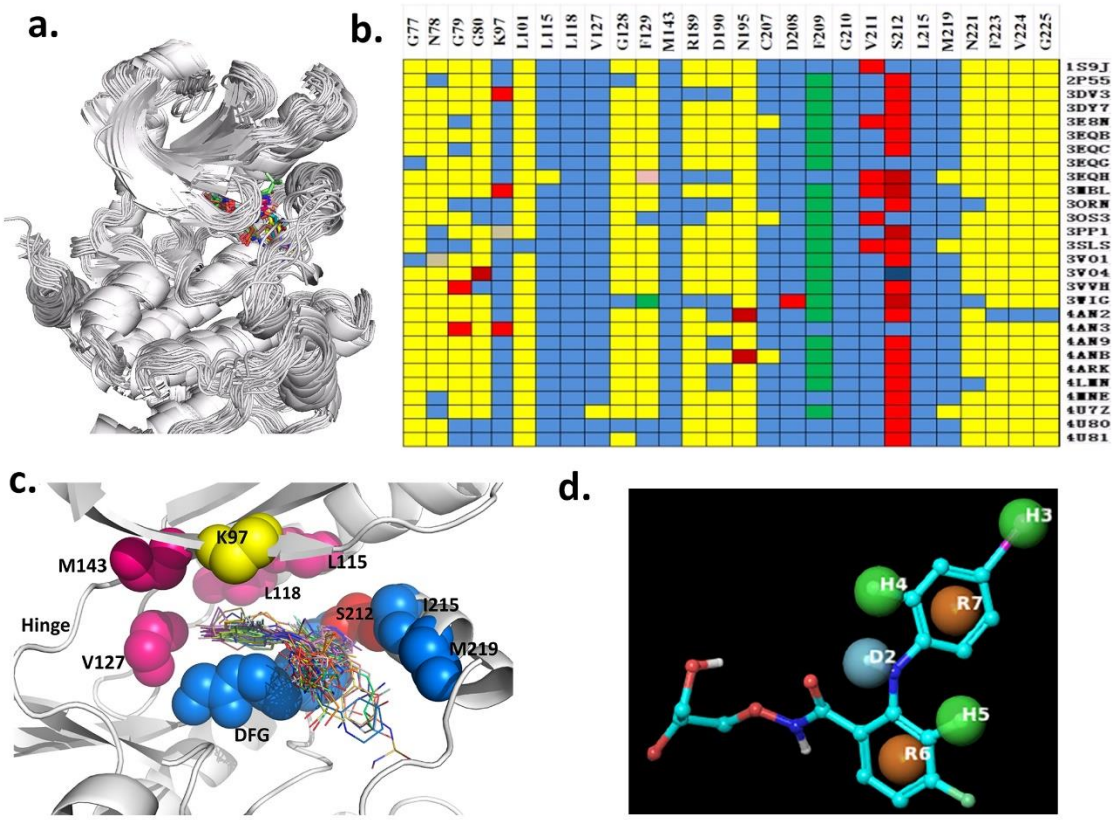

Figure 2. Ligands from the 28 MEK structures. The triangles highlight the conserved atoms that characterize these MEK Type-III inhibitors in 3D space.

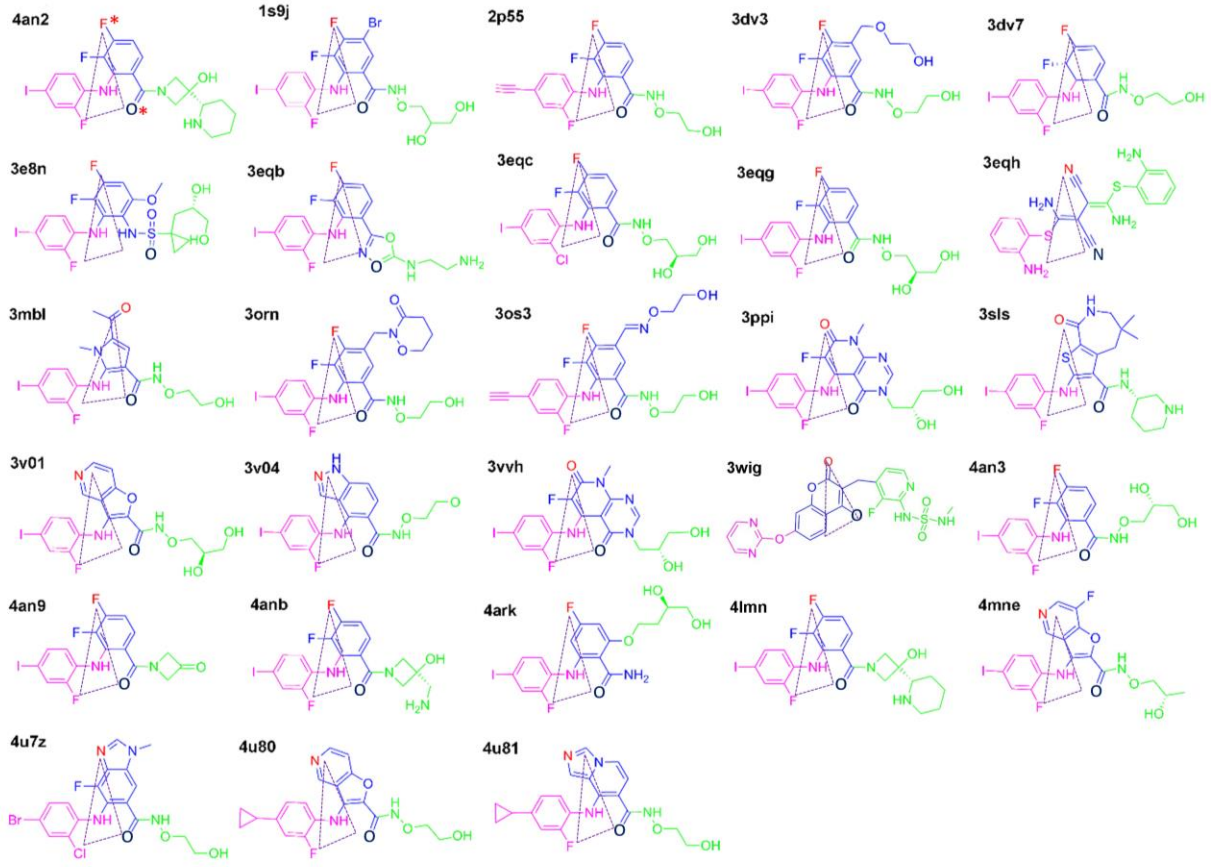

Figure 3. RMSF profiles and PCA projection. **a)** The RMSF profiles from the last 0.5 μm equilibrated MD trajectories of MEK and the MEK-Cobimetinib complex. Some secondary structure elements are shown on the abscissa. **b,c)** The Ca-atom projection along the first principal component. The displacements are shown as color-coded tubes from blue (small displacement) to orange (large displacement) for MEK and the MEK-Cobimetinib complex, respectively.

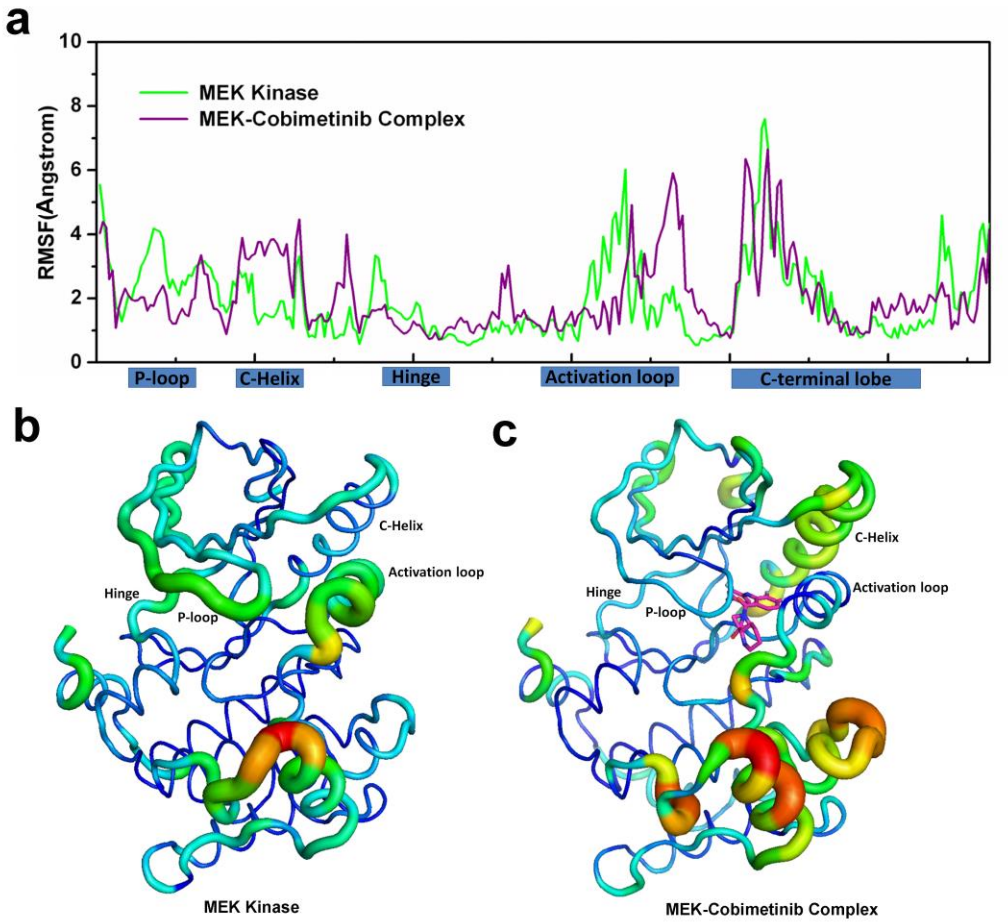

Figure 4. Conserved interatomic interactions between MEK and the ligand. a) Interatomic distance for every conformation in the MD trajectory; b) The probability distribution of interatomic distances.

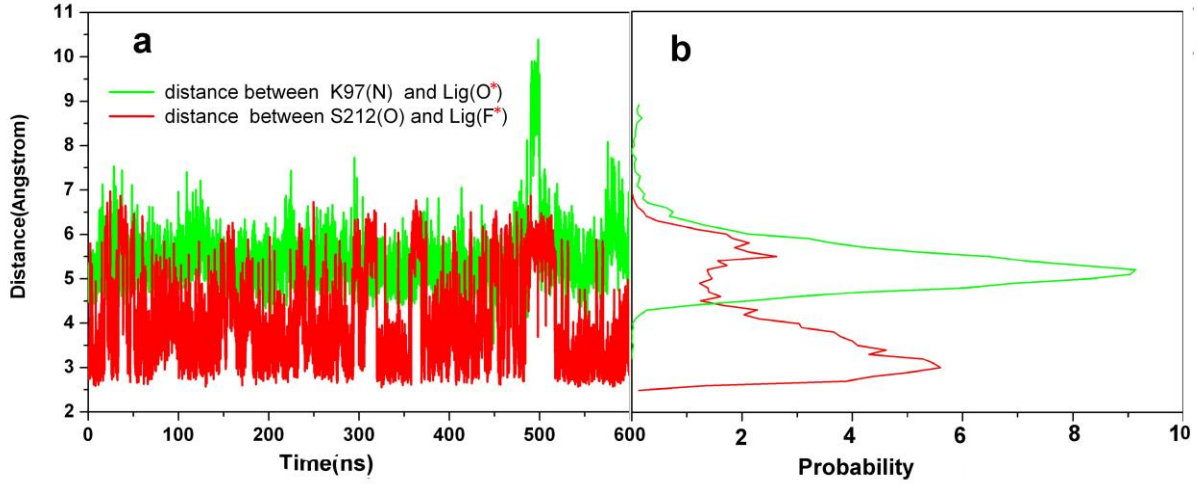

Figure 5.

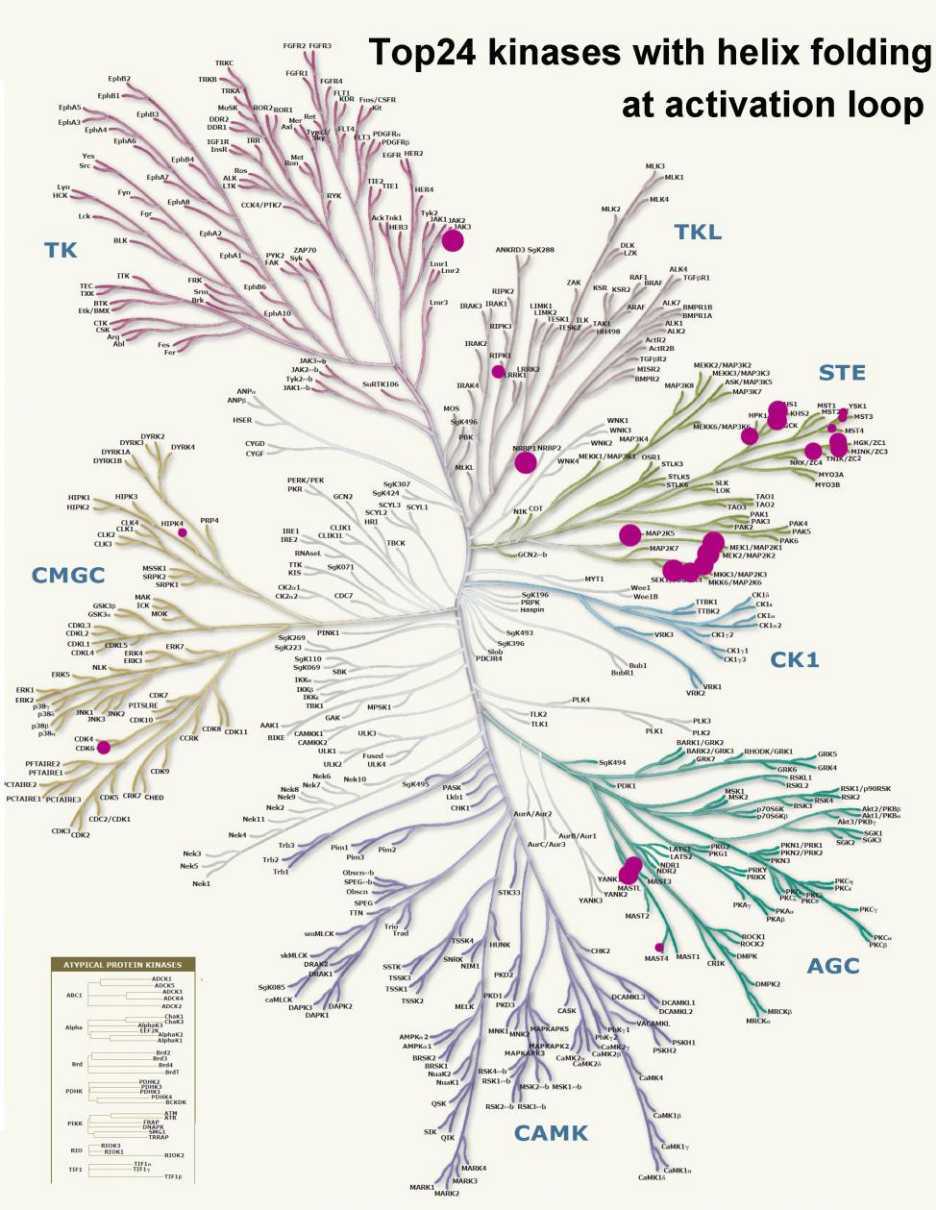

Table 1

Kinase Name	Uniprot ID	PDB id
MEK	Q02750	1S9J,2P55,3DV3,3DY7,3E8N,3EQB,3EQC,3EQD,3EQF,3EQG,3EQH,3EQI,3MBL,3ORN,3OS3,3PP1,3SLS,3V01,3V04,3VVH,3WIG,3ZLS,3ZLX,3ZLY,3ZM4,4AN2,4AN3,4AN9,4ANB,4ARK,4MNE,4U7Z,4U80,4U81,5BX0
P38a	Q16539	2YIX
BRAF	P15056	4PP7,4WO5

Figure S1. The Ca-atom RMSD.

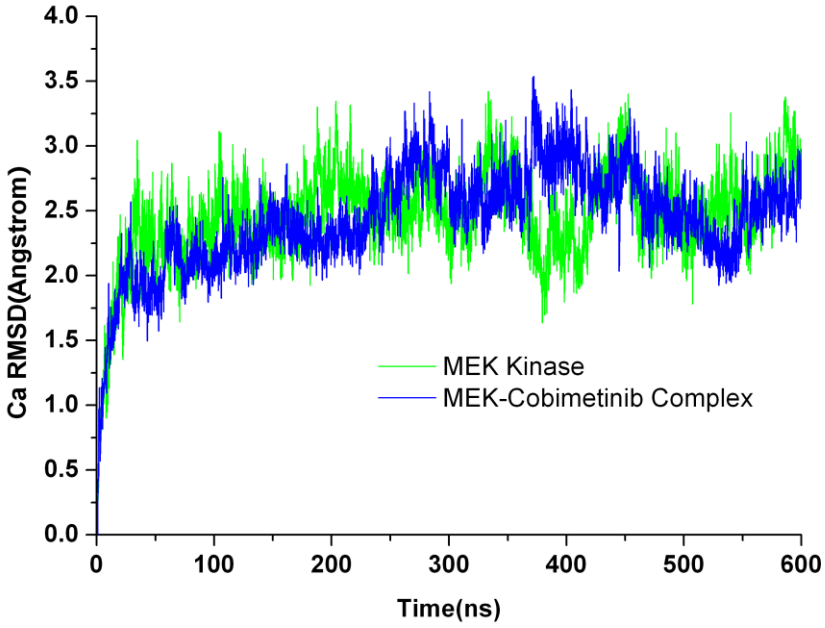

Figure S2. Sequence level similarity for the three kinases (BRAF, MEK, and P38a).

Table S1. The top 24 kinases predicted to contain a helix within the activation loop. The 17 kinases in red are predicted with the highest confidence.

Kinases	Sequences of Activation loop	Predicted secondary structure	Confidence of prediction
CMGC_DYRK_HIPK_HIPK4	DFGSASIFSEVRYVKEPYIQSR	------HH-HHHHHH------	9988100012310123468899
AGC_MAST__MAST4	DFGLSKVGLMSMTTNLYEGHIE	------HHHHHHHHHHHH---	9998710113311223215699
STE_STE20_YSK_MST3	DFGVAGQLTDTQIKRNTFVGTP	-----HHHHHHHHH---E---	9922113113321015111899
STE_STE20_YSK_MST4	DFGVAGQLTDTQIKRNTFVGTP	-----HHHHHHHHH---E---	9922113113321015111899
STE_STE20_YSK_YSK1	DFGVAGQLTDTQIKRNTFVGTP	-----HHHHHHHHH---E---	9922113113321015111899
TKL_RIPK__RIPK1	DLGLASFKMWSKLNNEEHNELR	------HHHHHHH--------	9998724543201112776899
CMGC_CDK_CDK4_CDK6	DFGLARIYSFQMALTSVVVTLW	----HHHHHHHHHHHHHHHH-	9981133321322324443069
TK_JakB__Domain2_JAK3	DPGVSPAVLSLEMLTDRIPWVA	-----HHHHHHHH-------	9998535655555404886148
STE_STE20_KHS_HPK1	DFGISAQIGATLARRLSFIGTP	------HHHHHHHH------	9998651226787750368999
AGC_MAST__MAST3	DFGLSKIGLMSMATNLYEGHIE	------HHHHHHHHHH----	9998601144567776216899
STE_STE20_MSN_ZC1_HGK	DFGVSAQLDRTVGRRNTFIGTP	----HHHHHHHHH-------	9963576888764136316899
STE_STE20_MSN_ZC2_TNIK	DFGVSAQLDRTVGRRNTFIGTP	----HHHHHHHHH-------	9963576888764136316899
STE_STE20_MSN_ZC3_MINK	DFGVSAQLDRTVGRRNTFIGTP	----HHHHHHHHH-------	9963576888764136316899
STE_STE20_KHS_KHS1	DFGVAAKITATIAKRKSFIGTP	----HHHHHHHHHH------	9965313789999811125899
AGC_MAST__MAST1	DFGLSKMGLMSLTTNLYEGHIE	-----HHHHHHHHHH----	9984235488888776425589
STE_STE7__MAP2K3	DFGISGYLVDSVAKTMDAGCKP	----HHHHHHHHHH------	9998146789999832058999
STE_STE7__MAP2K6	DFGISGYLVDSVAKTIDAGCKP	----HHHHHHHHHH------	9998146789999832058999
AGC_MAST__MAST2	DFGLSKIGLMSLTTNLYEGHIE	-----HHHHHHHHHH----	9984235488888887515589
STE_STE20_KHS_KHS2	DFGVSAQITATIAKRKSFIGTP	----HHHHHHHHHH------	9983799999999803213999
STE_STE7__MAP2K4	DFGISGQLVDSIAKTRDAGCRP	-----HHHHHHHHH------	9987278899999823378999
STE_STE7__MAP2K5	DFGVSTQLVNSIAKTYVGTNAY	----HHHHHHHHHHH-----	9941799999999851688999
Other_NRBP__NRBP1	VAPDTINNHVKTCREEQKNLHF	-----HHHHHHHHHH-----	9966457699999987058999
STE_STE7__MAP2K1	DFGVSGQLIDSMANSFVGTRSY	----HHHHHHHHHHH-----	9987178999999986027999
STE_STE7__MAP2K2	DFGVSGQLIDSMANSFVGTRSY	----HHHHHHHHHHH-----	9987178999999986027999

Table S2. List of active inhibitors.

PDB ID	Ligand Name	IC50 (nM)
3DV3	MEK	21
4MNE	573	35
4ARK	M3K	61*
4ANB	YQY	6.2
4AN9	2P7	44
4AN2	EUI	0.9
4AN3	5Y0	0.6*
3V01	3V0	25
3PP1	IZG	3.2
3OS3	3OS	235*
3ORN	3OR	8.8
3MBL	LSG	18.0
3EQH	5BM	60.0
3EQG	4BM	5.2
3E8N	VRA	19.0**
3DY7	1CX	2.0
3V04	V04	13.0
3VVH	4BM	5.2
4LMN	EUI	0.9

* Data from PDBbind; ** data from Binding MOAD; remaining data from bindingDB.