

Gatekeeper Mutations Activate FGF Receptor Tyrosine Kinases by Destabilizing the Autoinhibited State

Alida Besch¹, William M. Marsiglia^{1#}, Moosa Mohammadi²⁺, Yingkai Zhang^{1,3}, Nathaniel J.
Traaseth^{1*}*

¹Department of Chemistry, New York University, New York, NY, USA

²Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New
York, USA

³Simons Center for Computational Physical Chemistry at New York University, New York, NY, USA

[#]Current address: Department of Pharmacology and Toxicology, University of Alabama at Birmingham,
Birmingham, AL, USA

⁺Current address: Oujiang Laboratory (Zhejiang Lab for Regenerative Medicine, Vision and Brain Health), and
School of Pharmaceutical Sciences, Wenzhou Medical University, Wenzhou, Zhejiang 325000, China

* Corresponding authors: yingkai.zhang@nyu.edu, traaseth@nyu.edu

Abstract

Many types of human cancers are being treated with small molecule ATP-competitive inhibitors targeting the kinase domain of receptor tyrosine kinases. Despite initial successful remission, long term treatment almost inevitably leads to the emergence of drug resistance mutations at the gatekeeper residue hindering the access of the inhibitor to a hydrophobic pocket at the back of the ATP binding cleft. In addition to reducing drug efficacy, gatekeeper mutations elevate the intrinsic activity of the tyrosine kinase domain leading to more aggressive types of cancer. However, the mechanism of gain-of-function by gatekeeper mutations is poorly understood. Here, we characterized fibroblast growth factor receptor (FGFR) tyrosine kinases harboring two distinct gatekeeper mutations using kinase activity assays, NMR spectroscopy, bioinformatic analyses, and MD simulations. Our data show that gatekeeper mutations destabilize the autoinhibitory conformation of the DFG motif locally and of the kinase globally, suggesting they impart gain-of-function by facilitating the kinase's ability to populate the active state.

Significance Statement

Anti-cancer drugs inhibiting the kinase domain of receptor tyrosine kinases block downstream phosphorylation necessary for cellular signaling and growth. However, a major impediment to successful cancer therapy is the emergence of drug resistance mutations at a residue situated at the kinase hinge region between the N- and C-lobes which provides fitness advantages for tumor growth. In this work, we used kinase assays, NMR spectroscopy, and MD simulations to determine the mechanism of kinase activation from drug-resistant gatekeeper mutations. We found that these mutations destabilize the autoinhibited conformation of the kinase by weakening the hydrophobic spine and modulating the conformation of the conserved DFG motif, thereby leading to the conclusion that gatekeeper mutations increase in the population of the active state.

Introduction

Intra- and intercellular signaling by receptor tyrosine kinases (RTK) play essential roles in the regulation of cell growth, differentiation, metabolism, apoptosis, and senescence (1). Not surprisingly then, aberrant RTK signaling arising from amplifications, mutations, or their misexpression are implicated in a wide array of human developmental, metabolic and immune disorders, and cancer (2-4). Underscoring their importance in cancer, a recent study analyzing over 9000 tumor samples from 33 cancer types, found that alterations in genes encoding RTKs and constituents of the RAS signaling pathway dominate the landscape of altered genes in cancer (4). RTK gain-of-function mutations typically cluster within regulatory regions of the tyrosine kinase domain, including the juxtamembrane, activation loop (A-loop), and kinase hinge (5, 6). Epidermal growth factor receptor (EGFR) and fibroblast growth factor receptor (FGFR) are two of the most frequently mutated RTK families in human cancers (4). These mutations drive oncogenicity typically by subverting physiological mechanisms of ligand-induced kinase domain activation leading to uncontrolled cell growth and promotion of angiogenesis (1), key hallmarks of cancer. Accordingly, inhibition of the tyrosine kinase domain of RTKs via ATP-competitive small molecule inhibitors is the mainstay therapy for a multitude of solid and hematological cancers (6, 7).

Following the first successful kinase inhibitor Imatinib (Gleevec) (8, 9), there are now 71 FDA approved kinase drugs with 39 targeting RTKs involved in cancer (10). A major impediment to successful cancer therapy is the emergence of drug-resistant mutations at a residue within the kinase hinge region (11-14), which provides fitness advantages for tumor growth and hinders long-term remission (15). This location is often occupied by a threonine or valine, and in most instances, it becomes mutated to a larger hydrophobic residue most frequently isoleucine, methionine,

leucine, and phenylalanine (Fig. 1) (16). The bulkier side chains sterically block access of the inhibitor to the hydrophobic pocket at the back of the ATP binding cleft leading to poorer binding affinity and drug efficacy (14, 17-20). As a result, this residue is referred to as the *gatekeeper* in kinase jargon.

An underappreciated fact regarding gatekeeper mutations is that they elevate the kinase activity of the affected RTK. Understanding the molecular basis for this gain-of-function effect is important because the enhanced kinase activity of gatekeeper mutants correlates with greater signaling and worse patient outcomes (14, 21-23). It has been proposed that gatekeeper mutations encourage the active state of the kinase by strengthening the hydrophobic regulatory spine (R-spine) (24-26), a network of four hydrophobic residues connecting the N- and C-lobes of the kinase (Fig. 2a) (27). However, a couple of observations are inconsistent with this hypothesis. First, for many kinases, an intact hydrophobic R-spine occurs in both the autoinhibited and active states (28, 29). Second, certain gatekeeper mutations introduce hydrophilic residues, including in ALK (L1196Q), EPHA2 (T692N), FGFR3 (V555E), FGFR4 (V550E), and KIT (T670E) (16, 30) (Fig. 1). Although infrequent, such hydrophilic substitutions are difficult to reconcile with the hydrophobic R-spine hypothesis. Therefore, investigation of both hydrophobic and hydrophilic gatekeeper mutations would offer a comparative approach for elucidating mechanisms of activation.

Here we used a multipronged approach consisting of kinase assays, NMR spectroscopy, and MD simulations to determine the mechanism of RTK activation by drug-resistant gatekeeper mutations using FGFR2 kinase (FGFR2K) as the model RTK system. Our findings indicate that gatekeeper mutations destabilize the autoinhibited state of the kinase and modulate the

conformation of the conserved DFG motif, thereby leading to the proposed model where mutations shift the equilibrium to the active state of the kinase.

Results

Gatekeeper mutations elevate trans-autophosphorylation activity of FGFR2K

To assess the effects of gatekeeper mutations on FGFR2K activity, we compared kinase trans-autophosphorylation activities of V564I and V564E gatekeeper mutants with wild-type FGFR2K *in vitro*. V564I is an authentic FGFR2K mutation causing drug resistance (31) whereas the V564E mutation is analogous to the V550E mutation in the FGFR4 kinase domain detected in a childhood cancer (30). Each kinase sample was incubated with ATP and Mg^{2+} and analyzed for phosphorylation as a function of time using kinase mobility shifts in native gel electrophoresis and immunoblots with an antibody specific for A-loop phosphotyrosines (pY656/pY657; “2P” state). Note that kinase autophosphorylation occurs in an intermolecular manner (1), as depicted in our asymmetric dimer structure of FGFR kinase showing A-loop trans-autophosphorylation of Y656 (32). In our phosphorylation assays, both V564E and V564I gatekeeper mutants displayed an increased rate of intrinsic kinase activity as evident by earlier shifts in protein bands in native gels and earlier appearance of phosphorylated bands in immunoblots relative to wild-type FGFR2K (Fig. 2b). Quantification of the immunoblots revealed that gatekeeper mutants phosphorylated A-loop tyrosines ~3-4-fold faster than wild-type FGFR2K (compare 3 min time point in Fig. 2c). These results indicate that V564E and V564I gatekeeper mutants confer gain-of-function to wild-type FGFR2K *in vitro*.

Gatekeeper mutations enhance the conformational dynamics of FGFR2K when bound to ATP

Our kinase trans-autophosphorylation data imply that the gatekeeper mutations encourage the active state conformation of the kinase. To test this hypothesis, we applied Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion (33, 34) to characterize the conformational dynamics of the kinases in the absence of ATP and Mg^{2+} cofactor (i.e., apo state) on the μ sec-msec timescale. Wild-type FGFR2K and its gatekeeper mutants V564E and V564I were isotopically enriched with ^{13}C at methyl groups of isoleucine, leucine, and valine residues (i.e., ILV labeling). Surprisingly, CPMG results showed no significant relaxation dispersions for either wild-type FGFR2K or the gatekeeper mutants (SI Appendix, Fig. S1), indicating the absence of detectable higher energy conformations that convert with the lowest energy state on the μ sec-msec timescale.

We suspected that differences in conformational dynamics between wild-type FGFR2K and gatekeeper mutants may only be captured in the ATP/ Mg^{2+} bound state. Indeed, the nucleotide is essential for kinase function, as the triphosphate moiety serves as the energy source for phosphoryl transfer and the adenine base of ATP interacts with hydrophobic residues from the N- and C-lobes of the kinase to encourage formation of the catalytic spine (C-spine) necessary for kinase activation (35, 36). Hence, we repeated the CPMG experiments in the presence of ATP and Mg^{2+} . However, such an experiment was technically challenging since the kinase undergoes tyrosine trans-autophosphorylation in the presence of ATP. To overcome this problem, we engineered a non-phosphorylatable version of FGFR2K by mutating all five phosphorylatable tyrosines (Y466F/Y586L/Y588P/Y656F/Y657F; “0Y” nomenclature or WT^{0Y}). A $^1H/^{13}C$ HMQC spectral comparison of WT^{0Y} with wild-type FGFR2K showed only minor chemical shift perturbations near the mutation sites (SI Appendix, Fig. S2), which commonly arise from local changes in electrostatics introduced by the mutation. Since there were no large chemical shift perturbations in residues distant to the mutation sites, we inferred that the tyrosine mutations had

a minimal effect on the kinase structure. Accordingly, we prepared “0Y” versions of both gatekeeper mutants (V564E^{0Y} and V564I^{0Y}) for subsequent measurements.

Next, we performed CPMG relaxation dispersion experiments on WT^{0Y}, V564E^{0Y} and V564I^{0Y} mutants in the presence of saturating concentrations of ATP (20 mM) and Mg²⁺ (40 mM) as in the kinase assay (Fig. 3a, b; SI Appendix, Fig. S3). Due to the occurrence of ATP hydrolysis at longer incubation times (SI Appendix, Fig. S4), we reduced the total experimental time by collecting CPMG relaxation dispersion data at 50 Hz and 1000 Hz frequencies and at two magnetic field strengths. We found that gatekeeper mutants displayed significantly larger ΔR_2 values for residues near the gatekeeper residue (L550, L551), within or proximal to the α C helix (L528, L531, V562), the A-loop (I654), near the substrate binding site (L692, I707, L716), and within or proximal to the hydrophobic R-spine (L550, V562, I563, I623, I642) (Fig. 3b). Each of these regions are known to undergo conformational changes important in kinase activation, including the A-loop which blocks substrate access to the enzyme’s active site in the autoinhibited conformation (37).

The increased activity and ΔR_2 values for V564E^{0Y} and V564I^{0Y} mutants relative to WT^{0Y} suggested a population shift away from the major state of WT^{0Y} (see Methods section *CPMG data interpretation*). Thus, it became imperative to define this major conformation of WT^{0Y} when bound to nucleotide and Mg²⁺ (i.e., autoinhibited or active state). To do so, we designed a paramagnetic relaxation enhancement (PRE) experiment to measure distances between an MTSL spin label attached to a cysteine in the C-lobe (I707C) and tyrosine residues within the A-loop. Based on FGFR kinase crystal structures (PDB IDs: 3KY2 and 2PVF), Y656 and Y657 are within ~20 Å from the MTSL spin label when in the autoinhibited kinase conformation and >30 Å when in the active kinase conformation. Hence, the sensitivity of the PRE experiment (≤ 25 Å) would

be informative for deciphering the major state. $^1\text{H}/^{15}\text{N}$ HSQC spectra of ^{15}N -Tyr and MTSL labeled kinase were collected for the oxidized and reduced states in the presence of non-hydrolyzable nucleotide AMP-PCP and Mg^{2+} (SI Appendix, Fig. S5). Note that the rather long data collection times precluded the use of native ATP due to the occurrence of ATP hydrolysis. HSQC spectra for Y656 and Y657 within the A-loop displayed decreased intensities in the oxidized spectrum relative to the reduced dataset and corresponded to distances of ~ 20 Å from the spin label. This finding was consistent with the autoinhibited state and provided evidence that $\text{WT}^{0\text{Y}}$ in complex with nucleotide is primarily in the autoinhibited conformation.

These PRE findings taken together with results from CPMG relaxation dispersion experiments and activity assays support a model where gatekeeper mutations induce a population shift toward the active state of the kinase. Although the two-point CPMG approach used in this work does not directly identify the minor state as the kinase active conformation, these results do indicate a population shift away from the predominant autoinhibited conformation characterized in the PRE experiment. Thus, the correlation of increased CPMG relaxation dispersion with increased activity for gatekeeper mutants relative to wild-type FGFR2K supports our activation hypothesis where gatekeeper mutants induce an equilibrium shift toward the active conformation. This activation model is also consistent with FGFR kinase crystal structures which show the enzyme primarily crystallizes in two states, autoinhibited and active conformations (29).

ATP binding increases the conformational heterogeneity of gatekeeper mutants

Notwithstanding the insights gained from CPMG relaxation dispersion data, some residues near the ATP binding pocket could not be analyzed due to low signal intensities induced by the addition of ATP and Mg^{2+} (see purple spheres in Fig. 3b). These reductions in peak intensities

further supported the CPMG results that ATP binding induced conformational dynamics on the μ sec-msec timescale (38). Notably, these effects were independent of the kinase concentration (SI Appendix, Fig. S6), indicating that the reduced signal intensities were not due to kinase dimerization, as we previously reported for a gain-of-function mutation in the P+1 pocket of FGFR kinase (32). Thus, to complement CPMG data, we quantitatively compared $^1\text{H}/^{13}\text{C}$ HMQC peak intensities of V564I^{0Y} and V564E^{0Y} with WT^{0Y} in the absence and presence of saturating concentrations of ATP (20 mM) (Fig. 3a, c).

Mapping peak intensity changes onto the crystal structure revealed significantly greater reductions for V564E^{0Y} and V564I^{0Y} relative to WT^{0Y} in key regulatory regions of the kinase, including the R- and C-spines, DFG motif, and A-loop. Peaks corresponding to I547, I548, and L647 surrounding the DFG motif were completely absent for V564E^{0Y} and V564I^{0Y} but were only reduced in intensity for WT^{0Y}. C-lobe residues of the C-spine also displayed larger intensity reductions for the gatekeeper mutants than for WT^{0Y} (V632, L633, V634, L692, L696) (39). This observation is consistent with the role of ATP in promoting the hydrophobic C-spine and priming the active site for substrate binding and catalysis (40, 41). We also observed notable reductions in intensity for L550 in V564I^{0Y}, one of the residues within the R-spine, as well as other residues nearby the spine for V564I^{0Y} and V564E^{0Y} (L551, V562, I563, I564 of V564I only, I623, I642). In addition, peak intensity for I651 located in the A-loop and 20 Å away from the ATP binding site, decreased by ~80% for gatekeeper mutants (83.6% for V564I^{0Y} and 76.8% for V564E^{0Y}) compared to 60% for WT^{0Y}. These reductions/ablations in peak intensities for V564E^{0Y} and V564I^{0Y} observed in $^1\text{H}/^{13}\text{C}$ HMQC experiments indicated that the gatekeeper mutants are conformationally more heterogeneous than WT^{0Y} when bound to ATP. From these data, we conclude that gatekeeper mutations induce perturbations within the hydrophobic spine and A-loop,

regions both involved in the conformational change from the autoinhibited to the active state. This transition involves the C-terminal end of the A-loop moving away from its autoinhibitory position in the substrate binding pocket to enable the N-terminal portion of the A-loop to form new contacts with the α C helix (α C tether) and the catalytic loop (29).

Gatekeeper mutants bind ATP tighter than FGFR2K

We also performed NMR titration experiments to quantify binding affinities of the resulting WT^{0Y}, V564I^{0Y}, and V564E^{0Y} kinases using the native ATP nucleotide and Mg²⁺ cofactor (Fig. 4a; SI Appendix, Fig. S7). Fitting chemical shift perturbation data revealed K_d values of $450 \pm 70 \mu\text{M}$ for WT^{0Y}, $280 \pm 30 \mu\text{M}$ for V564I^{0Y}, and $140 \pm 20 \mu\text{M}$ for V564E^{0Y} (Fig. 4b). The 1.6-fold and 3.2-fold improvements in ATP binding affinities of V564I^{0Y} and V564E^{0Y} relative to WT^{0Y} implies that conformations of gatekeeper mutants are differentially affected upon binding to ATP/Mg²⁺ relative to unphosphorylated FGFR2K. Notably, the 20 mM ATP concentration used for phosphorylation assays is at least 44-fold above the K_d values; therefore, the elevated kinase activity of the gatekeeper mutants reported in Fig. 2 cannot be attributed to the improved ATP binding affinities of the gatekeeper mutants. Furthermore, the ATP concentration in cells is typically 1 mM and above (i.e., higher than the K_d values) (42), so ATP is expected to be bound to wild-type and mutants. Finally, we also measured the binding constant of FGFR2K to the non-hydrolyzable nucleotide AMP-PCP in the presence of Mg²⁺ and observed weaker binding ($K_d = 2.9 \pm 0.1 \text{ mM}$; SI Appendix, Fig. S8), which has been reported with AMP-PNP (43). This divergence from our K_d value determined for ATP as well as reported K_m values for ATP (70 to 840 μM) (24, 44, 45) corroborated our experimental approach to assess FGFR2K conformational dynamics using the native ATP nucleotide.

MD simulations reveal gatekeeper mutants destabilize the autoinhibited state

To complement the NMR experiments, we next performed MD simulations on the ATP/Mg²⁺ bound states of wild-type FGFR2K (unphosphorylated and phosphorylated forms) and the V564I and V564E gatekeeper mutants for a total of 5 μ s (five replicates) or 3 μ s (three replicates) starting from the autoinhibited or active conformation, respectively. We first focused on the hydrophobic R-spine due to its proposed role in kinase activation (27, 28) and our NMR observations on the gatekeeper mutants for this region. We quantified whether the hydrophobic R-spine was intact as defined as side chain contact distances less than 4.5 Å among the four constituents of the spine (M538, L550, H624, F645). The fraction of intact R-spine was calculated every 5 psec of the MD simulations and plotted as a function of each replicate (Fig. 5). MD simulations starting from the active conformation showed a similar extent of intact R-spine formed for wild-type FGFR2Ks and gatekeeper mutants (Fig. 5a, b). Next, we performed the same analysis on simulations starting from the autoinhibited conformation. The unphosphorylated FGFR2K possessed an intact R-spine for all replicates (Fig. 5c, d). In contrast, the V564I and V564E gatekeeper mutants showed a broken R-spine due to the disruption of contacts between M538 and F645 in the DFG motif (Fig. 5d). Likewise, we observed disruption of the hydrophobic R-spine in the phosphorylated FGFR2K simulations, albeit through disruption of contacts between H624 and F645. Notably, our simulations displayed a comparable effect for hydrophilic V564E or hydrophobic V564I gatekeeper mutations on the hydrophobic R-spine. We also analyzed whether the hydrophobic C-spine was intact in active and autoinhibited simulations, as defined as contact distances less than 4.5 Å among the adenine base of ATP and the eight residues encompassing the spine (V495, A515, L572, V632, L633, V634, L692, I696). We found the C-

spine to be intact in both wild-type FGFR2K and gatekeeper mutant simulations starting from the active state but slightly destabilized for V564E and phosphorylated wild-type in simulations starting from the autoinhibited state (SI Appendix, Fig. S9). From these data, we conclude that the gatekeeper mutations confer gain-of-function by destabilizing the autoinhibited state, most significantly at the hydrophobic R-spine. Notably, this conclusion stands at odds with the accepted view in the field that gatekeeper mutations act by strengthening the hydrophobic R-spine purported to be a unique hallmark of the active conformation (24-26).

Mechanism of autoinhibited state destabilization

To globally analyze MD simulations of wild-type FGFR2K and gatekeeper mutants, we derived a computational approach for classifying autoinhibited and active conformations through the analysis of 52 crystal structures from all four FGFR kinase isoforms. A matrix of the minimum residue distance for 192 conserved residues among the isoforms was calculated and subjected to principal component analysis (PCA) (46). Two distinct clusters emerged following transformation of the distance matrix onto the first and second principal components (Fig. 6a). Structures within these two clusters correlated to autoinhibited and active states of FGFR kinases, as concluded from phosphorylation assays (29, 47) (SI Appendix, Table S1). For instance, unphosphorylated FGFR kinases (e.g., 3KY2 and 1FGK) defined one cluster, while A-loop phosphorylated FGFR kinases (e.g., 2PVF) and gain-of-function mutants in the A-loop and/or hinge region (e.g., 4J97 and 5UI0) comprised a second cluster. Two out of the 52 PDBs analyzed did not fall in either cluster: 6PNX and 4TYG. The former structure is an asymmetric dimer complex of the A-loop transphosphorylation complex where the monomers correspond to the substrate-acting (6PNX^A) and enzyme-acting (6PNX^B) kinases (32). The enzyme-acting kinase was in close vicinity to the

active state cluster, in agreement with its role as the enzyme in A-loop phosphorylation, while the substrate-acting kinase significantly differed from both clusters. The 4TYG structure of FGFR4 kinase also significantly differed from both clusters, which likely stemmed from the inserted position of the A-loop into its own ATP binding pocket (48). Such an insertion has not been observed in other FGFR kinase structures. Based on this analysis, we conclude that 4TYG and the substrate-acting kinase of 6PNX (6PNX^A) represent unique conformations that do not resemble either the autoinhibited or active conformations.

Using the classification of autoinhibited and active clusters, we elucidated a set of distances that differed among kinase conformations and offered an unbiased and quantitative view of residue pairs experiencing distance changes upon activation. Namely, we found 20 *active contacts* formed in the active state but disrupted in the autoinhibited state and 22 *autoinhibited contacts* formed in the autoinhibited state but disrupted in the active state (Fig. 6b; SI Appendix, Table S2; see Methods for details). MD simulations were analyzed by constructing a heat map for each residue contact, where red corresponds to an intact contact (< 4.5 Å) and blue corresponds to a disrupted contact (> 4.5 Å) (Fig. 7a; SI Appendix, Fig. S10). The heat map representation indicates that our simulations do not capture the complete interconversion between autoinhibited and active conformations defined by the cluster analysis. Despite this time sampling limitation, several residue contacts displayed significant differences between wild-type and gatekeeper mutants. Simulations of the gatekeeper mutants starting from the autoinhibited conformation displayed a greater probability of disrupting autoinhibited contacts and forming active contacts (Fig. 7b-d). These contacts involve key residues in the $\beta 3$ strand (K517), αC helix (E530), DFG motif (D644, F645), A-loop (L647, A648), and catalytic loop (I623, H624, R625, N631). For example, in the gatekeeper mutant simulations, the DFG+1 and DFG+2 residues (L647, A648) had a higher

probability of forming active contacts with residues in the catalytic loop (I623, H624, R625) (Fig. 7b, d). Likewise, autoinhibited distances were disrupted for gatekeeper mutants, which remained intact for the wild-type kinase, including the DFG+1 residue (L647) with the aspartate of the DFG motif (D644) and the conserved lysine involved in ATP coordination (K517) (Fig. 7c). The DFG+2 residue (A648) also displayed a reduced probability to form an autoinhibited contact with the α C helix (E530).

In contrast to simulations starting from the autoinhibited conformation, we observed more similarities among wild-type and gatekeeper mutations for simulations starting from the active conformation (Fig. 7a; SI Appendix, Fig. S10). The heat map showed some disrupted active contacts with the greatest differences at residue pairs near the phosphorylated tyrosines within the C-lobe (R664 and R678) that forms interactions with the A-loop (L647, K659, and N662). It is likely the active state differences arise from the reduced stability of the A-loop conformation for unphosphorylated tyrosines in the A-loop relative to the phosphorylated forms. Nevertheless, these differences were relatively minor compared to those observed in simulations starting from the autoinhibited conformation. Overall, the analysis of MD simulations implies that gatekeeper mutations destabilize the autoinhibited conformation through structural changes primarily impacting the DFG, DFG+1, and DFG+2 residues without directly interacting with the phenylalanine residue from the DFG motif (SI Appendix, Fig. S11). These results are consistent with our hydrophobic spine analysis indicating a destabilization of the autoinhibited conformation. We propose that gatekeeper mutations disrupt interactions within the autoinhibited conformation, which induces a population shift to the active state such that the kinase occupies a catalytically competent conformation a greater fraction of time and leads to gain-of-function.

Discussion

Stabilization of the hydrophobic R-spine in the active conformation has been proposed to underlie the gain-of-function of gatekeeper mutations in the Abl- and Src-family tyrosine kinases (25, 26). This mechanism rests on the idea that an intact hydrophobic R-spine is a unique feature of the active state of tyrosine kinases. Contrary to this model, we discovered that hydrophobic and hydrophilic gatekeeper mutations of FGFR kinase manifest their gain-of-function by destabilizing the autoinhibited state. Unlike gatekeeper mutants, unphosphorylated FGFR2K displayed an intact hydrophobic spine in both the autoinhibited and active states. These findings lead us to propose that the observed gain-of-function for gatekeeper mutations occurs from a population shift away from the autoinhibited state toward the active state.

We hypothesize the mechanistic difference between Abl- and Src-family kinases and FGFR-like kinases stems from the ability of the former kinases to sample the DFG^{out} conformation (26, 28), which is an inactive state typically observed with type II inhibitors and one that possesses a broken hydrophobic R-spine. The DFG^{out} conformation has not been observed for FGFR kinase and other RTKs in the absence of drugs (28). Instead, these kinases possess two distinct DFGⁱⁿ conformations, where one conformer corresponds to an active state (DFG^{in, active}) and the other to an autoinhibited state (DFG^{in, inhibited}) (28, 29). Notably, we found that the χ_1 angle of the DFG motif phenylalanine residue strongly correlated with the active and autoinhibited states identified from our cluster analysis (Fig. 8a) (29, 37). Simulations of the autoinhibited conformation displayed a greater probability for gatekeeper mutations to sample the DFG^{in, active} conformation relative to wild-type FGFR kinase (Fig. 8b). Hence, we propose that the mechanism of gatekeeper mutations is to create steric repulsion forces with the hydrophobic R-spine at M538 in the autoinhibited state that influence the conformation of the DFG motif phenylalanine. Based on this

hypothesis, any mutation at the gatekeeper position, hydrophobic or hydrophilic, that destabilizes the hydrophobic R-spine may induce activation through a similar mechanism as the isoleucine and glutamate mutations studied in this work. Notably, mutation of the gatekeeper position to phenylalanine, somewhat common in FGFR kinases (13, 49) and RTKs in general (Fig. 1), is likely to increase steric interference with M538 and lead to activation. These steric interactions likely weaken the hydrophobic R-spine in the autoinhibited conformation by introducing conformational heterogeneity and enable the DFG phenylalanine to sample the active state conformation more easily (i.e., DFG^{in, active}) (Fig. 8c). The DFG^{in, active} conformation is more compatible with rotation of the N-lobe, a key step for transitioning to the active state (35).

We also found that mutations at the gatekeeper position influence the conformation of the DFG+1 and DFG+2 residues within the autoinhibited state. Namely, contacts between DFG+1 and DFG+2 residues and the α C-helix were seen in simulations for gatekeeper mutations but not those with wild-type. We envision these interactions encourage N-lobe rotation downward and the transition to the active state. Such interactions between the N-terminal portion of the A-loop and the α C-helix are reminiscent of the α C-tether we previously reported as a determinant of activated kinases (29). Our findings also concur with the importance of the DFG+1 residue in substrate specificity and allostery by modulating selectivity between serine and threonine containing substrates in Ser/Thr kinases (50). Furthermore, this residue is responsible for the most common oncogenic mutation (L858R) in EGFR that induces a ~50-fold increase in catalytic efficiency and is predicted to destabilize the inactive state by disrupting autoinhibitory interactions (51) and encourage the formation of kinase dimers (52, 53).

We surmise that the conformational changes induced by gatekeeper mutations at the R-spine and DFG motif occur on a relatively fast timescale (nsec- μ sec), while the global transition

from the autoinhibited state to the active state occurs on a slower timescale (μsec - msec) (Fig. 8c). Destabilization of the autoinhibited state and subsequent conformational changes at the DFG motif phenylalanine likely serves as the underlying source of fast motions that promotes N-lobe rotation and A-loop movement to the active conformation. This hypothesis is corroborated by our MD simulations capturing local conformational changes induced by gatekeeper mutations on the nsec - μsec timescale around the DFG motif and NMR experiments sensing dynamics at the DFG motif and A-loop on the μsec - msec timescale. Based on the similarity of FGFR kinase with other RTKs (SI Appendix, Fig. S12), such as PDGF, VEGF, KIT, and c-Met receptors, the mechanism discovered for FGFR kinase may be applicable to gatekeeper mutations in other tyrosine kinases to subvert autoinhibitory mechanisms, leading to enhanced phosphorylation activity and increased clinical severity for patients possessing such mutations.

Methods

Protein expression and purification

FGFR2K domain constructs, corresponding to P458 to E768 and with the C491A mutation, were expressed with an N-terminal His₆-tag to aid in protein purification. The C491A mutation of FGFR2K is referred to as wild-type throughout the text. The “0Y” construct of FGFR2K was prepared by mutating all five phosphorylatable tyrosines in the protein (i.e., Y466F/Y586L/Y588P/Y656F/Y657F). Note that Y586L and Y588P mutations mimic native positions in the FGFR4 isoform. Kinases were expressed in *E. coli* BL21(DE3) cells and induced with 0.1 mM IPTG overnight at 20 °C. Protein was purified using Ni-NTA resin and size exclusion chromatography. The N-terminal His₆-tag was cleaved using TEV protease and phosphorylated states were removed using FastAP alkaline phosphatase (Thermo Scientific). Anion exchange chromatography was performed to ensure purification of the unphosphorylated state. For NMR experiments, protein was expressed in minimal media with ¹⁵NH₄Cl, deuterated glucose, 99% D₂O, and ILV precursors, as previously described (29).

Activity assay of gatekeeper mutants

The activity assay reaction mixture consisted of 50 μM FGFR2K, 20 mM ATP, 40 mM MgCl₂, 20 mM Tris pH 7.5, and 150 mM NaCl. The addition of ATP initiated transphosphorylation on five tyrosine residues within FGFR2K, including the tyrosine residues within the A-loop. Reactions were quenched at 0, 0.5, 1, 3, 5, 10 and 20 minutes by adding EDTA to a final concentration of 50 mM. The transphosphorylation reaction was observed with native gel electrophoresis using a 7.5% acrylamide gel with a tris-glycine running buffer (25mM Tris, 200mM glycine, pH 8.3). The same reactions were analyzed using immunoblot analysis by

running samples on an SDS-PAGE gel followed by transfer to a nitrocellulose membrane. The membrane was blocked with 5% non-fat dry milk in PBS buffer at pH 7.4 for 1 hour at 37 °C and subsequently for 30 min at room temperature. Phosphorylation was then detected using an anti-phosphorylation-FGFR antibody for phosphorylation of A-loop tyrosines (Cell Signaling Technologies, part number 3471) by incubating the antibody overnight at 4 °C. Following incubation, the membrane was washed and incubated with anti-rabbit secondary antibody for 2 hr at 37 °C. Membranes were imaged for chemiluminescence by addition of horseradish peroxidase (HRP) enzyme.

ATPase activity assay of gatekeeper mutants

ATP hydrolysis rates were derived from ^{31}P one-dimensional NMR spectra. Each experiment was performed using a 15 μsec excitation pulse, 300 msec for acquisition (spectral width of 32467.533 Hz), and a 2.5 sec recycle delay. The total length of a single experiment was 5 min (160 scans). The kinetics of ATP hydrolysis was captured over a total time of 11.75 hours by collecting a series of experiments one after another. Each sample of WT^{0Y}, V564I^{0Y}, and V564E^{0Y} consisted of 600 μM kinase in the presence of 10 mM ATP and 20 mM MgCl_2 . Quantification of the time series was performed by integrating the γ -phosphate peak of ATP at each time point.

CPMG relaxation dispersion experiments

Methyl $^1\text{H}/^{13}\text{C}$ CPMG experiments (54) were performed using Bruker AVANCE III and AVANCE NEO NMR spectrometers operating at ^1H frequencies of 600 MHz and 800 MHz, respectively, each equipped with a TCI cryogenic probe. Data were acquired using a constant time

delay of 40 msec with frequencies for the 180° pulses of 50 and 1,000 Hz. ΔR_2 values were calculated using equation 1, where I_{1000} is the peak intensity at 1,000 Hz, I_{50} is the peak intensity at 50 Hz, and T is the constant time delay.

$$\Delta R_2 = \ln \left(\frac{I_{1000}}{I_{50}} \right) \frac{1}{T} \quad (1)$$

ΔR_2 values and errors reported reflect the average and standard deviation between two replicate experiments. CPMG experiments on the apo form of FGFR2K were performed at 10 °C with a protein concentration of 500 μ M in 25 mM HEPES pH 7.5, 150 mM NaCl, and 10% D₂O. For ATP bound experiments, 20 mM ATP and 40 mM MgCl₂ were used to saturate the kinase. In these experiments, ATP hydrolysis was observed and therefore a relatively low concentration of kinase was used (50 μ M) to slow the reaction while maintaining sufficient signal-to-noise for data interpretation. Experiments with ATP were performed at 25 °C with 16 scans and 60 complex points in indirect dimension. We quantified the amount of ATP hydrolysis that occurred before and after the experiment using ³¹P NMR. The highest amount of hydrolysis occurred with gatekeeper mutants, corresponding to 68% of the starting ATP concentration for V564E^{0Y}, 64%-71% for V564I^{0Y}, and 84%-89% for WT^{0Y}.

CPMG data interpretation

CPMG relaxation dispersion data are described by the following equation for two-site chemical exchange (states “A” and “B”) (55, 56):

$$R_{2eff} = \frac{R_2^A + R_2^B + k_{ex}}{2} - \frac{1}{2\tau_{cp}} \cosh^{-1} [D_+ \cosh \eta_+ - D_- \cosh \eta_-] \quad (2)$$

$$D_{\pm} = \frac{1}{2} \left[\pm 1 + \frac{\psi + 2\Delta\omega^2}{\sqrt{\psi^2 + \zeta^2}} \right] \quad \eta_{\pm} = \frac{\tau_{cp}}{\sqrt{2}} \left[\pm \psi + \sqrt{\psi^2 + \zeta^2} \right]^{1/2}$$

$$\psi = k_{ex}^2 - \Delta\omega^2 \quad \zeta = -2 \Delta\omega k_{ex} (p_A - p_B)$$

p_A and p_B are the populations of A and B, k_{ex} is the exchange rate which is equal to the forward (k_1) and reverse rate (k_{-1}) constants describing the equilibrium between states A and B, $\Delta\omega$ is the difference in chemical shifts between the two states, R_2^A and R_2^B are the transverse relaxation rates for states A and B, and τ_{cp} is the time between successive 180° refocusing pulses. This equation assumes the difference between R_2^A and R_2^B is negligible relative to the difference between k_1 and k_{-1} .

Gatekeeper mutants displayed larger ΔR_2 values calculated using equation 1 relative to the WT^{0Y} sample. There are three reasons (not mutually exclusive) why these ΔR_2 differences could arise according to equation 2: (1) change in $\Delta\omega$, (2) change in k_{ex} or (3) changes in p_A and p_B . For the first case, we anticipate that $\Delta\omega$ values will be the same for wild-type and gatekeeper mutants for residues not near the site of mutation. Indeed, many of the key changes we observed are greater than 10 Å from the gatekeeper position. We consider this a somewhat safe assumption considering FGFR kinases almost exclusively crystallize in either the autoinhibited or active conformation (47). Indeed, our quantitative and unbiased clustering analysis in Fig. 6 displays two clusters representing 52 FGFR kinase crystal structures of wild-type and mutated variants. In other words, the active state (or autoinhibited state) is essentially identical whether a mutation is present or not and therefore is expected to have the same $\Delta\omega$ for wild-type or mutants. The second possible reason for a change in ΔR_2 values is a change in k_{ex} . However, since k_{ex} is a sum of k_I and k_{-I} , the only way k_{ex} can change without changing populations is if both k_I and k_{-I} change by exactly the same factor. This narrow condition is very unlikely and would not explain the increased activity of gatekeeper mutants relative to wild-type displayed in Fig. 2. In the more likely case that k_{ex} changes, which stems from a change in k_I and/or k_{-I} , this would result in a change in populations. Therefore, based on these arguments and the increased activity of gatekeeper mutants, we conclude

that the increased ΔR_2 values for gatekeeper mutants correspond to a change in the population relative to wild-type.

Intensity retention NMR experiments

Intensity retention experiments were performed on ILV labeled kinases using a Bruker AVANCE NEO NMR spectrometer operating at a ^1H frequency of 800 MHz equipped with a TCI cryogenic probe. $^1\text{H}/^{13}\text{C}$ HMQC spectra were collected at 25 °C on 50 μM and 250 μM samples of WT^{0Y}, V564I^{0Y}, and V564E^{0Y} in 25 mM HEPES pH 7.5, 150 mM NaCl, and 40 mM MgCl_2 in the presence and absence of 20 mM ATP. Intensity retention values were calculated as the peak intensity in the ATP/ Mg^{2+} bound state divided by the intensity in the absence of ATP. $^1\text{H}/^{13}\text{C}$ HMQC spectra for 50 μM samples in the absence of ATP were collected with 8 scans, while spectra in the presence of 20 mM ATP were collected with 12 scans to account for the dilution upon nucleotide addition (50 μL of nucleotide into 200 μL kinase sample). $^1\text{H}/^{13}\text{C}$ HMQC spectra for 250 μM samples in the presence and absence of ATP were each collected with 4 scans. To compensate for the dilution factor and/or scan difference in HMQC spectra, peak heights of kinase samples at 250 μM in the presence of ATP were multiplied by 1.25, while peak heights of kinase samples at 50 μM in the absence of ATP were multiplied by 1.2. The peak height multiplications were done in an identical manner for WT^{0Y}, V564I^{0Y}, and V564E^{0Y}.

Paramagnetic relaxation enhancement experiment

A ^{15}N tyrosine selectively labeled sample was prepared using the I707C mutant of FGFR2K, as described previously (29). Prior to labeling, the kinase was treated with 20 mM DTT for 15 min in 20 mM Tris pH 7.5 and 150 mM NaCl. DTT was buffer exchanged, and the pH was

adjusted to 6.5. MTSL (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate) was added in half-molar equivalents relative to the kinase concentration (32 μ M) and incubated on ice for 10 minutes at each of two total additions. The progress of the reaction was monitored using a Bruker ultrafleXtremeTM MALDI-TOF. The sample was buffer exchanged to 25 mM HEPES pH 7.5 and 150 mM NaCl for NMR experiments. The kinase sample concentration for NMR experiments was 106 μ M and included the addition of 10 mM AMP-PCP, 20 mM MgCl₂, and 10% D₂O.

¹H/¹⁵N HSQC experiments were collected for MTSL labeled I707C FGFR2K in the oxidized and reduced forms using a Bruker AVANCE NEO NMR spectrometer operating at a ¹H frequency of 800 MHz equipped with a TCI cryogenic probe. The oxidized sample was obtained by treating the reduced sample with 10 molar equivalents of ascorbic acid relative to the kinase concentration. ¹H/¹⁵N HSQC experiments were performed with 704 scans, spectral widths of 12,500 Hz (¹H) and 1,469.734 Hz (¹⁵N), acquisition time of 81.92 msec, 24 complex points in the indirect dimension, and a recycle delay of 4 sec. ¹H/¹⁵N HSQC intensity retentions were calculated by dividing peak heights of the oxidized HSQC spectrum over that of the reduced spectrum. Distances were calculated as previously described using a correlation time of 22.2 ns (57, 58).

NMR nucleotide titrations

ATP titrations were performed at 25 °C on WT^{0Y}, V564I^{0Y}, and V564E^{0Y} using a Bruker AVANCE NEO NMR spectrometer operating at a ¹H frequency of 800 MHz equipped with a TCI cryogenic probe. ¹H/¹³C HMQC experiments were acquired on ILV labeled kinases with 4 scans and 60 complex points in the indirect dimension to minimize the experiment time due to ATP hydrolysis. The entire titration was acquired in ~3 hr. Titrations were carried out with 50 μ M

kinase in 25 mM HEPES pH 7.5, 150 mM NaCl, and 40 mM MgCl₂. ATP was titrated at several concentrations ranging from low μ M to mM, as displayed in the figure legends. The AMP-PCP titration was performed at 10 °C on FGFR2K using a Bruker AVANCE III NMR spectrometer operating at a ¹H frequency of 600 MHz equipped with a TCI cryogenic probe. ¹H/¹³C HMQC experiments were performed on 533 μ M ILV labeled kinases with 2 scans and 100 complex points in the indirect dimension. Binding constants were fit using a non-linear least-squares regression as shown in equation 3.

$$\Delta\delta_{obs} = \Delta\delta_{max} \frac{([P]_t + [L]_t + K_d) - \sqrt{([P]_t + [L]_t + K_d)^2 - 4[P]_t[L]_t}}{2[P]_t} \quad (3)$$

FGFR kinase structural analysis

52 crystal structures that correspond to each of the four FGFR isoforms were analyzed: 32 FGFR1, 9 FGFR2, 3 FGFR3, and 8 FGFR4. PDB's with unresolved A-loops were removed from the analysis as the A-loop was found to be an important part of the classification. Of the 208 conserved residues among all FGFR isoforms of the kinase domain, 192 of these residues were used in the analysis since some regions were consistently not resolved in crystal structures (e.g., P-loop and kinase insert). Using the 192 residues, a distance matrix of all PDBs was calculated by using the minimum residue distance (MRD). Principal component analysis was then used on MRD of the 52 high-resolution crystal structures and plotted according to the PC1 and PC2. The first two principal components (PCs) accounted for 66% and 7% of the total variance. From there, PDBs were clustered by the PC1 and PC2 using DBSCAN clustering algorithm (59) using Scikit-learn (60).

To determine the distances that account for the variance between clusters, a *t-score* was calculated as shown in equation 4, where d_{ij}^i and s_{ij}^i represent the average distance pair and the standard deviation of the distance within the cluster, respectively.

$$t\text{-score} = \frac{d_{ij}^1 - d_{ij}^2}{\sqrt{s_{ij}^1 + s_{ij}^2}} \quad (4)$$

The *t-score* was then scaled (*scaled t-score*) by the minimum of the distance pair (d_{ij}) to enhance the *t-score* of smaller distances.

$$\text{scaled } t\text{-score} = \frac{t\text{-score}}{\min(d_{ij})} \quad (5)$$

All distance pairs with a scaled *t-score* greater than 3.5 Å in magnitude and a minimum distance less than 4.5 Å were extracted. This yielded 43 distances formed in the active state and 45 distances formed in the autoinhibited state. Next, distance pairs were filtered to represent unique contacts formed in the active and autoinhibited states. The first filtering was performed to remove long-range distances by removing distances that have an average PDB distance greater than 5 Å. The next filtering removed distances shared in active and autoinhibited states by removing average distances between the clusters that varied by 1 Å or less. Lastly, distances were filtered based on the stability of the contact in MD simulation. The stability of active and autoinhibited contacts was determined by the MD simulations starting from 2PVF and the FGFR2K homology model of 3KY2, respectively. Distances were removed if the percent contact formed was less than 25%, where a contact is defined as formed if the minimum residue heavy atom distance is less than 4.5 Å. Active distances removed include L647-L665, L647-P666, R625-L665, and K658-D677. Autoinhibited distances removed include T660-L665, R664-S702, R630-T660, R664-E695, R573-N662, R573-R664. This yielded 20 *active contacts* formed in the active state but disrupted in the autoinhibited state and 22 *autoinhibited contacts* formed in the autoinhibited state but disrupted in

the active state (SI Appendix, Table S2). These contacts represented ~0.5% of the total number of conserved residue pairs

MD simulations

Molecular dynamic simulations were performed starting from autoinhibited and active states. The active state was represented by an FGFR2K structure with A-loop tyrosines phosphorylated with ATP analog and substrate peptide (PDB ID: 2PVF (45)). The inactive state was modelled from a homology model of FGFR1 kinase (PDB ID: 3KY2 (61)) using Rosetta comparative homology modelling (62, 63) since there are no FGFR2K crystal structures in the authentically autoinhibited conformation. Simulations were performed on wild-type, V564I, V564E, and phosphorylated state (3P) with ATP and Mg^{2+} bound and by removing the substrate peptide from PDB ID 2PVF. The phosphorylated state was simulated with pY656, pY657, pY586 to mimic the phosphorylation state in the crystal structure (PDB ID: 2PVF). The protein was solvated with TIP3P water with a 15 Å buffer region and neutralized with Na^+ (e.g., 8 ions for wild-type, 9 ions for V564E, and 14 ions for 3P). For ATP bound simulations, parameters for ATP and Mg^{2+} were implemented from a previously parameterized system and coordinates were modeled from PDB ID 2PVF (45, 64, 65). Simulations were built in either Amber 16.06 or Amber 20.11 and ran in either Gromacs 5.1.4 and Gromacs 2020.4 for 2PVF and 3KY2, respectively. Parameter files were converted from Amber to Gromacs using ACPYPE (66). All simulations were performed using the *amber14sb* forcefield (67). Hydrogen bonds to heavy atoms were constrained using the LINCS algorithm with a 2 fs timestep (68). The van der Waals interactions were switched off from 1.0 to 1.2 nm. Electrostatic interactions were treated with particle-mesh Ewald with a cutoff of 1.2 nm (69). A steepest descent minimization was completed until the

maximum force was below $1,000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ with heavy atoms constrained. The system was equilibrated for 1 ns under constant volume and temperature with the heavy atoms constrained. Following this equilibration, a constant pressure and temperature simulation was performed where the maximum force on heavy atoms of the protein was reduced from $1,000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ to $0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ over 3 ns. From there, 30 ns of unconstrained MD simulations were performed prior to the production runs. For simulations starting from the autoinhibited state, heavy atoms of ATP were released prior to the release of force constants of protein because ATP was modelled into the structure. Temperature was maintained at 300 K using the Nose-Hoover thermostat for protein and solvent separately (70). Pressure was maintained at 1 atm using the Parrinello-Rahman coupling (71). Simulations were repeated 3 or 5 times for 1 μs each saving coordinates every 5 ps for the analyses. Mutant structures were obtained using the Dunbrack rotamer library implemented in Chimera (72). Loops missing in PDB ID 2PVF were modelled in using Modeler (73).

Acknowledgements

This work was supported by NIH (R01 GM117118, R01 AI108889) and NSF (MCB 1902449) awards to N.J.T. and NIH award (R35 GM127040) to Y.Z. A.B. acknowledges support from a Margaret Strauss Kramer Fellowship. Computational resources were provided by NYU-ITS. NMR data collected using the 600 MHz cryoprobe at NYU was supported by an NIH S10 grant (OD016343).

Author Contributions

A.B. designed research, prepared kinase samples, performed activity measurements, performed NMR experiments to assess ATP/Mg^{2+} binding affinities and CPMG relaxation

dispersion experiments, performed all computational analysis and simulations, and wrote the manuscript. W.M.M. performed NMR experiments to assess AMP-PCP/Mg²⁺ binding affinities and acquired preliminary datasets establishing the feasibility of the experimental NMR approach. M.M. designed research. Y.Z. directed and designed the project and wrote the manuscript. N.J.T. directed and designed the project and wrote the manuscript. All authors participated in data analysis and in revising the manuscript.

References

1. M. A. Lemmon, J. Schlessinger, Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-1134 (2010).
2. J. R. Pon, M. A. Marra, Driver and passenger mutations in cancer. *Annu Rev Pathol* **10**, 25-50 (2015).
3. M. H. Bailey *et al.*, Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **174**, 1034-1035 (2018).
4. F. Sanchez-Vega *et al.*, Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321-337 e310 (2018).
5. A. Torkamani, N. J. Schork, Prediction of cancer driver mutations in protein kinases. *Cancer Res* **68**, 1675-1682 (2008).
6. P. Cohen, D. Cross, P. A. Janne, Kinase drug discovery 20 years after imatinib: progress and future directions. *Nat Rev Drug Discov* **20**, 551-569 (2021).
7. F. M. Ferguson, N. S. Gray, Kinase inhibitors: the road ahead. *Nat Rev Drug Discov* **17**, 353-377 (2018).
8. I. Melnikova, J. Golden, Targeting protein kinases. *Nat Rev Drug Discov* **3**, 993-994 (2004).
9. B. J. Druker *et al.*, Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med* **2**, 561-566 (1996).
10. R. Roskoski, Jr., Properties of FDA-approved small molecule protein kinase inhibitors: A 2022 update. *Pharmacol Res* **175**, 106037 (2022).
11. M. E. Gorre *et al.*, Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* **293**, 876-880 (2001).
12. C. R. Antonescu *et al.*, Acquired resistance to imatinib in gastrointestinal stromal tumor occurs through secondary gene mutation. *Clin Cancer Res* **11**, 4182-4190 (2005).
13. L. Goyal *et al.*, Polyclonal Secondary FGFR2 Mutations Drive Acquired Resistance to FGFR Inhibition in Patients with FGFR2 Fusion-Positive Cholangiocarcinoma. *Cancer Discov* **7**, 252-263 (2017).
14. S. Kobayashi *et al.*, EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* **352**, 786-792 (2005).
15. M. R. Stratton, P. J. Campbell, P. A. Futreal, The cancer genome. *Nature* **458**, 719-724 (2009).

16. R. Hu, H. Xu, P. Jia, Z. Zhao, KinaseMD: kinase mutations and drug response database. *Nucleic Acids Res* **49**, D552-D561 (2021).
17. Y. Liu, K. Shah, F. Yang, L. Witucki, K. M. Shokat, A molecular gate which controls unnatural ATP analogue recognition by the tyrosine kinase v-Src. *Bioorganic & Medicinal Chemistry* **6**, 1219-1226 (1998).
18. K. S. Gajiwala *et al.*, KIT kinase mutants show unique mechanisms of drug resistance to imatinib and sunitinib in gastrointestinal stromal tumor patients. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1542-1547 (2009).
19. J. Cools *et al.*, A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *N Engl J Med* **348**, 1201-1214 (2003).
20. J. Mondal, P. Tiwary, B. J. Berne, How a Kinase Inhibitor Withstands Gatekeeper Residue Mutations. *Journal of the American Chemical Society* **138**, 4608-4615 (2016).
21. A. Hochhaus *et al.*, Molecular and chromosomal mechanisms of resistance to imatinib (STI571) therapy. *Leukemia* **16**, 2190-2196 (2002).
22. F. E. Nicolini *et al.*, The BCR-ABL T315I mutation compromises survival in chronic phase chronic myelogenous leukemia patients resistant to tyrosine kinase inhibitors, in a matched pair analysis. *Haematologica* **98**, 1510-1516 (2013).
23. Y. Lee *et al.*, Clinical outcome according to the level of preexisting epidermal growth factor receptor T790M mutation in patients with lung cancer harboring sensitive epidermal growth factor receptor mutations. *Cancer* **120**, 2090-2098 (2014).
24. C. D. Sohl, M. R. Ryan, B. Luo, K. M. Frey, K. S. Anderson, Illuminating the molecular mechanisms of tyrosine kinase inhibitor resistance for the FGFR1 gatekeeper mutation: the Achilles' heel of targeted therapy. *ACS Chem Biol* **10**, 1319-1329 (2015).
25. M. Azam, M. A. Seeliger, N. S. Gray, J. Kuriyan, G. Q. Daley, Activation of tyrosine kinases by mutation of the gatekeeper threonine. *Nat Struct Mol Biol* **15**, 1109-1118 (2008).
26. T. Xie, T. Saleh, P. Rossi, C. G. Kalodimos, Conformational states dynamically populated by a kinase determine its function. *Science* **370** (2020).
27. A. P. Kornev, N. M. Haste, S. S. Taylor, L. F. Eyck, Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 17783-17788 (2006).
28. V. Modi, R. L. Dunbrack, Jr., Defining a new nomenclature for the structures of active and inactive kinases. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 6818-6827 (2019).

29. H. Chen *et al.*, Elucidation of a four-site allosteric network in fibroblast growth factor receptor tyrosine kinases. *Elife* **6**, e21137 (2017).
30. J. G. Taylor VI *et al.*, Identification of FGFR4-activating mutations in human rhabdomyosarcomas that promote metastasis in xenotransplanted models. *The Journal of clinical investigation* **119**, 3395-3407 (2009).
31. S. A. Byron *et al.*, The N550K/H mutations in FGFR2 confer differential resistance to PD173074, dovitinib, and ponatinib ATP-competitive inhibitors. *Neoplasia* **15**, 975-988 (2013).
32. L. Chen *et al.*, Molecular basis for receptor tyrosine kinase A-loop tyrosine transphosphorylation. *Nat Chem Biol* **16**, 267-277 (2020).
33. A. K. Mittermaier, L. E. Kay, Observing biological dynamics at atomic resolution using NMR. *Trends in Biochemical Sciences* **34**, 601-611 (2009).
34. J. P. Loria, R. B. Berlow, E. D. Watt, Characterization of enzyme motions by solution NMR relaxation dispersion. *Acc Chem Res* **41**, 214-221 (2008).
35. S. S. Taylor, A. P. Kornev, Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci* **36**, 65-77 (2011).
36. J. Hu *et al.*, Kinase regulation by hydrophobic spine assembly in cancer. *Mol Cell Biol* **35**, 264-276 (2015).
37. W. M. Marsiglia *et al.*, A Conserved Allosteric Pathway in Tyrosine Kinase Regulation. *Structure* **27**, 1308-1315 e1303 (2019).
38. L. R. Masterson, A. Mascioni, N. J. Traaseth, S. S. Taylor, G. Veglia, Allosteric cooperativity in protein kinase A. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 506-511 (2008).
39. A. P. Kornev, S. S. Taylor, L. F. Ten Eyck, A helix scaffold for the assembly of active protein kinases. *Proceedings of the National Academy of Sciences* **105**, 14377-14382 (2008).
40. L. R. Masterson *et al.*, Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy. *Proceedings of the National Academy of Sciences* **108**, 6969-6974 (2011).
41. A. P. Kornev, S. S. Taylor, Defining the conserved internal architecture of a protein kinase. *Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics* **1804**, 440-444 (2010).
42. F. M. Gribble *et al.*, A novel method for measurement of submembrane ATP concentration. *J Biol Chem* **275**, 30046-30049 (2000).

43. K. Yoza *et al.*, Biophysical characterization of drug-resistant mutants of fibroblast growth factor receptor 1. *Genes to Cells* **21**, 1049-1058 (2016).
44. G. Sun, R. J. Budde, Requirement for an additional divalent metal cation to activate protein tyrosine kinases. *Biochemistry* **36**, 2139-2146 (1997).
45. H. Chen *et al.*, A molecular brake in the kinase hinge region regulates the activity of receptor tyrosine kinases. *Mol Cell* **27**, 717-730 (2007).
46. J. Lever, M. Krzywinski, N. Altman (2017) Points of significance: Principal component analysis. (Nature Publishing Group).
47. H. B. Chen *et al.*, Cracking the Molecular Origin of Intrinsic Tyrosine Kinase Activity through Analysis of Pathogenic Gain-of-Function Mutations. *Cell Reports* **4**, 376-384 (2013).
48. E. Lesca, A. Lammens, R. Huber, M. Augustin, Structural analysis of the human fibroblast growth factor receptor 4 kinase. *J Mol Biol* **426**, 3744-3756 (2014).
49. L. Tan *et al.*, Development of covalent inhibitors that can overcome resistance to first-generation FGFR kinase inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E4869-4877 (2014).
50. C. Chen *et al.*, Identification of a major determinant for serine-threonine kinase phosphoacceptor specificity. *Mol Cell* **53**, 140-147 (2014).
51. C. H. Yun *et al.*, Structures of lung cancer-derived EGFR mutants and inhibitor complexes: Mechanism of activation and insights into differential inhibitor sensitivity. *Cancer Cell* **11**, 217-227 (2007).
52. Y. Shan *et al.*, Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell* **149**, 860-870 (2012).
53. L. Sutto, F. L. Gervasio, Effects of oncogenic mutations on the conformational free-energy landscape of EGFR kinase. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 10616-10621 (2013).
54. D. M. Korzhnev, K. Klover, V. Kanelis, V. Tugarinov, L. E. Kay, Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *J Am Chem Soc* **126**, 3964-3973 (2004).
55. D. G. Davis, M. E. Perlman, R. E. London, Direct Measurements of the Dissociation-Rate Constant for Inhibitor-Enzyme Complexes Via the T-1-Rho and T-2 (Cpmg) Methods. *J Magn Reson Ser B* **104**, 266-275 (1994).
56. J. P. Carver, R. E. Richards, General 2-Site Solution for Chemical Exchange Produced Dependence of T2 Upon Carr-Purcell Pulse Separation. *Journal of Magnetic Resonance* **6**, 89-& (1972).

57. J. L. Battiste, G. Wagner, Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355-5365 (2000).
58. H. Chen *et al.*, Elucidation of a four-site allosteric network in fibroblast growth factor receptor tyrosine kinases. *elife* **6**, e21137 (2017).
59. M. Ester, H.-P. Kriegel, J. Sander, X. Xu (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. in *kdd*, pp 226-231.
60. F. Pedregosa *et al.*, Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825-2830 (2011).
61. J. H. Bae *et al.*, Asymmetric receptor contact is required for tyrosine autophosphorylation of fibroblast growth factor receptor in living cells. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 2866-2871 (2010).
62. S. Raman *et al.*, Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins: Structure, Function, and Bioinformatics* **77**, 89-99 (2009).
63. Y. Song *et al.*, High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735-1742 (2013).
64. K. L. Meagher, L. T. Redman, H. A. Carlson, Development of polyphosphate parameters for use with the AMBER force field. *Journal of Computational Chemistry* **24**, 1016-1025 (2003).
65. O. Allner, L. Nilsson, A. Villa, Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. *J Chem Theory Comput* **8**, 1493-1502 (2012).
66. A. W. Sousa da Silva, W. F. Vranken, ACPYPE-Antechamber python parser interface. *BMC research notes* **5**, 1-8 (2012).
67. J. A. Maier *et al.*, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713 (2015).
68. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463-1472 (1997).
69. U. Essmann *et al.*, A Smooth Particle Mesh Ewald Method. *Journal of Chemical Physics* **103**, 8577-8593 (1995).
70. D. J. Evans, B. L. Holian, The Nose-Hoover Thermostat. *Journal of Chemical Physics* **83**, 4069-4074 (1985).
71. M. Parrinello, A. Rahman, Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. *Journal of Applied Physics* **52**, 7182-7190 (1981).

72. M. V. Shapovalov, R. L. Dunbrack, Jr., A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **19**, 844-858 (2011).
73. A. Sali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815 (1993).

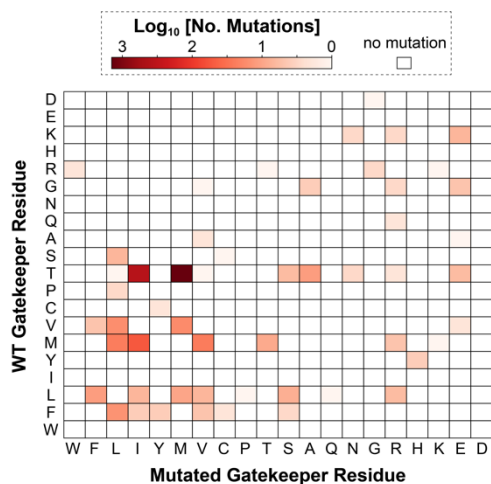


Figure 1. Distribution of gatekeeper mutations within tyrosine kinases from the kinase mutations and drug response database (16).

Heatmap of gatekeeper mutations, where the native gatekeeper residue is displayed on the y-axis and the residue it is mutated to is shown on the x-axis. The greatest occurrence of mutations is colored in red and stems from the same mutation observed in different tyrosine kinases and in multiple tumor samples; white boxes correspond to no observed mutations. Amino acid residues are ordered from hydrophobic to hydrophilic in the left to right and bottom to top directions.

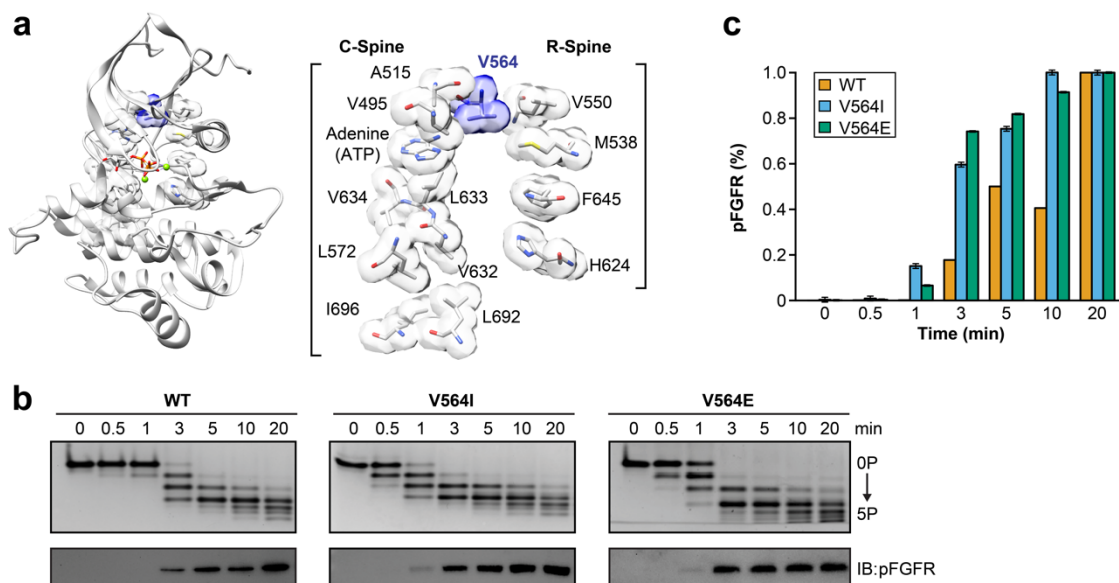


Figure 2. Gatekeeper mutants display increased phosphorylation activity *in vitro*.

a. Cartoon representation of FGFR2K structure (PDB ID: 2PVF) highlighting the catalytic and regulatory spines (C- and R-spine). Left: Whole view of the kinase structure with the non-hydrolyzable nucleotide AMP-PCP (grey), gatekeeper residue (blue), and hydrophobic R- and C-spine residues (grey) represented in sticks. The surface of gatekeeper and hydrophobic spine residues are also displayed in blue and grey, respectively. Mg^{2+} atoms are shown in green spheres. Right: expanded view of the gatekeeper and hydrophobic spine residues represented as in the left panel.

b. Phosphorylation assays of wild-type FGFR2K and its gatekeeper mutants evaluated using native gel electrophoresis (top panels) and immunoblot (IB) analyses with an antibody specific for A-loop phosphorylated FGFR at both Y656 and Y657 (bottom panels).

c. Quantification of phosphorylation at A-loop tyrosines (Y656/Y657) from the IB data in panel (b) for FGFR2K and gatekeeper mutants. Intensities for each respective kinase sample are normalized to a value of 1 for the 20 min time point.

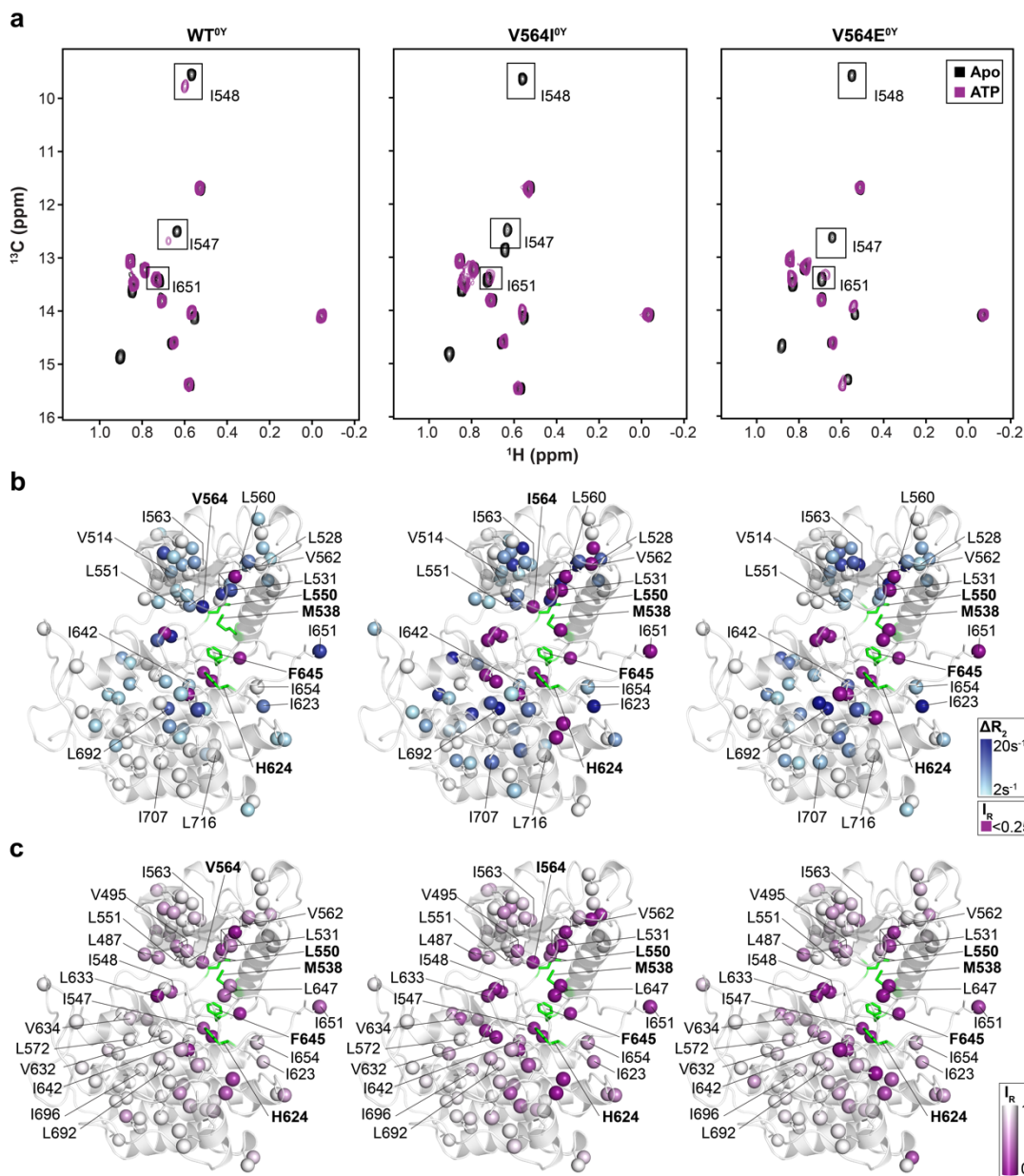


Figure 3. ATP binding increases conformational heterogeneity of gatekeeper mutants.

a. ¹H/¹³C HMQC spectra of WT^{0Y}, V564I^{0Y}, and V564E^{0Y} in the absence (apo state; black) and presence of ATP (20 mM) and Mg²⁺ (40 mM; purple). Residues highlighted with a box indicate example residues with notable differences between WT^{0Y} and the gatekeeper mutants.

b. Results from CPMG relaxation dispersion experiments and intensity retention calculations plotted on the homology model of autoinhibited FGFR2K. The difference between CPMG

relaxation dispersion experiments performed at two frequencies, ΔR_2 , was calculated as in equation 1 and plotted on the homology model of autoinhibited FGFR2K. Significant ΔR_2 value are shown from 20 s⁻¹ (blue) to 2 s⁻¹ (light blue). ΔR_2 value less than 2 s⁻¹ are shown in white. Peaks with intensity retention (I_R) values less than 0.25 from panel (a) are displayed in purple. Hydrophobic R-spine residues are displayed in green; residue labels in bold correspond to the gatekeeper residue and hydrophobic R-spine residues. Note that some residues displayed missing peaks in the CPMG experiment for WT^{0Y} (L647, L675) and V564E^{0Y} (L675) and could not be analyzed; these residues are not displayed in the respective plots.

c. I_R values for HMQC spectra from panel (a) plotted on the homology model of autoinhibited FGFR2K. I_R values were calculated as the intensities of ATP and Mg²⁺ bound states divided by the intensities of the apo states. I_R values plotted range from 0 (purple) to 1 (white). Hydrophobic R-spine residues are displayed in green; residue labels in bold correspond to the gatekeeper residue and hydrophobic R-spine residues.

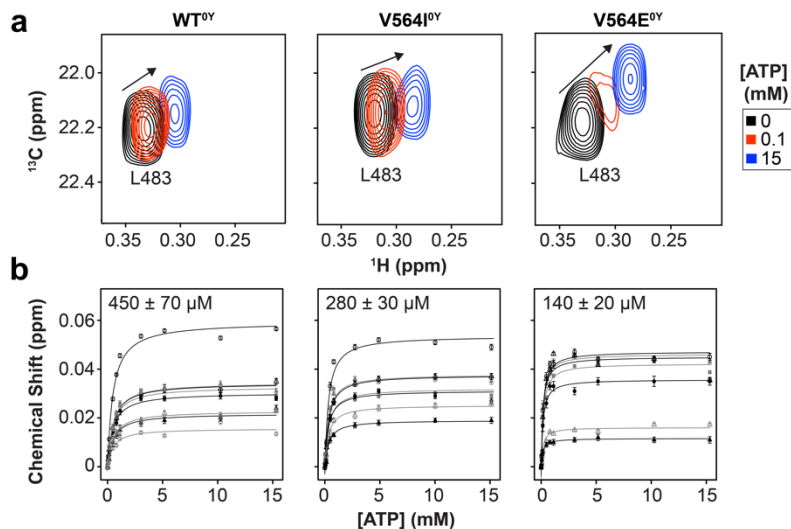


Figure 4. Gatekeeper mutations of FGFR2K enhance the ATP binding affinity.

a. $^1\text{H}/^{13}\text{C}$ methyl HMQC spectra of L483 ($\beta 1$ strand) at three concentrations of ATP using WT^{0Y} (left), V564I^{0Y} (middle), and V564E^{0Y} (right). Full titrations for all peaks are displayed in the supporting information. Note that the L617 peak (22.2 ^{13}C ppm and 0.25 ^1H ppm) was removed from the spectra for clarity.

b. Chemical shift perturbation as a function of ATP concentrations (absolute value shown). ^{13}C chemical shifts were scaled by a factor of 0.25 to account for the ^1H chemical shift range. Binding affinity was quantified using a global non-linear least-square fit to assigned residues experiencing chemical shift perturbations in fast exchange. Individual residues correspond to the symbols as represented below.

WT^{0Y}: $\circ$ I548^{C δ 1}, $\blacktriangle$ V516^{H γ} , $\triangle$ I548^{H δ 1}, $\bullet$ L483^{H δ} , $\bullet$ V516^{H γ} , $\triangle$ L572^{H δ} , $\blacktriangle$ L560^{H δ} , $\circ$ L483^{C δ}

V564I^{0Y}: $\circ$ L496^{H δ} , $\blacktriangle$ L483^{H δ} , $\triangle$ V516^{H γ} , $\bullet$ V516^{C γ} , $\bullet$ V516^{H γ} , $\triangle$ L560^{H δ} , $\blacktriangle$ L572^{H γ} ;

V564E^{0Y}: $\circ$ V516^{H γ} , $\blacktriangle$ L483^{C δ} , $\triangle$ I623^{C δ 1}, $\bullet$ L483^{C δ} , $\bullet$ L560^{H δ} , $\triangle$ L603^{C δ} , $\blacktriangle$ L603^{H δ} .

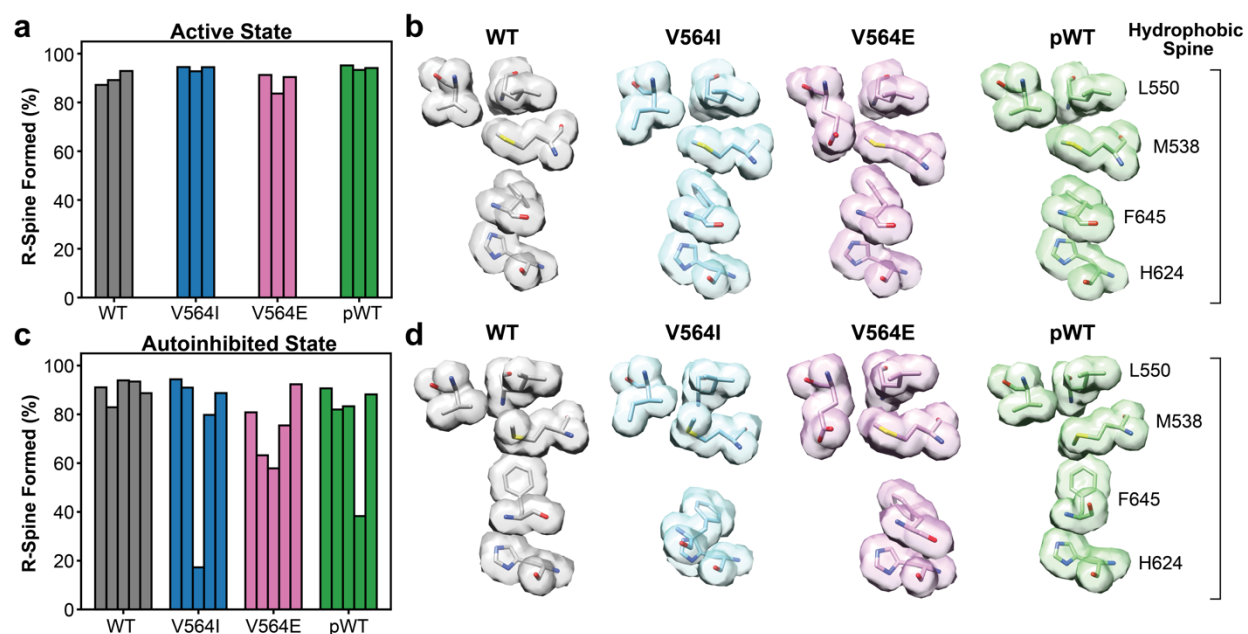


Figure 5. MD simulations reveal disruption of the hydrophobic R-spine within the autoinhibited state.

a, c. The percentage of R-spine formed calculated from MD simulations every 5 psec starting from the active (panel a) and autoinhibited states (panel c) for unphosphorylated wild-type FGFR2K (WT; grey), V564I (blue), V564I (pink), and phosphorylated wild-type FGFR2K (pWT; green). Each bar represents one replicate of 1 μ s sampling time.

b, d. Screenshots of representative structures for the gatekeeper (position 564) and hydrophobic R-spine residues (M538, L550, H624, F645) sampled during MD simulations starting from the active (panel b) and autoinhibited states (panel d). In each panel, the hydrophobic R-spine residues correspond to the order displayed in the right panel. Gatekeeper and R-spine residues are shown in sticks and a surface representation.

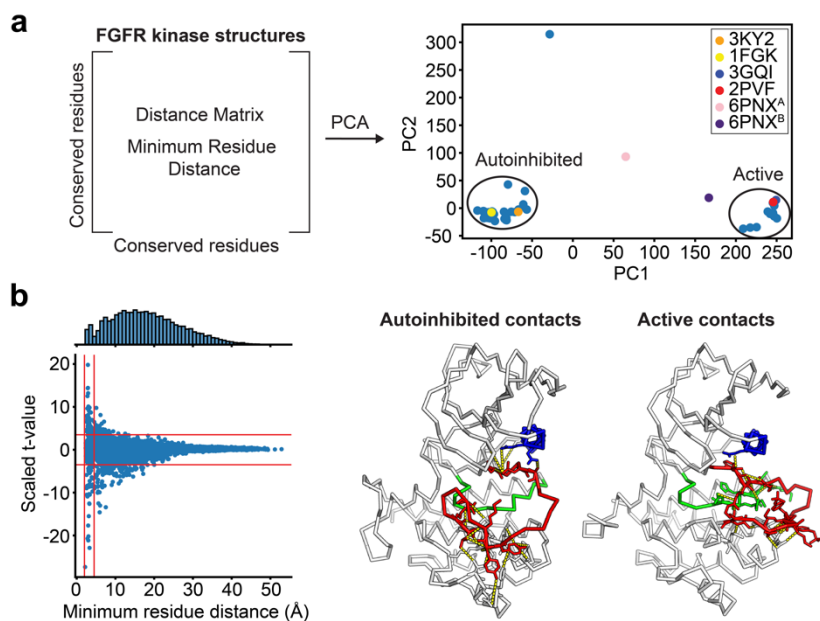


Figure 6. Classification of PDB structures and contacts distinguishing active and autoinhibited conformations.

a. Left: Approach to classify FGFR kinase structures using a PCA analysis of the minimum residue distance matrix for conserved residues among the four FGFR isoforms. Right: Crystal structures are plotted on the first two principal components showing two distinct clusters highlighted as autoinhibited and active states. Representative structures are highlighted from each cluster where 3KY2 and 1FGK represent unphosphorylated wild-type FGFR kinases and 3GQI and 2PVF represent phosphorylated wild-type FGFR kinases. 6PNX^A and 6PNX^B correspond to the substrate-acting and enzyme-acting kinases, respectively, in the A-loop transphosphorylation asymmetric dimer structure of FGFR3 kinase (32). A complete list of PDBs and where they cluster is displayed in SI Appendix, Table S1.

b. Left: Scaled t-value between active and inactive clusters of each distance-distance pair is plotted against the minimum residue distance. Red lines represent the cutoff region used to extract

significant autoinhibited and active contacts. Right: Autoinhibited and active contacts are plotted on their representative structure and contacts are highlighted in yellow dotted lines. Key regions in the kinase are colored as follows: α C helix (blue), A-loop (red), and catalytic loop (green).

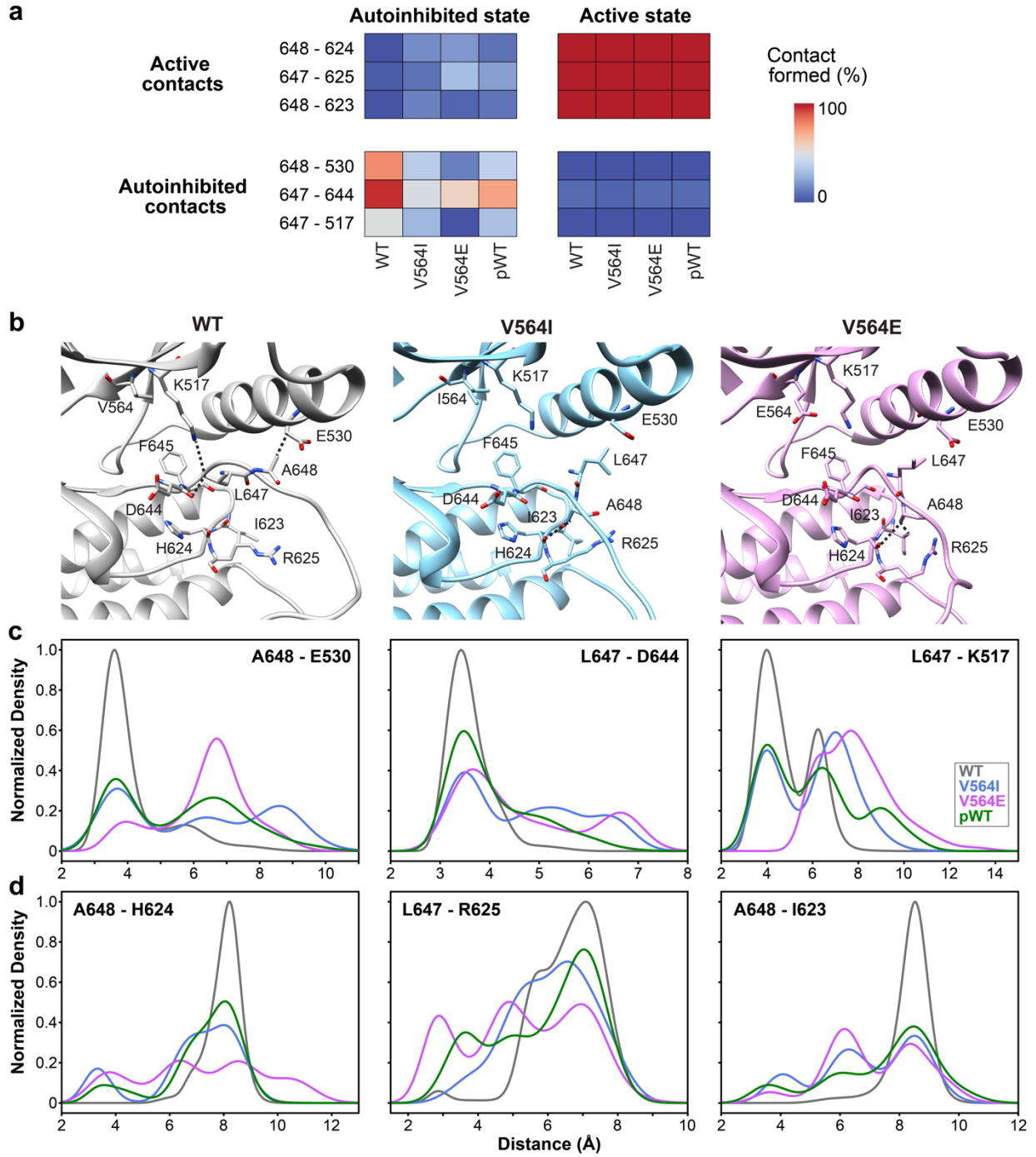


Figure 7. Gatekeeper mutants destabilize the autoinhibited state around the DFG motif and DFG+1 and DFG+2 residues.

a. Heatmap of active or autoinhibited contacts that are formed (red) and broken (blue) during MD simulations. The two heatmap columns refer to simulations starting from the autoinhibited state

(left) or active state (right). Residue pairs displayed for autoinhibited and active contacts represent the three greatest differences between wild-type and gatekeeper mutants from simulations initiated from the autoinhibited state structure (PDB ID: 3KY2).

b. Representative snapshots of MD simulations of wild-type FGFR2K (grey), V564I (light blue), and V564E (pink) that display contacts formed and disrupted starting from the autoinhibited state. Select residues are displayed that correspond to those identified in panel (a) and those of the DFG motif and gatekeeper position. Black dotted lines represent distances 4.5 Å or less for autoinhibited contacts for the wild-type FGFR2 snapshot and active contacts for the gatekeeper mutation snapshots.

c, d. Density plots corresponding to the minimum residue distance of the indicated autoinhibited (c) or active contact (d) for MD stimulations starting from the autoinhibited conformation. Distances were calculated for each 5 psec of the MD stimulation. Colors of the density plots are displayed within the right panel of (c).

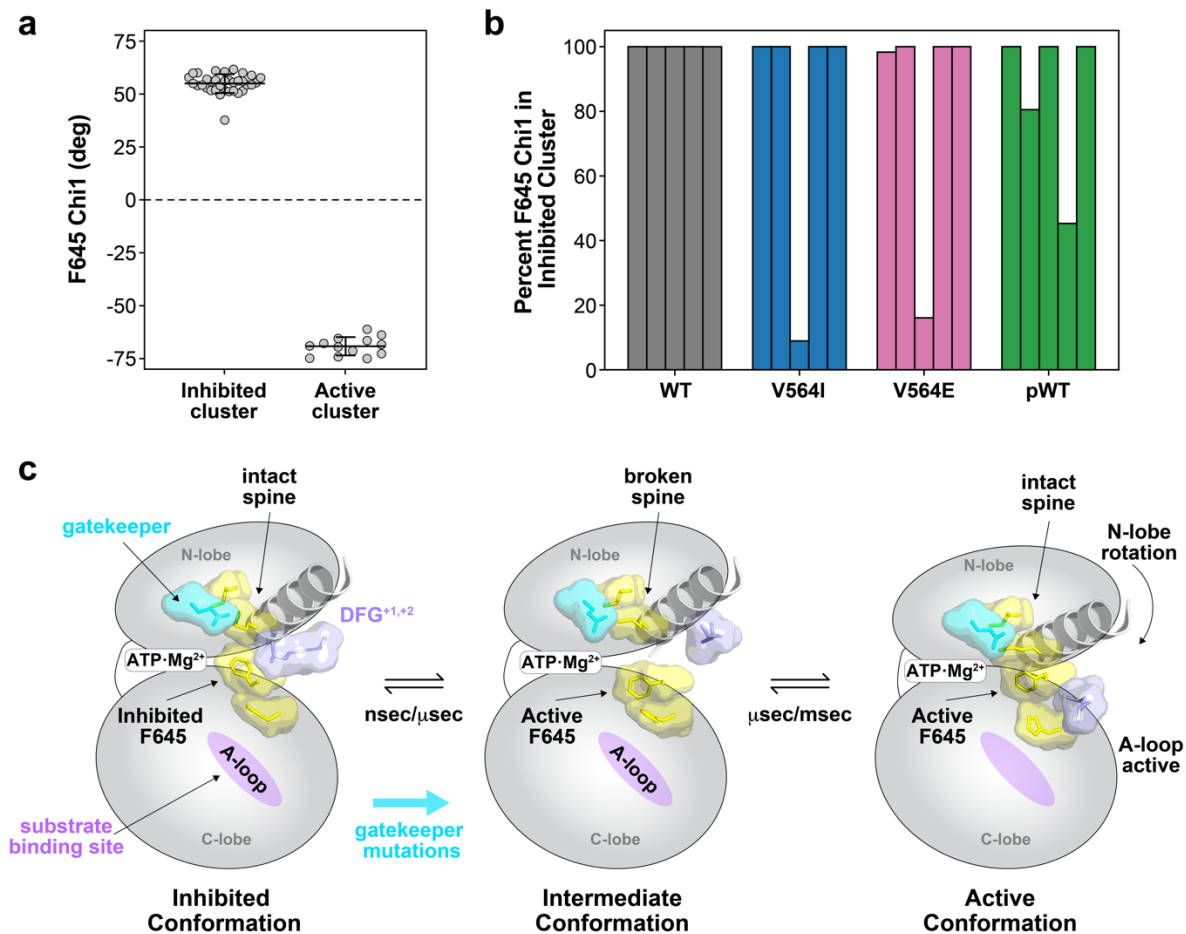


Figure 8. Mechanism of gatekeeper mutation activation occurring through the hydrophobic R-spine and DFG motif.

- a.** F645 chi1 dihedral angle calculated for the autoinhibited (DFG^{in, inactive}) and active (DFG^{in, active}) clusters displayed in Fig. 6a. Two PDBs (4UXQ, 4QRC) within the autoinhibited cluster were excluded from the plot since they were bound by type II inhibitors and in the DFG^{out} conformation.
- b.** The percentage of F645 in the autoinhibited conformation (i.e., chi1 angle > 0°) from MD stimulations starting from the autoinhibited state for unphosphorylated wild-type FGFR2K (WT; grey), V564I (blue), V564E (pink), and phosphorylated wild-type FGFR2K (pWT; green). The

dihedral angle was calculated every 5 psec of the MD simulation. Each bar in the graph corresponds to one replicate of 1 μ s sampling time.

c. Model depicting mutations at the gatekeeper residue (in cyan) disrupting the hydrophobic R-spine (in yellow) of the autoinhibited conformation and influencing the conformation of the DFG phenylalanine (F645). The α C-helix is depicted in grey, the substrate binding site is depicted as a purple oval, and the N- and C-lobes of the kinase are shown as grey circles. The A-loop text in the substrate binding site forms an autoinhibited interaction, while the A-loop text outside depicts the active conformation of the loop.

Supporting Information for

Gatekeeper Mutations Activate FGF Receptor Tyrosine Kinases by Destabilizing the Autoinhibited State

Alida Besch, William M. Marsiglia, Moosa Mohammadi, Yingkai Zhang*, Nathaniel J. Traaseth*

*Corresponding authors: yingkai.zhang@nyu.edu, traaseth@nyu.edu

This PDF file includes:

Figures S1 to S12
Tables S1 to S2

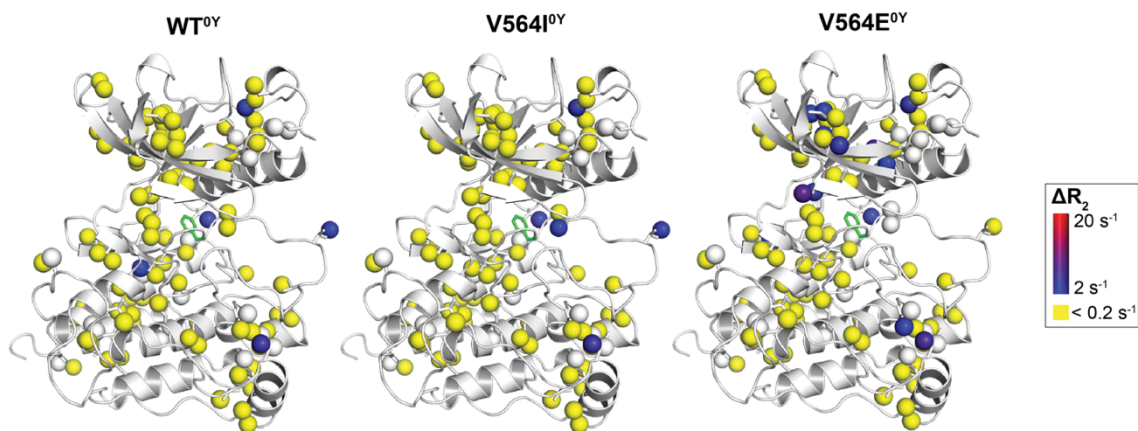


Fig. S1. CPMG relaxation dispersion experiments for apo FGFR2K and gatekeeper mutations display no significant relaxation dispersions.

WT^{0Y}, V564I^{0Y}, and V564E^{0Y} were subject to CPMG experiments and the ΔR_2 value is plotted on the structure of the homology model of FGFR2K. ΔR_2 was calculated using equation 1 from two CPMG relaxation dispersion experiments correspond to pulsing rates of 50 Hz and 1,000 Hz.

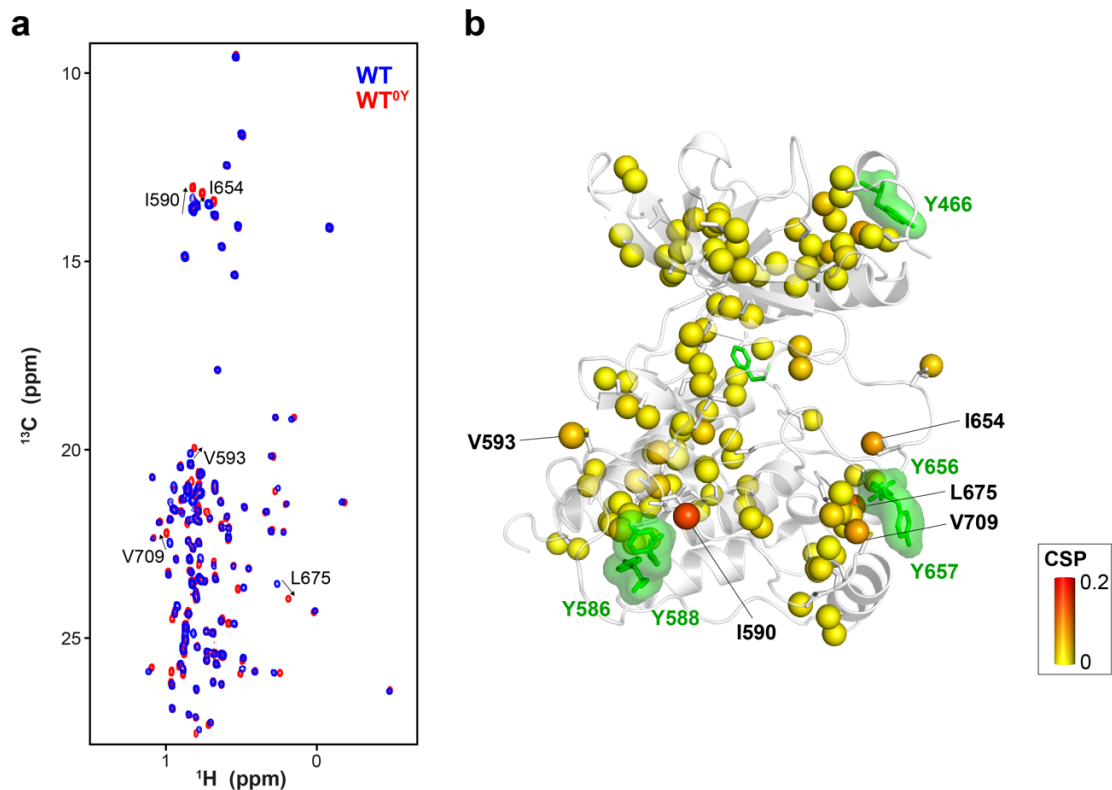


Fig. S2. WT^{0Y} shows minimal structural perturbations compared to wild-type FGFR2K.

a. $^1\text{H}/^{13}\text{C}$ HMQC spectra of ILV methyl groups comparing wild-type FGFR2K (blue) to WT^{0Y} (red). Residues experiencing a chemical shift perturbation (CSP) > 0.04 ppm are shown by arrows.

b. CSPs plotted on the FGFR2K autoinhibited homology model from 0 (yellow) to 0.2 (red). Residues experiencing CSPs > 0.04 ppm are labeled on the structure. Tyrosines that were mutated in the “0Y” construct are shown in green. CSPs occur only nearby the mutated residues indicating a change in local electrostatics and no major changes to the protein’s conformation.

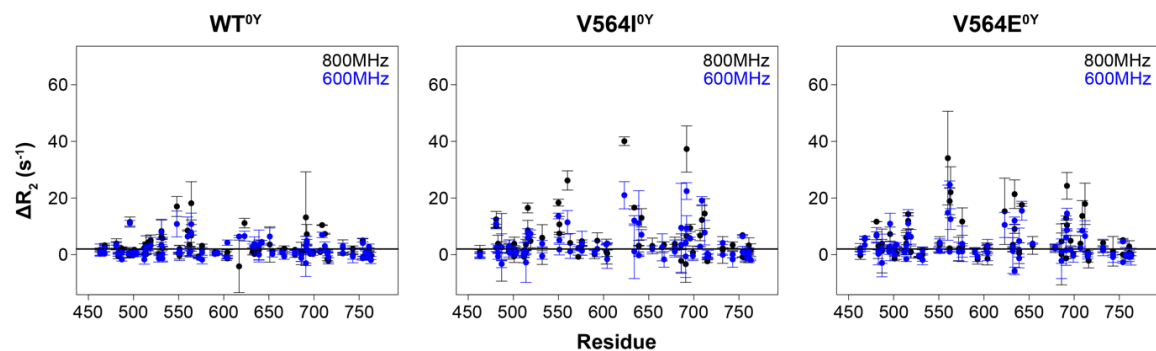


Fig. S3. CPMG relaxation dispersion values are larger at a higher magnetic field strength suggesting exchange is in the fast or fast/intermediate regime.

ΔR_2 values from CPMG relaxation dispersion experiments plotted as a function of residue for WT^{0Y} (left), V564I^{0Y} (middle), and V564E^{0Y} (right). Experiments were conducted at two magnetic field strengths corresponding to 600 MHz (blue) and 800 MHz (black). A black line is plotted at 2 s⁻¹, indicating significant ΔR_2 values. Peaks with intensity retention values less than 0.25 from HMQC experiments are not included in the plots as these peaks are broadened beyond detection in CPMG experiments.

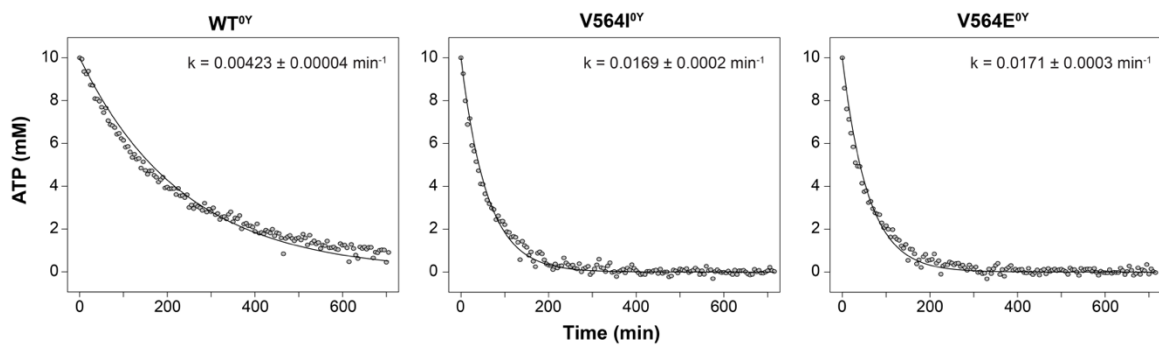


Fig. S4. Gatekeeper mutants result in a higher amount of ATPase activity assay.

WT^{0Y}, V564I^{0Y}, and V564E^{0Y} were incubated with ATP and Mg²⁺ and ³¹P 1D NMR experiments were collecting in series to quantify the amount of ATP hydrolysis. Data were fit to a single exponential to yield the rate constants displayed in each panel.

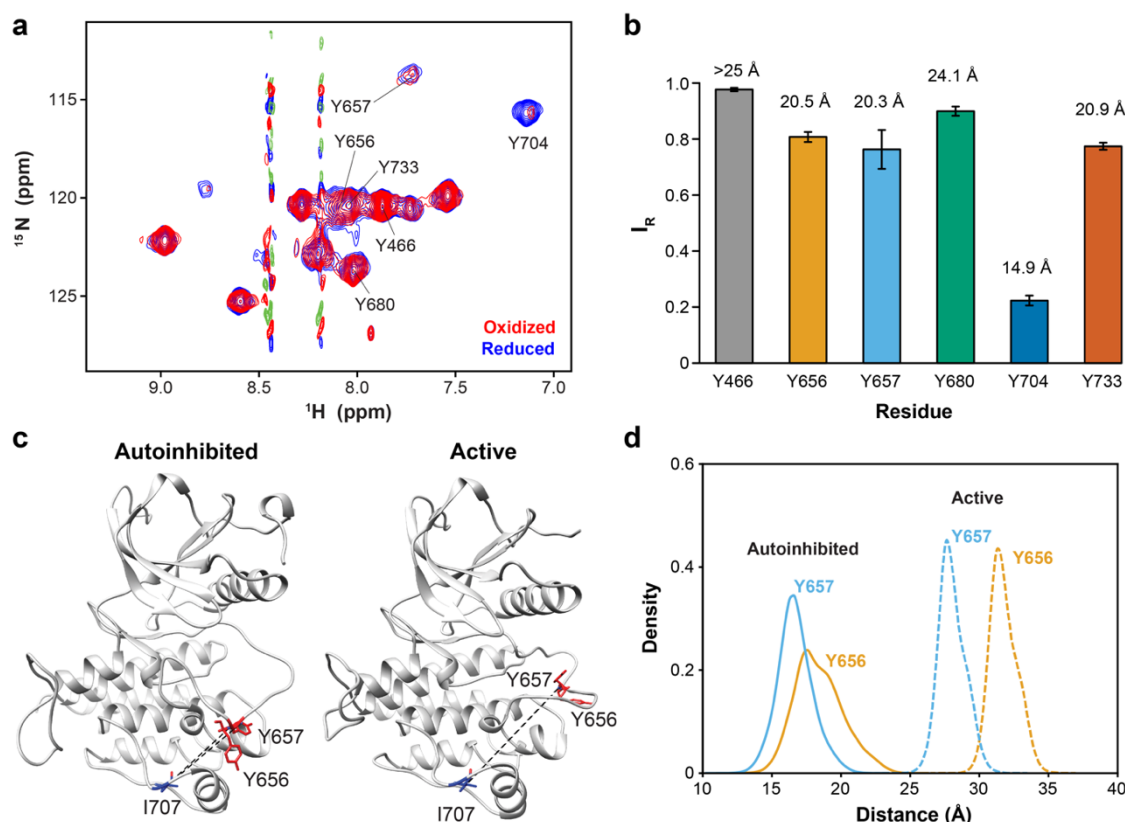


Fig. S5. PRE experiments indicate FGFR2K bound to nucleotide is primarily in the autoinhibited A-loop conformation.

a. $^1\text{H}/^{15}\text{N}$ HSQC spectra of ^{15}N -Tyr I707C FGFR2K labeled with MTSL in the oxidized (red) and reduced (blue) forms. Each sample contains 10 mM AMP-PCP and 20 mM MgCl_2 . The vertical streaks at ~8.2 and ~8.4 ppm corresponds to natural abundance nucleotide.

b. Intensity retention (I_R) plot of assigned residues calculated as the ratio of peak intensities for the oxidized and reduced samples. Calculated distances from the I_R values are displayed above the bars.

c. Autoinhibited structure (homology model of PDB ID: 3KY2) and active structure (PDB ID: 2PVF) highlighting the key distances that change regarding autoinhibited or active states of Y656 and Y657. I707 is represented in blue and is the location for the MTSL labeled site. Y656 and Y657, labeled in red, change their distances from residue 707 depending on whether the kinase is in the autoinhibited or active conformation.

d. Density plot of distances from wild-type FGFR2K autoinhibited state (solid lines) and active state (dotted lines) MD simulations. Distances were calculated from the C γ of I707 to the amide hydrogen on the backbone every 5 psec from the MD trajectories. The experimentally determined distance of ~20 Å indicates better agreement with the autoinhibited conformation. Note that the average tyrosine distances from MD simulations were calculated in a similar manner and were in good agreement with the PRE calculated distances: Y466 = 48.7 Å, Y680 = 25.3 Å, Y704 = 7.1 Å, and Y733 = 20.4 Å. MD simulations indicated a shorter distance for Y704 than the experimentally calculated value, which could arise from a fraction of incomplete MTSL labeling. However, this does not change the conclusion that the autoinhibited conformation is the predominant population.

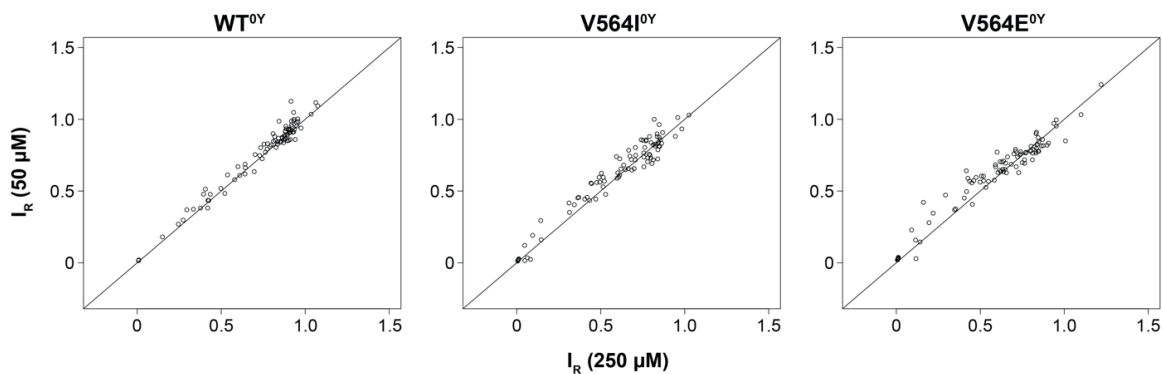


Fig. S6. Intensity retention values upon ATP binding are independent of the kinase concentration.

Correlation plots of intensity retention values (I_R) upon addition of 20 mM ATP and 40 mM Mg^{2+} for WT^{0Y} (left), V564I^{0Y} (middle), and V564E^{0Y} (right) at concentrations of 50 μM and 250 μM . Intensity retention values were calculated as the peak intensities in the ATP/ Mg^{2+} bound state divided by the intensities in absence of ATP. Solid black lines represent slopes equal to 1.

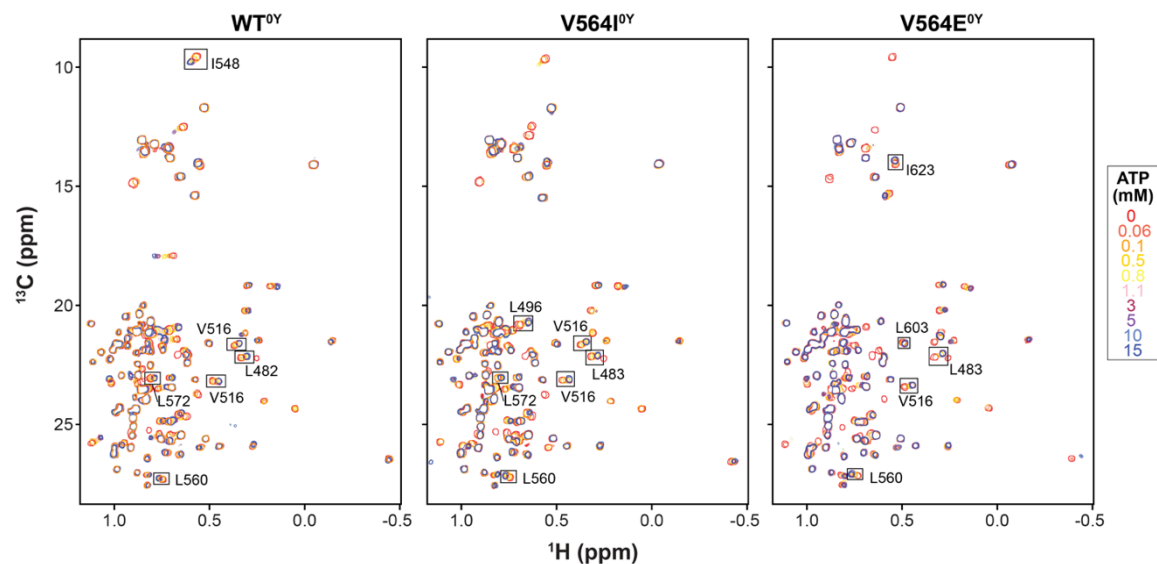


Fig. S7. $^1\text{H}/^{13}\text{C}$ HMQC spectra for ATP titrations.

ATP/ Mg^{2+} titration data for non-phosphorylatable kinases (0Y) corresponding to wild-type (left), V564I (middle), and V564E (right) kinases. The ATP concentrations used for the titration are shown on the right of V564E spectra. Note that V564I^{0Y} does not have an HMQC spectrum corresponding to 1.1 mM ATP concentration; this titration also has 2.7 mM ATP in place of the 3 mM ATP point in magenta.

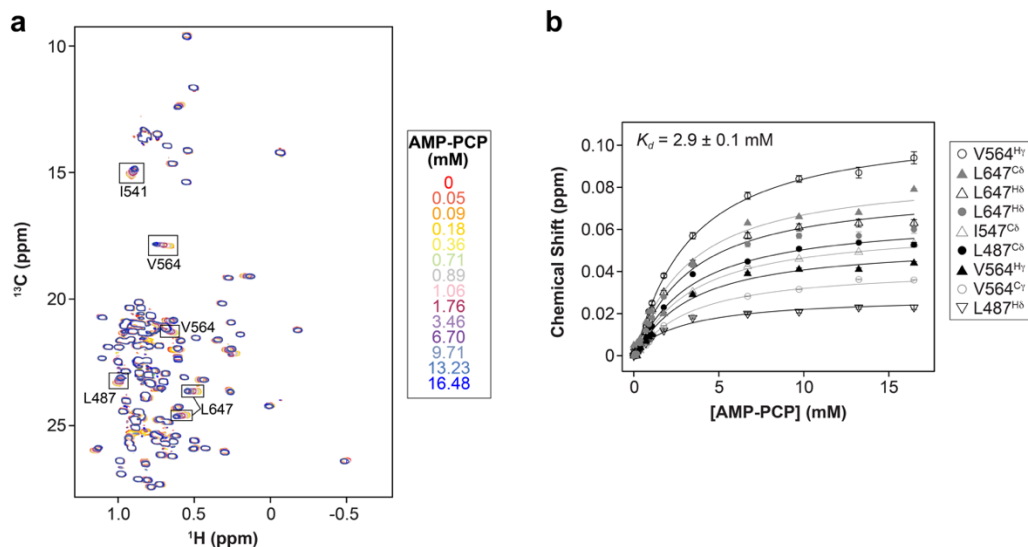


Fig. S8. $^1\text{H}/^{13}\text{C}$ HMQC spectra of AMP-PCP titration and determination of binding affinity.

a. $^1\text{H}/^{13}\text{C}$ methyl HMQC spectra of wild-type FGFR2K titrated with AMP-PCP. Residues highlighted display the largest chemical shift perturbations that were subsequently fit to obtain the binding affinity.

b. Chemical shift perturbation as a function of AMP-PCP concentrations (absolute value shown). ^{13}C chemical shifts were scaled by a factor of 0.25 to account for the ^1H chemical shift range. Binding affinity was quantified using a global non-linear least-square fit to residues experiencing chemical shift perturbations highlighted in panel (a).

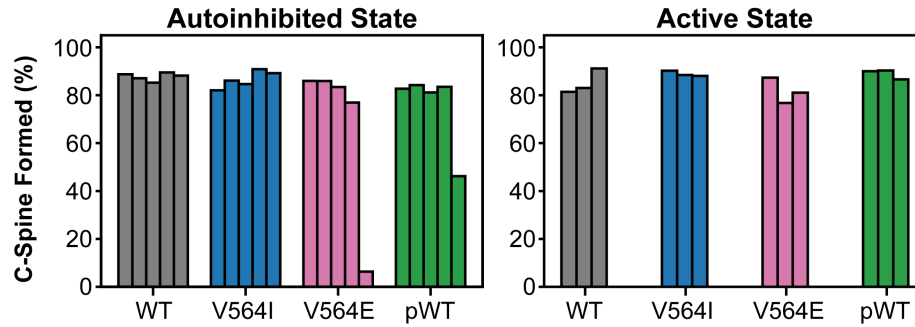


Fig. S9. MD simulations show a stable hydrophobic C-spine for the active state and a slight destabilization of the hydrophobic C-spine for V564E and pWT of the autoinhibited state.

The percentage of C-spine formed calculated from MD simulations every 5 psec starting from the autoinhibited and active states for unphosphorylated wild-type FGFR2K (WT; grey), V564I (blue), V564I (pink), and phosphorylated wild-type FGFR2K (pWT; green). Each bar represents one replicate of a total sampling time of 1 μ s. The C-spine was defined as intact if the minimum side-chain contact distance of any residue associated with the C-spine and the heavy atoms of the adenine base of ATP was less than 4.5 Å from one another. Disruption of the C-spine for one of the replicates in the V564E simulation stemmed from the adenine base moving out of adenine pocket.

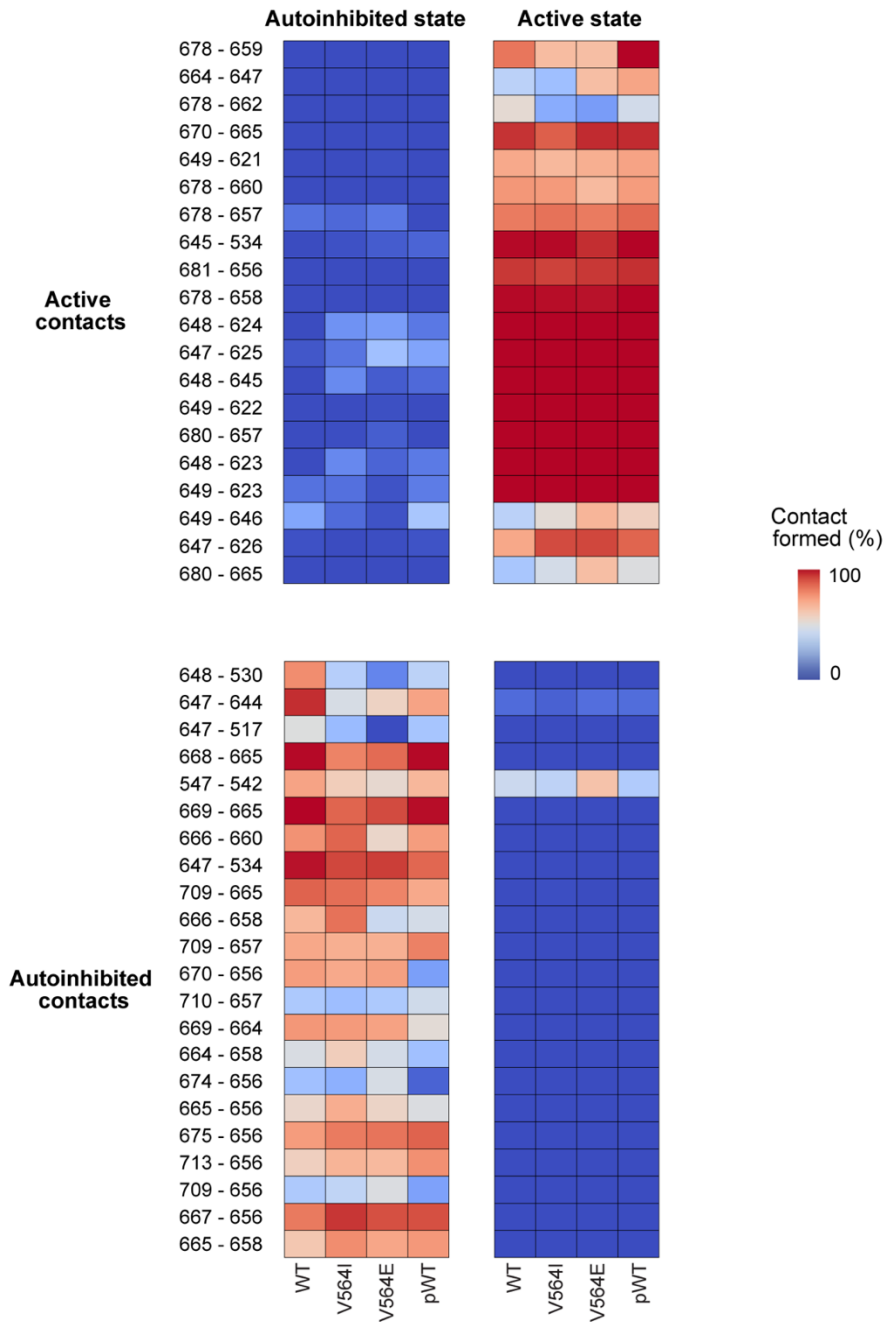


Fig. S10. MD simulation analysis using the active and autoinhibited contacts for unphosphorylated and phosphorylated wild-type FGFR2Ks and the gatekeeper mutants.

Heatmap of active contacts that are formed (red) and broken (blue) during MD simulations. Top: Displayed contacts are sorted by ranking the active contacts broken in gatekeeper mutants compared to the phosphorylated active state of FGFR2K. Bottom: Displayed contacts are sorted by ranking the autoinhibited contacts broken in gatekeeper mutants compared to the unphosphorylated autoinhibited state of FGFR2K.

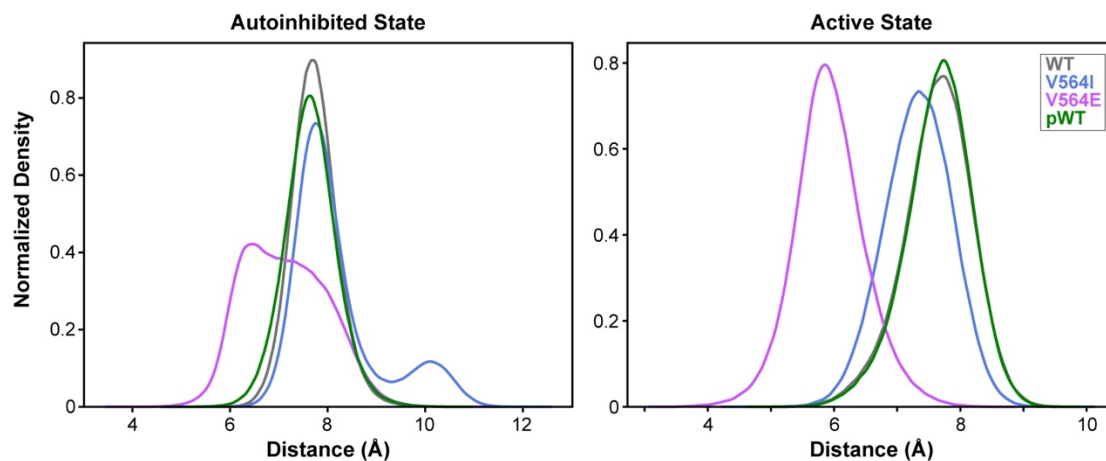


Fig. S11. MD simulations display no molecular contact between the gatekeeper residue and the phenylalanine from the DFG motif.

Density plots corresponding to the minimum residue distance of the gatekeeper residue and F645 from the DFG motif for MD simulations starting from the autoinhibited conformation (left) or active conformation (right). Distances were calculated for each 5 psec of the MD stimulation. Colors of the density plots are displayed within the right panel.

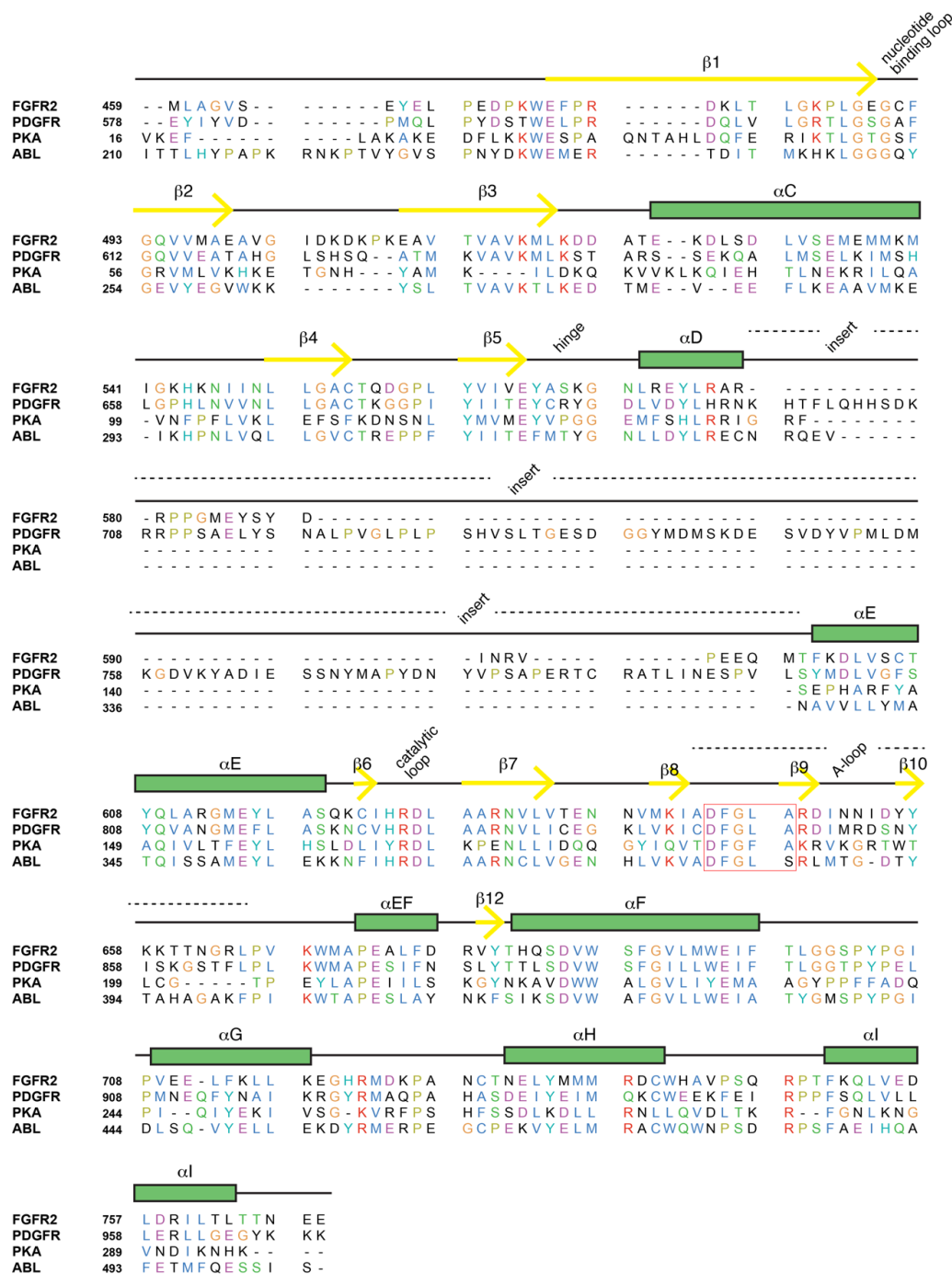


Fig. S12. Sequence alignment of FGFR2K with PDGFR, PKA and Abl kinases.

Multiple sequence alignment on the kinase domains of FGFR2 (P21802), platelet-derived growth factor receptor (PDGFR; P09619), protein kinase A (PKA; P17612), and Abl kinase (ABL; P00519) with the corresponding UniProt IDs displayed in the parentheses. Secondary motifs according to FGFR2K are displayed above the sequence; the DFG motif and DFG+1 and DFG+2 residues are indicated with a red box.

Table S1. Classification of FGFR kinase PDB structures.

Cluster	PDB	Isoform	Gain of Function Mutation	Phosphorylated	Drug Bound
Autoinhibited	6P69	FGFR1			x
	6P68	FGFR1			x
	1FGI	FGFR1			x
	1FGK	FGFR1			
	3TT0	FGFR1			x
	3KY2	FGFR1			
	6ITJ	FGFR1			x
	3C4F	FGFR1			x
	4UWC	FGFR1			x
	4F63	FGFR1			x
	3RHX	FGFR1			x
	4UWB	FGFR1			x
	5O49	FGFR1			x
	4F64	FGFR1			x
	4ZSA	FGFR1			x
	5ZV2	FGFR1			x
	5AM6	FGFR1			x
	5AM7	FGFR1	x		x
	5VND	FGFR1			x
	5EW8	FGFR1			x
	4V05	FGFR1			x
	1AGW	FGFR1			x
	4UWY	FGFR1			
	4RWL	FGFR1			x
	4RWI	FGFR1	x		
	5B7V	FGFR1			x
	4RWJ	FGFR1			x
	5Z0S	FGFR1			x
	4RWK	FGFR1	x		x
	4QQT	FGFR4	x		
	4UXQ	FGFR4			x
	4XCU	FGFR4			x
	4QRC	FGFR4	x		x
	6JPJ	FGFR4	x		x
	4QQJ	FGFR4	x		
	6IUP	FGFR4	x		x
	4TYG	FGFR4			
Active	3GQI	FGFR1			
	5FLF	FGFR1	x		
	3KXX	FGFR1			
	4J96	FGFR2	x		
	4J97	FGFR2	x		
	2PVF	FGFR2		x	
	5UGX	FGFR2	x		
	2PWL	FGFR2	x		
	5UI0	FGFR2	x		
	3RI1	FGFR2			x
	5EG3	FGFR2	x	x	
	2PSQ	FGFR2			
Other	4K33	FGFR3	x		
	6PNX	FGFR3	x		
	4TYG	FGFR4			

Table S2. Active and autoinhibited contacts with corresponding location within the kinase.

	Residue 1	Location 1	Residue 2	Location 2
Active Contacts	678	Loop (α EF- α F)	659	A-Loop
	664	A-Loop	647	DFG+1
	678	Loop (α EF- α F)	662	A-Loop
	670	α EF Helix	665	A-Loop
	649	A-Loop	621	Catalytic Loop
	678	Loop (α EF- α F)	660	A-Loop
	678	α EF Helix	657	A-Loop
	645	DFG Motif	534	α C Helix
	681	Loop (α EF- α F)	656	A-Loop
	678	α EF Helix	658	A-Loop
	648	DFG+2	624	HRD Motif
	647	DFG+1	625	HRD Motif
	648	DFG+2	645	DFG Motif
	649	A-Loop	622	Catalytic Loop
	680	Loop (α EF- α F)	657	A-Loop
	648	DFG+2	623	Catalytic Loop
	649	A-Loop	623	Catalytic Loop
	649	A-Loop	646	DFG Motif
	647	DFG+1	626	HRD Motif
	680	Loop (α EF- α F)	665	A-Loop
Autoinhibited Contacts	648	DFG+2	530	α C Helix
	647	DFG+1	644	DFG Motif
	647	DFG+1	517	β 3 strand
	668	α EF Helix	665	A-Loop
	547	DFG-Latch	542	DFG-Latch
	669	α EF Helix	665	A-Loop
	666	α EF Helix	660	A-Loop
	647	DFG+1	534	α C Helix
	709	α G Helix	665	A-Loop
	666	α EF Helix	658	A-Loop
	709	α G Helix	657	A-Loop
	670	α EF Helix	656	A-Loop
	710	α G Helix	657	A-Loop
	669	α EF Helix	664	A-Loop
	664	A-Loop	658	A-Loop
	674	α EF Helix	656	A-Loop
	665	A-Loop	656	A-Loop
	675	α EF Helix	656	A-Loop
	713	α G Helix	656	A-Loop
	709	α G Helix	656	A-Loop
	667	α EF Helix	656	A-Loop
	665	A-Loop	658	A-Loop