

Supplementary Materials for

Using ancient protein kinases to unravel a modern cancer drug's mechanism

C. Wilson, R. V. Agafonov, M. Hoemberger, S. Kutter, A. Zorba, J. Halpin,
V. Buosi, R. Otten, D. Waterman, D. L. Theobald, D. Kern*

*Corresponding author. E-mail: dkern@brandeis.edu

Published 20 February 2015, *Science* **347**, 882 (2015)
DOI: 10.1126/science.aaa1823

This PDF file includes:

Material and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 and S2
References

Materials and Methods:

Ancestral protein sequence reconstruction

Seventy-six sequences were selected from the NCBI non-redundant protein sequence database spanning the Tec, Src, and Abl kinase subfamilies. Both phylogeny and alignment were co-estimated using the Bayesian BAli-Phy software package (Fig. S1) (34). The analysis was performed using the RS07 insertion/deletion model, LG amino acid substitution matrix, estimating equilibrium amino acid frequencies, with gamma distributed rates across sites (four categories). Two independent chains were run until the ASDSF and PSRF-80%CI criteria fell below 0.01 and 1.01, respectively. Ancestral sequences were inferred using the marginal likelihood method implemented in PAML (35), with the maximum *a posteriori* phylogeny and expected parameters (normalized equilibrium frequencies and gamma shape parameter) from the BAli-Phy run.

We note that although ancestral reconstruction is a well-established method (14, 17) it is still a developing field, and the underlying assumptions should be considered (18). The reconstructed proteins are probabilistic inferences. The estimated probability of reconstructing the exact actual ancestral sequence is the product of the probabilities for each site in the protein, and hence the overall probability is vanishingly small. However, the histograms of the posterior probabilities associated with each inferred position in the ancestral proteins (Fig. S1) show that the estimated confidence is high (PP > 95%) for the great majority of ancestral residues. In fast evolving regions of the protein the majority of the ambiguous residues are expected to be selectively neutral or nearly neutral, and the sequence alternatives involve chemically conservative substitutions. These mathematical considerations also reflect the fact that, like modern proteins, the ancestral proteins existed in large populations of organisms (in our case single-celled eukaryotes), comprising a polymorphic ensemble of similar proteins that changed over time. From a practical perspective, reconstructed sequences can be viewed as representatives of groups of proteins that are likely similar to ancestral sequences in biophysically relevant ways.

Expression and Purification

Ancestral sequence cDNAs were constructed by Genscript. Ancestral and extant inserts were sub-cloned into pET-41M vector containing a His-tag and MBP-tag on the N-terminus. Vector was co-transformed with the YOPH phosphatase (36) to ensure de-phosphorylated protein and to lower toxicity of the insert into GroEL competent BL-21 cells Arctic Express from Invitrogen (GroEL under tetracycline induction). Cells were grown in TB media to an OD of 0.8 at 37 °C then switched to 18 °C for 1 hour before induction with 100 µM of IPTG. Cells were allowed to grow for 16 hours at 18 °C. Cells were lysed in the presence of benzonase by sonication. After purification via a Talon- and MBP-column, the tags were cleaved with His-tagged TEV-protease overnight at 4 °C while dialyzing against storage buffer (25 mM Tris-HCl pH 8, 500 mM NaCl, 5% (v/v) glycerol). Cleaved sample was collected and run over Ni-NTA column to remove His-tagged TEV, cleaved MBP and uncleaved His-MBP-Kinase contaminants. Flow-through was collected, concentrated to 5 mL and passed over a 16/60 S-100 gel filtration column. All columns were run at 4 °C. Samples were confirmed to be unphosphorylated by western blot using a standard phosphorylated-Tyr antibody.

Activity, IC₅₀, and K_d measurements

Protein activity was assayed using the Antibody Beacon™ Tyrosine Kinase Assay Kit (Molecular Probes). In addition to kit components the reaction mixture contained 10-50 nM of protein, 500 µM of standard peptide EAIYAAPFAKKK, 20 mM MgCl₂, and 1mM ATP. Phosphorylated peptides of known concentration were used for fluorescence level calibration. All reactions were performed at 25 °C. K_i values for Gleevec were calculated from IC₅₀'s using the standard equation:

$$K_i = IC_{50}/(1 + \left(\frac{[ATP]}{K_m}\right))$$
, where we used a K_m for ATP of 70 µM.

For several samples, the resulting rates were validated by HPLC analysis of the reaction products using Agilent Infinity 1260 and C18-AR columns from ACE. Phosphorylated and unphosphorylated peptides were separated using a linear gradient between 0 and 40% of acetonitrile with 0.1% TFA as a mobile phase. The results of both methods were identical within experimental error.

For dissociation constant (K_d) measurements of Gleevec to the ancestors, 10 nM of kinase was mixed with 2-1000 nM of Gleevec. Binding was monitored via changes in

intrinsic Trp fluorescence. Measurements were done using the FluoroMax-4 (Jobin-Yvon) fluorimeter. Tryptophanes were excited at 295 nm and fluorescence was detected at 350 nm. Extracted intensities were fitted to a generalized binding equation:

$$F = F_0 + A \cdot \frac{[I] + [E_t] + K_d - \sqrt{([I] + [E_t] + K_d)^2 - 4 \cdot [E_t] \cdot [I]}}{2 \cdot [E_t]},$$

where $[E_t]$ is total enzyme concentration, $[I]$ concentration of Gleevec, F_0 and A are background fluorescence and a scaling factor, respectively.

The dissociation constant (K_d) of Gleevec Abl (T315I) could not be determined by Trp fluorescence because of too weak binding and severe inner filter effects of the drug at the high concentrations. Only ITC at 25 °C gave a reliable data for the Gleevec affinity to Abl (T315I). Titrations were carried out on a Nano ITC (TA instruments) and analyzed with the NanoAnalyze software. Injectant was added in 1 µL volume, every 180 s. The concentrations used were 25 µM Abl (T315I) and 340 µM Gleevec.

X-ray crystallography

Hexagonal crystals of ancestor ANC-AS with bound AMPPCP (with dimensions $h = 50$ - 100 µm, $a = 20$ µm) were grown for three days and were flash frozen in liquid nitrogen. 6.3 mg/mL of lysine modified (ethylated) protein was crystallized at 18 °C using the hanging-drop method in 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% (v/v) glycerol, 20 mM MgCl₂, 2 mM imidazole, 1 mM AMPPCP, mixed 1:1 with 2.2 M ammonium sulfate. The data were indexed and integrated with XDS (37) while scaling was done with AIMLESS (38) (CCP4). Molecular replacement was performed with CCP4 MOLREP (39) using a human Abl kinase structure (pdb code 2HYY) as an initial search model. Model refinement was carried out using PHENIX (40) and CCP4 REFMAC (41). Models were built using COOT and WINCOOT (42) Molecular replacement and the first refinement cycles were done without the nucleotide and the magnesium ion in the model. Later, AMPPCP was placed into the positive peak of the difference electron density map. No density could be confidently determined for the magnesium ion. In an effort to minimize model bias, simulated annealing (both Cartesian and torsion angles) was performed with PHENIX (40) using default parameters for several rounds. Tables 1 and 2 summarize the data

collection/processing statistics and the refinement statistics. Model validation was done with MOLPROBITY (43).

Two-dimensional plates of ANC-AS with Gleevec bound (with dimensions 300 $\mu\text{m} \times$ 300 μm) grew within one week on dust particles using the sitting-drop method. These crystals were later used for microseeding using the hanging-drop method. Smaller but three-dimensional plates (100 $\mu\text{m} \times$ 100 $\mu\text{m} \times$ 15 μm) were flash frozen in liquid nitrogen. For both steps, 10mg/ml of lysine modified (ethylated) protein was used in 30 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM Gleevec, mixed 1:1 with 200 mM ammonium acetate, 100 mM sodium acetate trihydrate, pH 4.6 and 30% (w/v) PEG 4000, at 18 °C. Data processing and model building was done as described for protein in the presence of AMPPCP.

ThermoFluor experiments

Solutions of 10 μl of 225X Sypro Orange, 15 μl of storage buffer (50 mM HEPES pH 8, 500 mM NaCl, 5% (v/v) glycerol and 10 mM TCEP) and 5 μl of 100 μM protein was added to a 96-well PCR plate. A control containing the storage buffer + Sypro Orange was added. The plates were sealed with optical sealing tape and heated in an Applied Biosystems 9600 real-time PCR machine from 20 to 100 degrees Celsius with increments of 0.2 °C. Fluorescence of the Sypro orange dye was measured by exciting at 490 nm and measuring at 575 nm.

Stopped-flow kinetics experiments and data analysis

All stopped-flow experiments were performed with the Applied Photophysics SX-20 instrument at 5 °C or 25 °C as specified in the text. Binding was monitored via changes in intrinsic tryptophan fluorescence, samples were excited at 295 nm (9 nm bandwidth) and emission was detected using a long-pass 320 nm cut-off filter. After mixing the concentration of kinase was 0.1 μM , and the concentration of Gleevec was varied. To study dissociation kinetics, protein (at 0.1-1 μM) was pre-incubated with 0.1-100 μM of Gleevec (depending on the K_d of the kinase) for 10 minutes, placed into the 0.5 mL syringe and then diluted 11-fold. All experiments were performed in a buffer containing 50 mM Tris-HCl, 500 mM NaCl, 1 mM MgCl₂, 1 mM TCEP and 5% (v/v) DMSO (pH

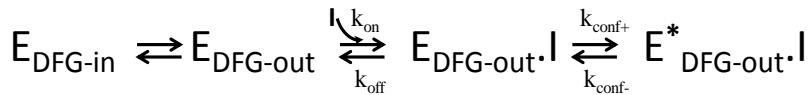
8.0). Data were analyzed using Applied Photophysics software. Kinetic fluorescence traces were fitted to a single or multi-exponential function. To account for photobleaching, an additional exponential term was included into the fitting function. This rate was fixed to the value determined in control experiments where protein was mixed with buffer in the absence of Gleevec.

Analysis of kinetic data

The following naming convention is used throughout the text:

- Different states of the enzyme without or with bound inhibitor in two different enzyme conformations are called E, E.I and E*.I, respectively.
- The conformation of the DFG-loop is specified with “DFG-in” or “DFG-out” subscripts (e.g., E_{DFG-in}, E_{DFG-out}, E_{DFG-out.I}).
- Rates describing the time dependence of experimentally observed changes in fluorescence are called “observed rates”.
- k_{on}, k_{off}, k_{conf+}, and k_{conf-} are called “rate constants” and correspond to individual microscopic steps in the reaction schemes.
- F denotes the amplitude of the observed fluorescent signal and is generated by combined fluorescence from all enzyme species.

As detailed in (11) Gleevec binding is defined by the following kinetic scheme:



In this scheme the first step (conformational selection), E_{DFG-in} ⇌ E_{DFG-out}, is fast and not directly observed in the kinetic experiment. However, the equilibrium between these two states affects the population of the binding competent state (E_{DFG-out}) and hence is reflected in the amplitude of the next step in the scheme (the binding step). This phenomenon allowed us to qualitatively track the evolutionary change in the DFG-in/DFG-out equilibrium along the phylogenetic tree (Fig. 2A).

In all of our binding kinetic experiments, concentration of the inhibitor was much higher than concentration of the enzyme ($[I] \gg [E]$). Under such conditions the binding is a pseudo-first-order reaction ($E_{DFG-out} + I \rightleftharpoons E_{DFG-out}\cdot I$) and thus characterized by a linear dependence of the observed binding rate (k_{fast}) on inhibitor concentration (Figs. 1G, 4A, and S11A). This linear dependence is the feature that allows clear identification of the phase corresponding to binding in the multi-exponential kinetic traces. In contrast, the observed rate that characterizes the conformational change after binding (the induced fit step, k_{slow}) has a non-linear dependence on inhibitor concentration, since the transient concentration of the $E_{DFG-out}\cdot I$ depends on inhibitor concentration (Figs. 1H, 4B, and S11B) (11).

These plots of k_{fast} , k_{slow} as a function of inhibitor concentration can be used to extract the microscopic rate constants for different steps of the binding scheme. From the linear plot of k_{slow} vs. $[I]$ one can extract the k_{on}^{obs} (which is equal to the slope of the line) and the k_{off} (which is equal to the intercept). We note that k_{on}^{obs} is not a microscopic rate constant, but is the product of k_{on} and the fractional population of the kinase in the binding capable state $P_{DFG-out}$: $k_{on}^{obs} = k_{on} \times P_{DFG-out}$. As a consequence, k_{on}^{obs} reflects both the $E_{DFG-in} \rightleftharpoons E_{DFG-out}$ equilibrium and the rate of the physical binding step.

The Gleevec dissociation experiment was used to determine the k_{conf-} rate constant. Since the fluorescent change observed in this experiment was mono-exponential and much slower than the k_{off} (determined as described above), the rate constant characterizing the dissociation must be attributed to the conformational change k_{conf-} . In addition, the value of the plateau in the k_{slow} vs $[I]$ graph (Figs. 1H, 4B, and S11B) determines the sum $k_{conf+} + k_{conf-}$, which allows calculating the value of k_{conf+} . Thereby the system is fully determined (11).

Knowledge of the individual microscopic constants enables calculation of the overall K_d^{calc} :

$$K_d^{calc} = \frac{K_{bind}^{obs} \cdot K_{IF}}{(1 + K_{IF})}$$

where K_d^{calc} is the overall dissociation constant, K_{bind}^{obs} is the observed dissociation constant for binding that includes the conformational selection equilibrium, and K_{IF}

corresponds to the equilibrium constant for the induced fit step, respectively. This calculated K_d^{calc} can be compared with the value of K_d^{measured} (Fig. S6), which was determined in an independent thermodynamic experiment (Figs. S4 and S10). Such a comparison serves as an independent verification of the model and the determined parameters.

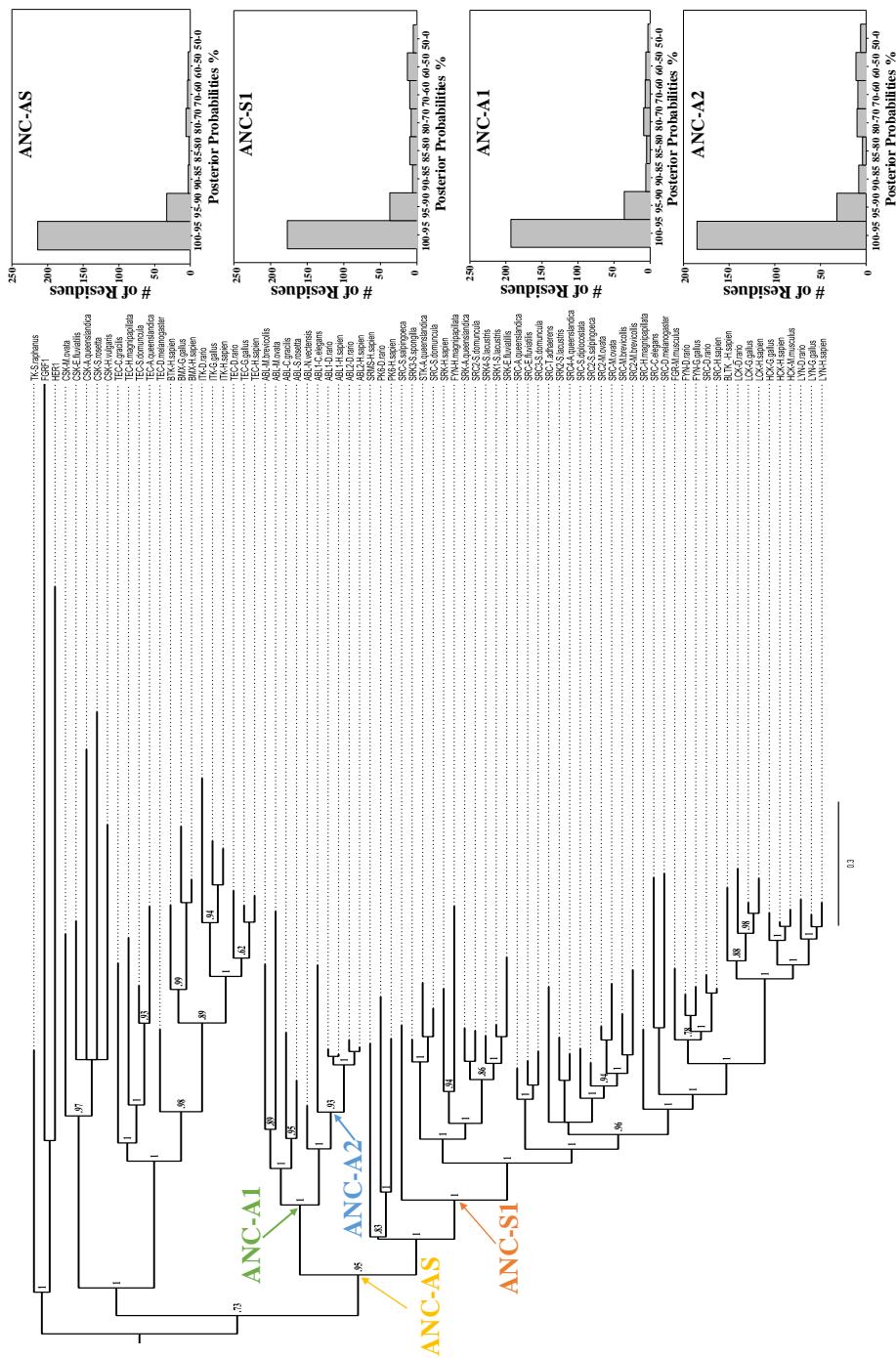
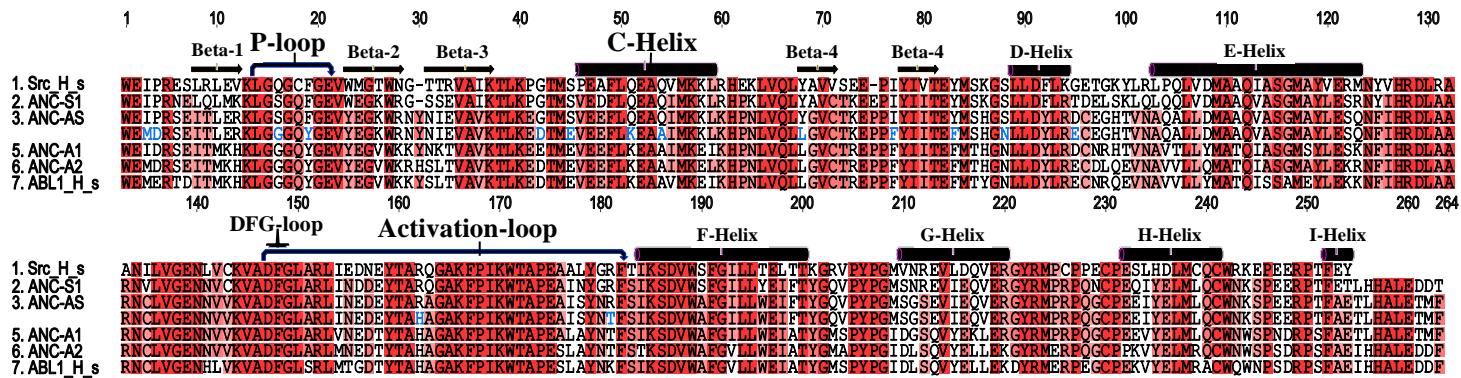


Fig. S1.

Full phylogenetic tree of the Src/Abl/Tec tyrosine kinases and binned probabilities of ancestral state reconstruction. Left: Phylogenetic tree of Src/Abl/Tec tyrosine kinases. Posterior probabilities $>.50$ for each node are shown. Resurrected nodes are labeled and colored using the color scheme presented in the manuscript. Right: Histograms of the posterior probabilities at each residue.

Fig. S2

Multiple sequence alignment of Src, Abl, and reconstructed ancestral sequences. Conserved and similar residues are colored in red and pink, respectively. Mutations between ANC-AS and AS(+15) responsible for changing Gleevec affinity to levels comparable to Abl are marked in blue.



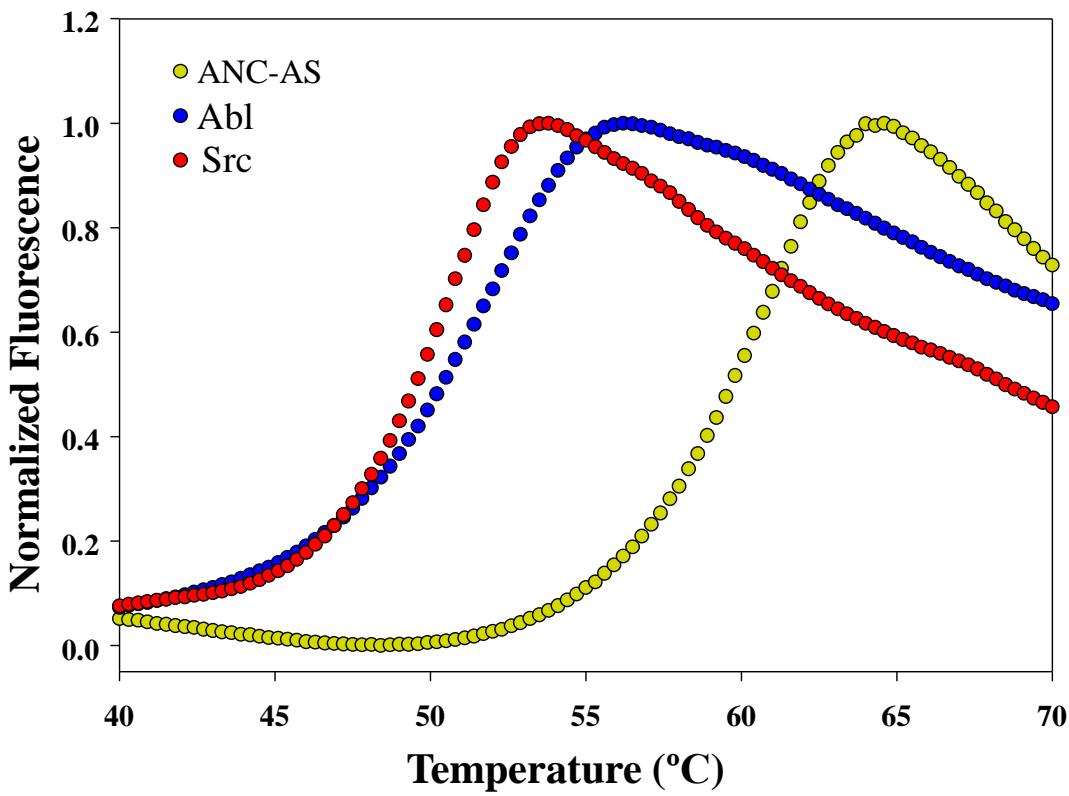


Fig. S3

ANC-AS has a higher melting temperature than Src and Abl. We used the ThermoFluor assay (15) to measure the melting temperatures of Src, Abl, and ANC-AS. The dye Sypro Orange is similar in structure to ANS (1-anilinonaphthalene-8-sulfonate), and when bound to hydrophobic patches that are exposed in denatured proteins, is fluorescent. Utilizing a RT-PCR machine the temperature was gradually changed between 20 and 100 °C and the increase in fluorescence was monitored. The melting temperature of ANC-AS is 11 degrees higher than for modern Src and Abl, suggesting an increase in thermal stability. This is in accordance with other resurrections in which ancestral proteins show higher melting temperatures than their modern day counter parts.

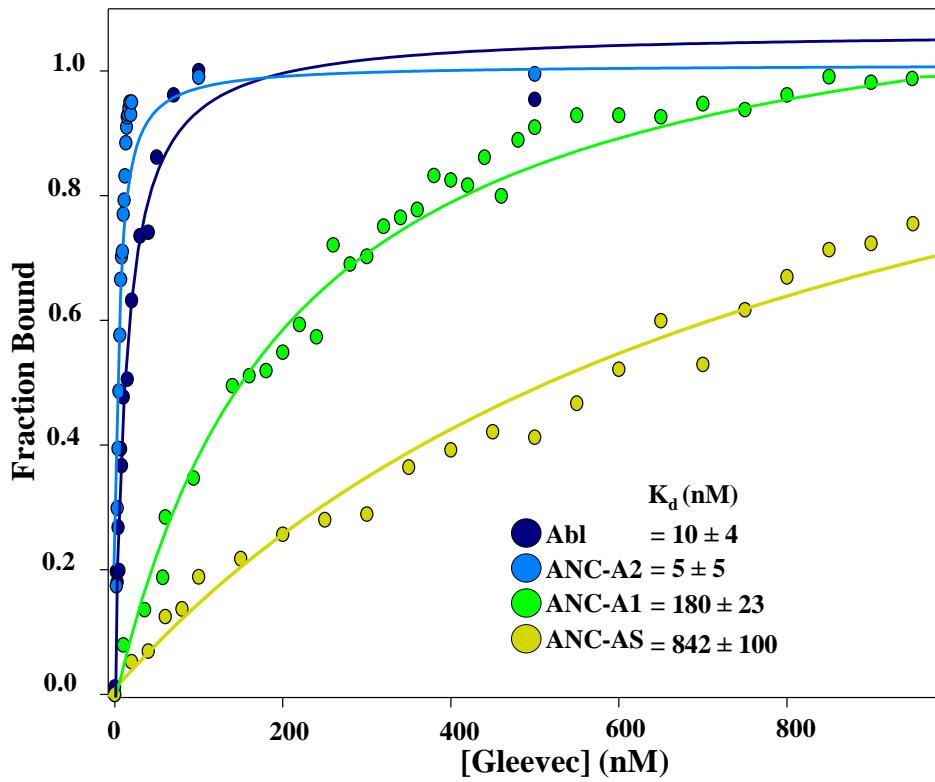


Fig. S4

Dissociation constants (K_d) for Gleevec at 5 °C. K_d 's were measured by fluorescence quenching (K_d values for the weak binders Src and ANC-S1 could not be well determined because of inner-filter effects at high Gleevec concentrations).

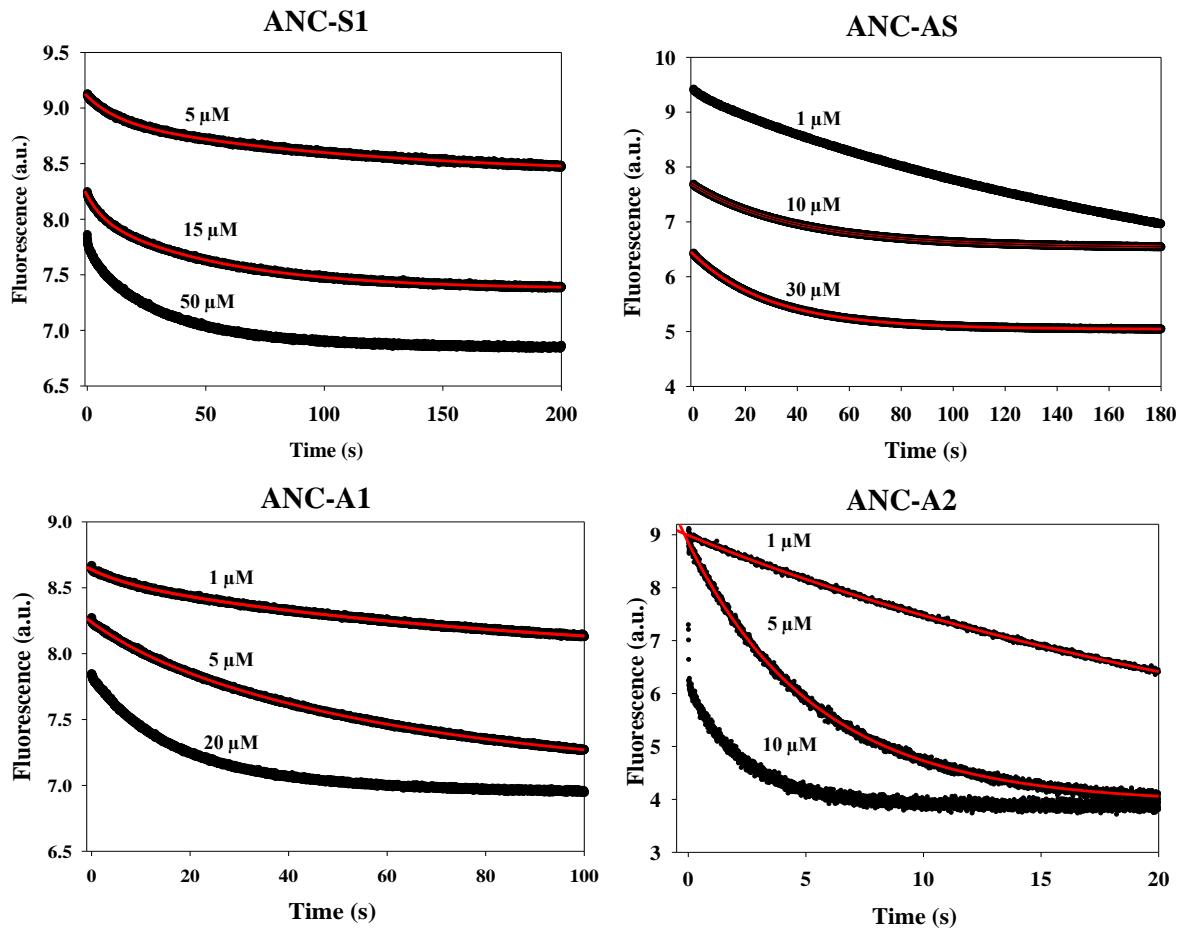


Fig. S5

Kinetics of Gleevec binding to all ancestors at 5 °C. Representative time traces of tryptophan fluorescence quenching at different Gleevec concentrations, measured by stopped-flow fluorescence, are shown in black. 50 nM kinase was mixed with varying Gleevec concentrations and fit to double-exponentials (red). Complete results are shown in Fig. 1G and 1H.

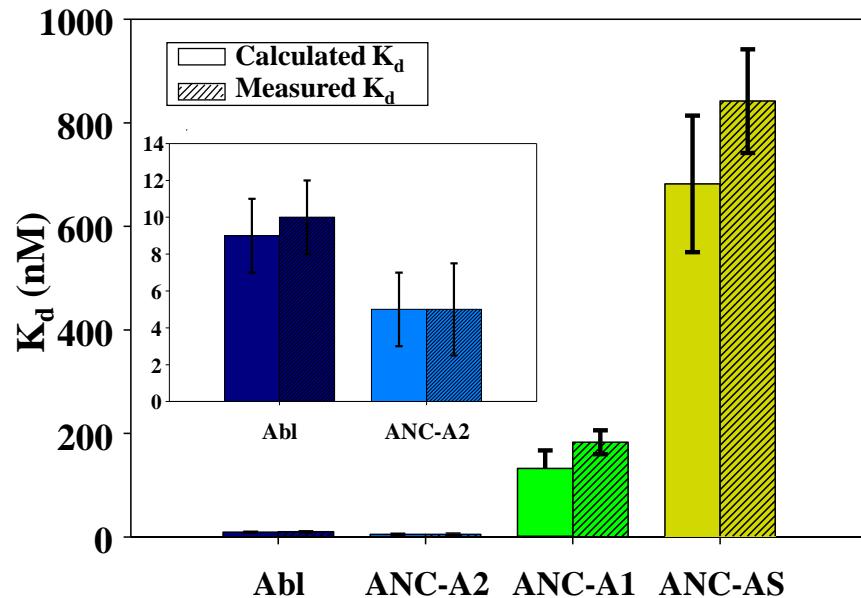


Fig. S6

Stringent test for validity of kinetic scheme and measured rate constants for Abl and several ancestors. Measured K_d 's (Fig. S6) are within experimental error to K_d 's calculated from all microscopic rate constants (Fig. 2). For the very weak Gleevec binders Src and ANC-S1, the K_d is too weak to measure accurately (3) and also the K_d 's cannot be calculated from the kinetic rate constants because the amplitude of the fast phase is too small (Fig. 2A).

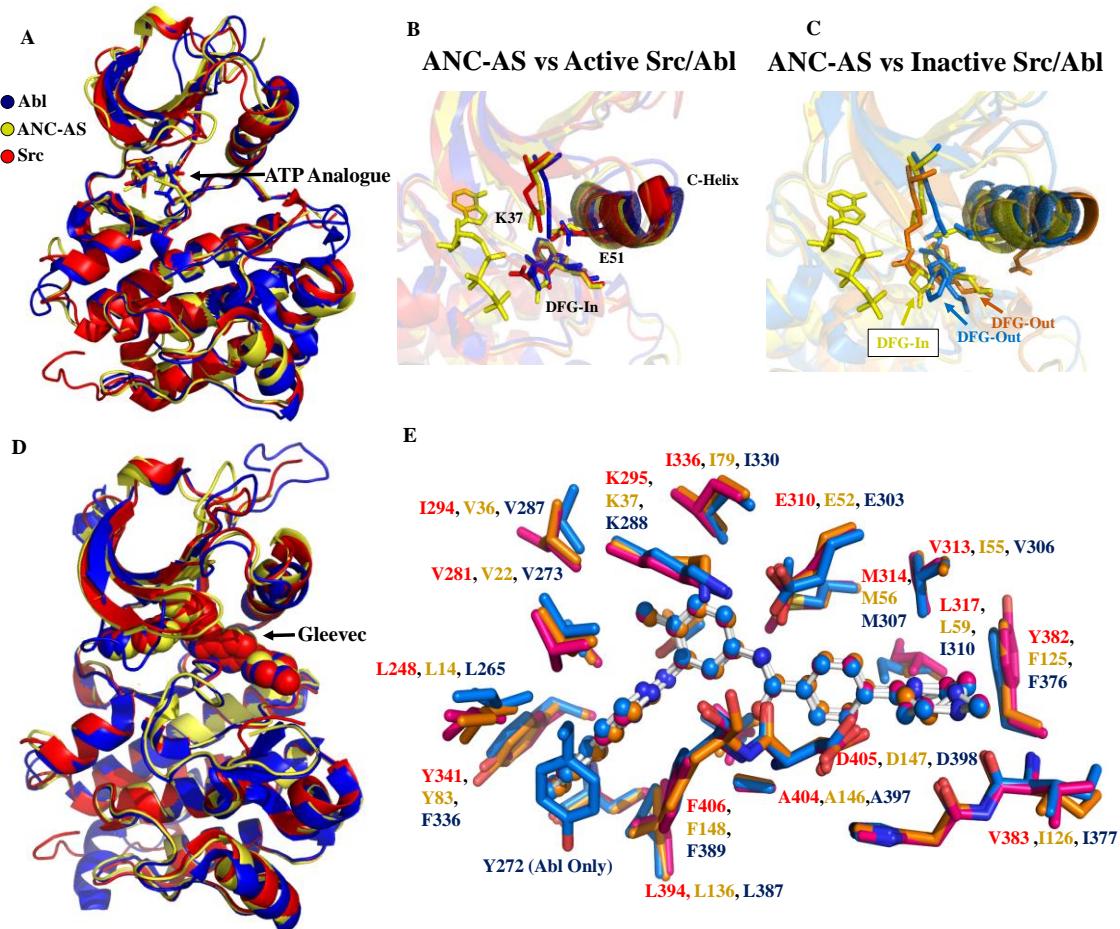


Fig. S7

Comparison of ANC-AS bound to AMPPCP with active and inactive Src and Abl, and inactive ANC-AS bound to Gleevec with the corresponding Abl and Src structures. ANC-AS samples a fully active state without phosphorylation of Tyr 159 in the activation loop. (A) 2.9 Å x-ray structure of the last common ancestor ANC-AS bound to AMPPCP (4UEU, gold) is superimposed with active conformations of Src (2BDF, red) (45) and Abl (2G2I, blue) (46). ANC-AS is very similar in structure to the modern day kinases with an overall RMSD of 1 Å to Abl and 0.89 Å to Src. Zoom into the DFG loop and active site highlights that nonphosphorylated ANC-AS (gold) crystallized in the active state with DFG-in (blue is Abl active state and red is Src active state) (B), and there is a poor superposition with the corresponding inactive structures of Abl (light blue, 2G2F) and Src (orange, 2SRC) (C). (D) 2.1 Å x-ray structure of the last common ancestor ANC-AS bound to Gleevec (4CSV, gold) is superimposed with the Gleevec-bound structures of Src (2OIQ, red) (3) and Abl (1OPJ, blue) (47). (E) Comparison of the corresponding Gleevec binding pockets show that they are nearly identical. Gleevec is shown in ball-and-stick representation.

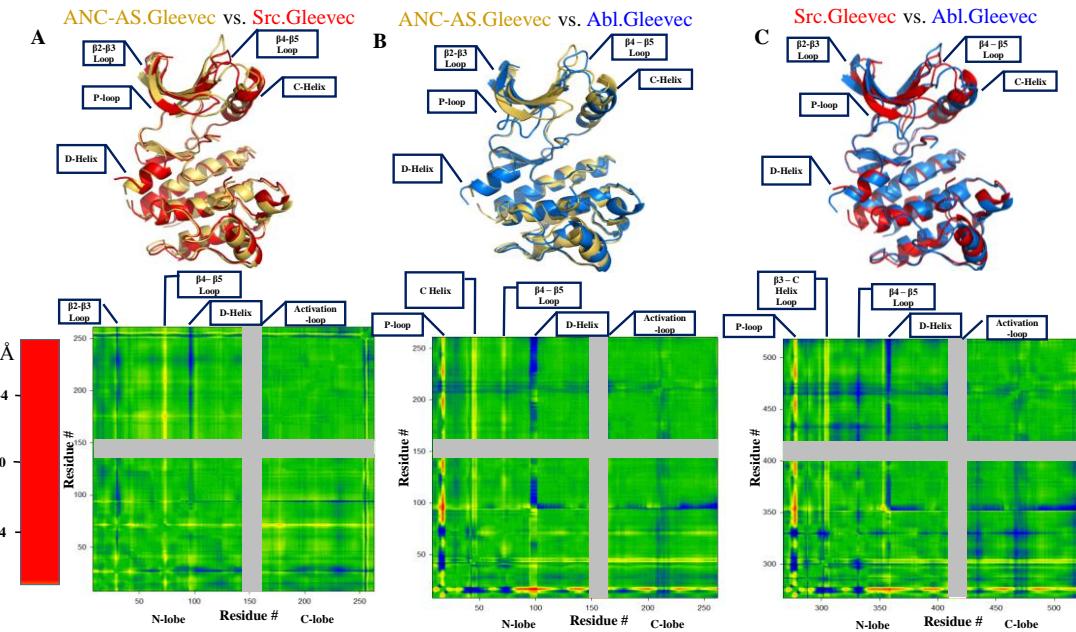


Fig. S8

Comparison of crystal structure of the ANC-AS.Gleevec complex with structures of Src and Abl bound to Gleevec (same pdb's as in S7). X-ray structures are superimposed using SuperPose (48) and the corresponding $\text{Ca}-\text{Ca}$ distance maps are shown below. Ca atoms that are moving closer in distance relative to each Ca atom in the other structure are indicated in blue, and atoms that move further away are colored in yellow/red. The activation loop is not included in the difference maps (grey) because of too weak electron density. (A) The Ca RMSD between ANC-AS.Gleevec and Src.Gleevec structures is 1.35 Å with the major differences in $\beta 2-\beta 3$ loop, $\beta 4-\beta 5$ loop and D-helix. (B) The Ca RMSD between ANC-AS.Gleevec and Abl.Gleevec is 1.90 Å illustrating that ANC-AS.Gleevec aligns better to Src compared to Abl. Three major differences between ANC-AS.Gleevec and Abl.Gleevec are in the P-loop, the D-helix, and $\beta 3-\text{C-Helix}$ loop (C) The Ca RMSD between Abl.Gleevec and Src.Gleevec structures is 2.02 Å with the major changes as reported before (3).

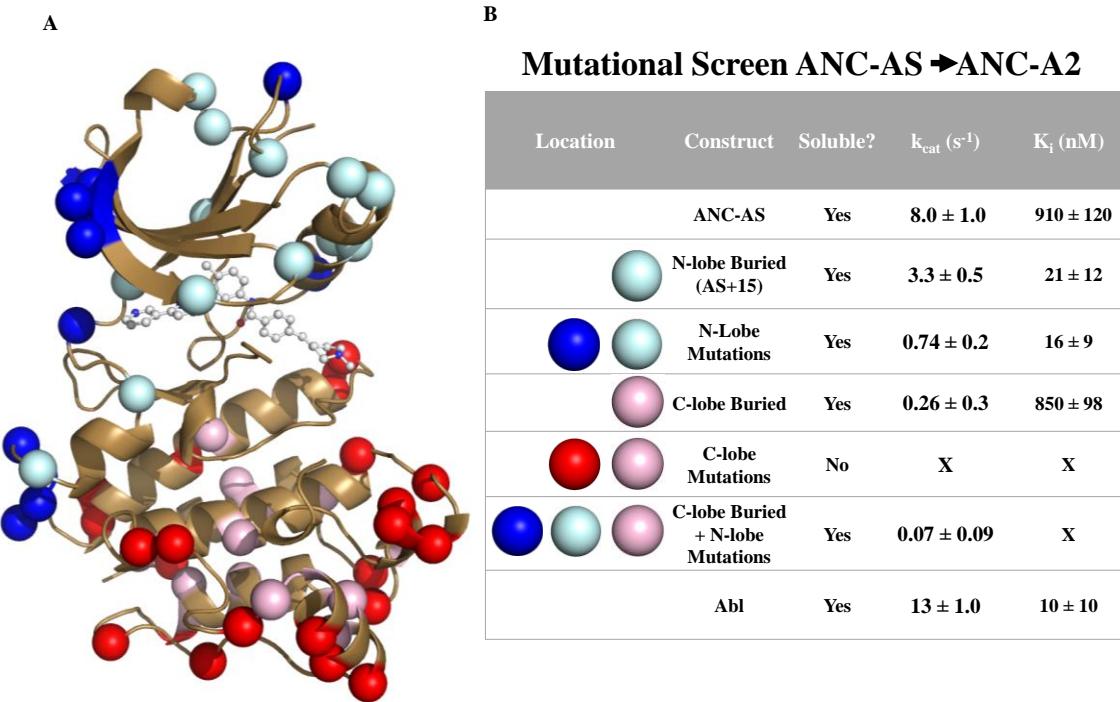


Fig. S9

Mutational screen between ANC-AS and ANC-A2 identifies subsets of mutations sufficient for major increase in Gleevec affinity. (A) Sequence differences between ANC-AS (weak binder) and ANC-A2 (tight binder) are plotted onto the ANC-AS.Gleevec structure as spheres. These residues were separated into two sets, N-lobe (blue) and C-lobe (red) mutations and then further split into buried residues (light red and light blue) and solvent-exposed residues (dark red and dark blue) (see B). Such crude divide-and-conquer approach was motivated by the simple scientific intuition that solvent exposed residues are less likely be important for binding of a hydrophobic compound to the interior of the protein or intramolecular interactions. Dividing by N-terminal and C-terminal lobe was aimed at figuring out whether long-range interactions all the way into the C-terminal domain could play a role, which is not the case as revealed by the results. (B) Biochemical characterization of the constructs carrying different combinations of mutations. Solubility, activity and inhibition constants for Gleevec for each of the constructs, ANC-AS and Abl are added for comparison. All constructs with mutational sets involving the C-lobe are either insoluble, or have severely compromised activity. Importantly, N-lobe buried mutations are sufficient to switch affinity for Gleevec to levels corresponding to tight binders.

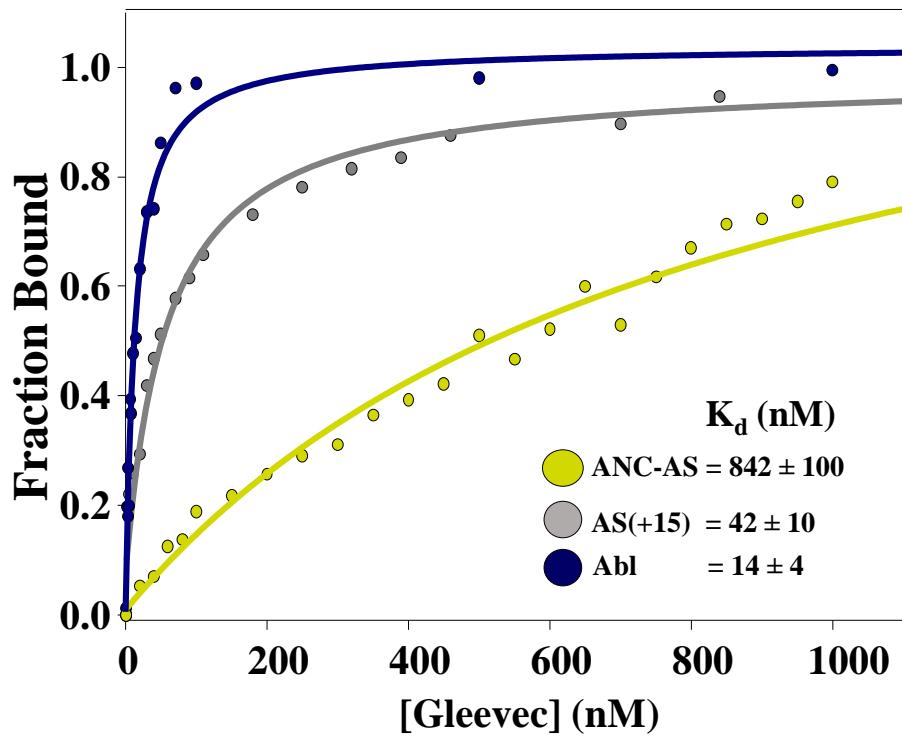


Fig. S10

Comparison of the dissociation constants (K_d) for Gleevec between ANC-AS, AS(+15), and Abl. Dissociation constants were measured by fluorescence quenching at 5 °C.

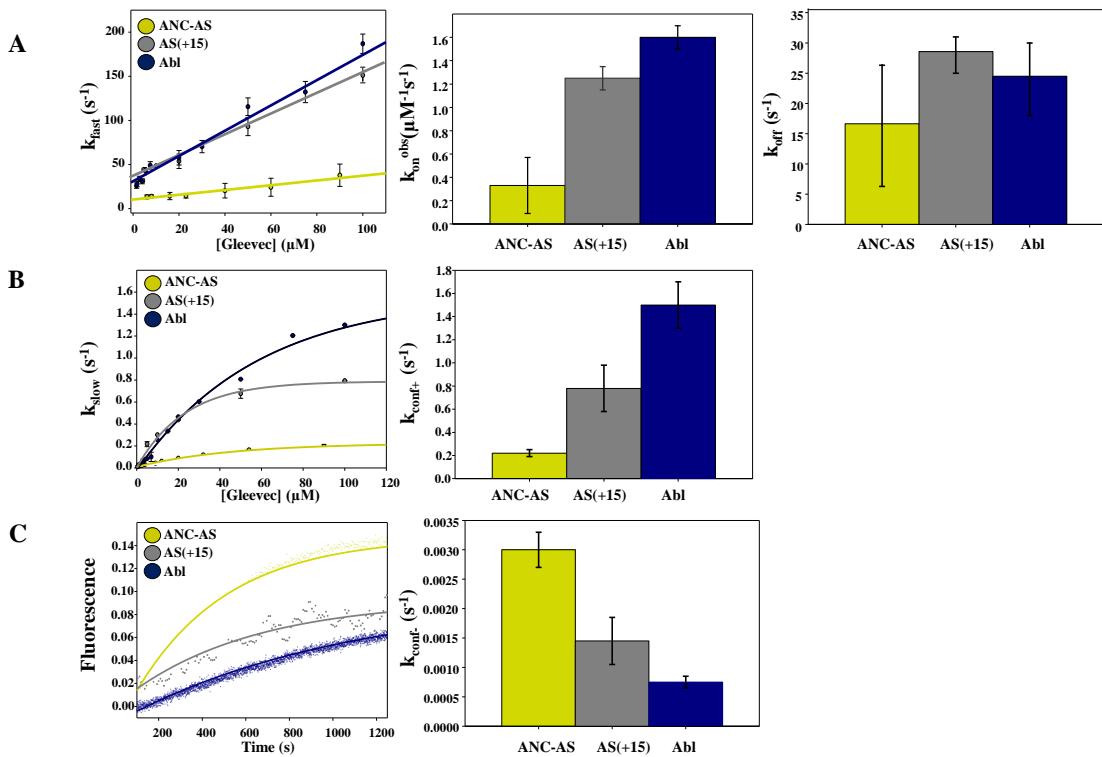


Fig. S11

Comparison of the kinetic properties between ANC-AS, AS(+15) and Abl. (A, B) Kinetic of binding (left) was measured by stopped-flow fluorescence at 5° C. 50-100 nM of kinase was mixed with 1-100 μM of Gleevec. The fast phase (A) corresponds to the binding step and the slow step (B) corresponds to the induced fit step (see Scheme 1 and Fig. 1). Values for each of the observed rate are shown on the right. (C) Gleevec dissociation initiated by 11-fold dilution of equimolar kinase-Gleevec complex (experimental conditions are the same as in Fig. 1F). The rate of fluorescence change is determined by E*.I to E.I interconversion ($k_{\text{conf-}}$), while k_{off} is much faster (intercept in A).

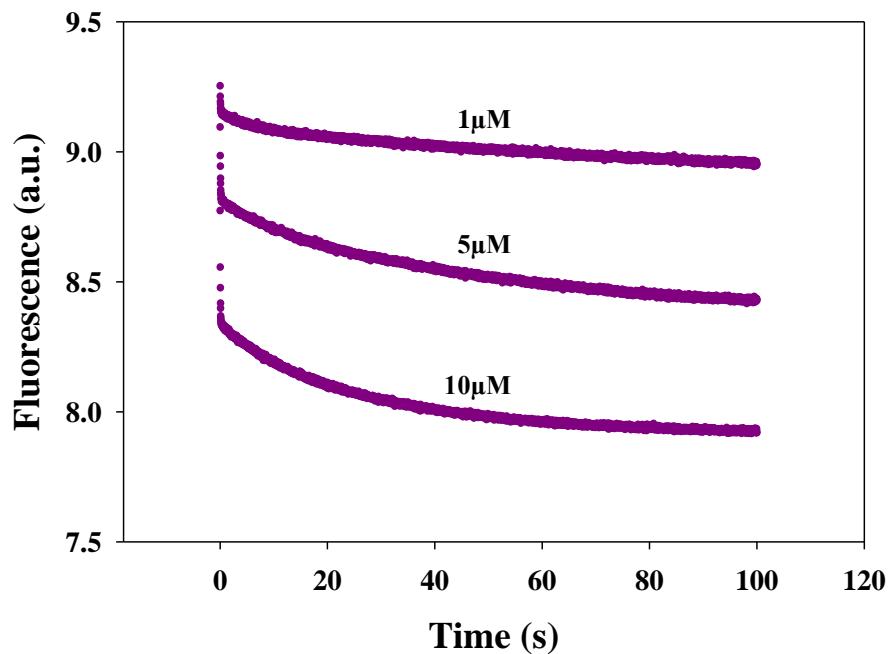


Fig. S12

Kinetics of Gleevec binding to Abl (T315I) at 5 °C. Representative time traces of tryptophan fluorescence quenching at different Gleevec concentrations, measured by stopped-flow fluorescence, are shown in magenta. 50 nM kinase was mixed with varying Gleevec concentrations and fit to double-exponentials.

Table S1.

X-ray data collection and processing statistics. Values in parentheses correspond to the highest-resolution shell.

	ANC-AS bound to AMPPCP (pdb id: 4UEU)	ANC-AS bound to Gleevec (pdb id: 4CSV)
Beamline	BL8.2.2 ALS	BL8.2.2 ALS
Detector	Quantum 315 ADSC Area Detector	Quantum 315 ADSC Area Detector
Wavelength (Å)	0.999995	0.999956
Temperature (K)	100	100
Crystal-detector distance (mm)	375	285
Oscillation range per image (°)	0.5	0.5
Total oscillation range (°)	180	122
Space group	P6 ₂	I2
Unit cell parameter (Å, °)	128.49×128.49×52.81, 90×90×120	71.03×56.87×76.12, 90×116.62×90
Mosaicity	0.36	0.32
Resolution limits (Å)	64.24-2.949 (3.18-2.949)	42.36-2.049 (2.11-2.049)
Total number of reflection	104864	40153
Unique reflections	10022	15954
Redundancy	10.5	2.5
I/σ (I)	15.5 (3.1)	10.1 (2.0)
Completeness (%)	93.4 (98.2)	93.1 (96.8)
R _{merge} (%)	9.7 (69.1)	4.3 (41.0)
R _{meas} (%)	10.7 (76.3)	6.1 (57.8)
R _{p.i.m.} (%)	4.5 (31.9)	4.3 (40.7)
Overall B-factor from Wilson plot (Å ²)	60.9	43.9

Table S2.

Refinement statistics. Values in parentheses correspond to the highest-resolution shell.

Resolution range (Å)	64.33-2.949 (3.026-2.949)	62.43-2.049 (2.102-2.049)
Total number of atoms (nonhydrogen)	2095	2061
Number of protein atoms	2064	1968
R _{cryst} (%)	19.2 (31.3)	19.0 (30.5)
R _{free} (%)	25.3 (37.0)	24.2 (37.3)
RMSD from ideality; bonds (Å), angles (°)	0.012, 1.616	0.010, 1.420
Ramachandran plot, favored regions (%)	86.1	97.5
Rotamer outliers (%)	0.9	3.3
C-beta outliers	0	0
Molprobity clashscore	13.53	3.78
Molprobity overall score	2.28	1.67
Average B-factor (Å ²)	91.4	59.5

The structures of the catalytic domain of ANC-AS in its active and inhibited state were solved at 2.95 Å and 2.05 Å, respectively. One monomeric kinase molecule can be found in the asymmetric unit cell of both crystal structures. The activation loop and the C-terminal residues in both models, and the residues 97-100 (inhibited state model only) could not be traced into the electron density map. The P-loop (residues 17-24) and the loop between the D- and E-helix (residues 95-100) of the active state model have high B-factors (\approx two times of the average B-factor). However, there is enough main chain density to model these amino acids.

References

1. D. J. Richter, N. King, The genomic and cellular foundations of animal origins. *Annu. Rev. Genet.* **47**, 509–537 (2013). [Medline](#) [doi:10.1146/annurev-genet-111212-133456](https://doi.org/10.1146/annurev-genet-111212-133456)
2. G. Manning, G. D. Plowman, T. Hunter, S. Sudarsanam, Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* **27**, 514–520 (2002). [Medline](#) [doi:10.1016/S0968-0004\(02\)02179-5](https://doi.org/10.1016/S0968-0004(02)02179-5)
3. M. A. Seeliger, B. Nagar, F. Frank, X. Cao, M. N. Henderson, J. Kuriyan, c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty. *Structure* **15**, 299–311 (2007). [Medline](#) [doi:10.1016/j.str.2007.01.015](https://doi.org/10.1016/j.str.2007.01.015)
4. S. S. Taylor, A. P. Kornev, Protein kinases: Evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* **36**, 65–77 (2011). [Medline](#) [doi:10.1016/j.tibs.2010.09.006](https://doi.org/10.1016/j.tibs.2010.09.006)
5. A. P. Kornev, S. S. Taylor, Defining the conserved internal architecture of a protein kinase. *Biochim. Biophys. Acta* **1804**, 440–444 (2010). [Medline](#) [doi:10.1016/j.bbapap.2009.10.017](https://doi.org/10.1016/j.bbapap.2009.10.017)
6. Y.-L. Lin, Y. Meng, W. Jiang, B. Roux, Explaining why Gleevec is a specific and potent inhibitor of Abl kinase. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1664–1669 (2013). [Medline](#) [doi:10.1073/pnas.1214330110](https://doi.org/10.1073/pnas.1214330110)
7. A. Aleksandrov, T. Simonson, Molecular dynamics simulations show that conformational selection governs the binding preferences of imatinib for several tyrosine kinases. *J. Biol. Chem.* **285**, 13807–13815 (2010). [Medline](#) [doi:10.1074/jbc.M110.109660](https://doi.org/10.1074/jbc.M110.109660)
8. S. Lovera, L. Sutto, R. Boubeva, L. Scapozza, N. Dölker, F. L. Gervasio, The different flexibility of c-Src and c-Abl kinases regulates the accessibility of a druggable inactive conformation. *J. Am. Chem. Soc.* **134**, 2496–2499 (2012). [Medline](#) [doi:10.1021/ja210751t](https://doi.org/10.1021/ja210751t)
9. S. W. Cowan-Jacob, G. Fendrich, P. W. Manley, W. Jahnke, D. Fabbro, J. Liebetanz, T. Meyer, The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13**, 861–871 (2005). [Medline](#) [doi:10.1016/j.str.2005.03.012](https://doi.org/10.1016/j.str.2005.03.012)
10. A. C. Dar, K. M. Shokat, The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. *Annu. Rev. Biochem.* **80**, 769–795 (2011). [Medline](#) [doi:10.1146/annurev-biochem-090308-173656](https://doi.org/10.1146/annurev-biochem-090308-173656)
11. R. V. Agafonov, C. Wilson, R. Otten, V. Buosi, D. Kern, Energetic dissection of Gleevec's selectivity toward human tyrosine kinases. *Nat. Struct. Mol. Biol.* **21**, 848–853 (2014). [Medline](#) [doi:10.1038/nsmb.2891](https://doi.org/10.1038/nsmb.2891)
12. Y. Shan, M. A. Seeliger, M. P. Eastwood, F. Frank, H. Xu, M. Ø. Jensen, R. O. Dror, J. Kuriyan, D. E. Shaw, A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 139–144 (2009). [Medline](#) [doi:10.1073/pnas.0811223106](https://doi.org/10.1073/pnas.0811223106)

13. M. A. DePristo, D. M. Weinreich, D. L. Hartl, Missense meanderings in sequence space: A biophysical view of protein evolution. *Nat. Rev. Genet.* **6**, 678–687 (2005). [Medline](#) [doi:10.1038/nrg1672](#)
14. M. J. Harms, J. W. Thornton, Evolutionary biochemistry: Revealing the historical and physical causes of protein properties. *Nat. Rev. Genet.* **14**, 559–571 (2013). [Medline](#) [doi:10.1038/nrg3540](#)
15. L. Pauling, E. Zuckerkandl, Molecular “restoration studies” of extinct forms of life. *Acta Chem. Scand.* **17**, S9–S16 (1963).
16. A. M. Dean, J. W. Thornton, Mechanistic approaches to the study of evolution: The functional synthesis. *Nat. Rev. Genet.* **8**, 675–688 (2007). [Medline](#) [doi:10.1038/nrg2160](#)
17. D. A. Liberles, *Ancestral Sequence Reconstruction* (Oxford Univ. Press, Oxford, 2007).
18. P. D. Williams, D. D. Pollock, B. P. Blackburne, R. A. Goldstein, Assessing the accuracy of ancestral protein reconstruction methods. *PLOS Comput. Biol.* **2**, e69 (2006). [Medline](#)
19. N. M. Krishnan, H. Seligmann, C. B. Stewart, A. P. De Koning, D. D. Pollock, Ancestral sequence reconstruction in primate mitochondrial DNA: Compositional bias and effect on functional inference. *Mol. Biol. Evol.* **21**, 1871–1883 (2004). [Medline](#) [doi:10.1093/molbev/msh198](#)
20. T. Hunter, Tyrosine phosphorylation: Thirty years and counting. *Curr. Opin. Cell Biol.* **21**, 140–146 (2009). [Medline](#) [doi:10.1016/j.ceb.2009.01.028](#)
21. W. Eckhart, M. A. Hutchinson, T. Hunter, An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell* **18**, 925–933 (1979). [Medline](#) [doi:10.1016/0092-8674\(79\)90205-8](#)
22. Materials and methods are available as supplementary materials on *Science* Online.
23. M. A. Seeliger, P. Ranjitkar, C. Kasap, Y. Shan, D. E. Shaw, N. P. Shah, J. Kuriyan, D. J. Maly, Equally potent inhibition of c-Src and Abl by compounds that recognize inactive kinase conformations. *Cancer Res.* **69**, 2384–2392 (2009). [Medline](#) [doi:10.1158/0008-5472.CAN-08-3953](#)
24. M. E. Gorre, M. Mohammed, K. Ellwood, N. Hsu, R. Paquette, P. N. Rao, C. L. Sawyers, Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* **293**, 876–880 (2001). [Medline](#) [doi:10.1126/science.1062538](#)
25. M. Modugno, E. Casale, C. Soncini, P. Rosettani, R. Colombo, R. Lupi, L. Rusconi, D. Fancelli, P. Carpinelli, A. D. Cameron, A. Isacchi, J. Moll, Crystal structure of the T315I Abl mutant in complex with the aurora kinases inhibitor PHA-739358. *Cancer Res.* **67**, 7987–7990 (2007). [Medline](#) [doi:10.1158/0008-5472.CAN-07-1825](#)
26. H. Daub, K. Specht, A. Ullrich, Strategies to overcome resistance to targeted protein kinase inhibitors. *Nat. Rev. Drug Discov.* **3**, 1001–1010 (2004). [Medline](#) [doi:10.1038/nrd1579](#)
27. A. Ingles-Prieto, B. Ibarra-Molero, A. Delgado-Delgado, R. Perez-Jimenez, J. M. Fernandez, E. A. Gaucher, J. M. Sanchez-Ruiz, J. A. Gavira, Conservation of protein structure over four billion years. *Structure* **21**, 1690–1697 (2013). [Medline](#) [doi:10.1016/j.str.2013.06.020](#)

28. R. Perez-Jimenez, A. Inglés-Prieto, Z. M. Zhao, I. Sanchez-Romero, J. Alegre-Cebollada, P. Kosuri, S. Garcia-Manyes, T. J. Kappock, M. Tanokura, A. Holmgren, J. M. Sanchez-Ruiz, E. A. Gaucher, J. M. Fernandez, Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. *Nat. Struct. Mol. Biol.* **18**, 592–596 (2011). [Medline](#) [doi:10.1038/nsmb.2020](#)
29. E. A. Gaucher, S. Govindarajan, O. K. Ganesh, Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* **451**, 704–707 (2008). [Medline](#) [doi:10.1038/nature06510](#)
30. E. A. Gaucher, J. M. Thomson, M. F. Burgan, S. A. Benner, Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature* **425**, 285–288 (2003). [Medline](#) [doi:10.1038/nature01977](#)
31. M. J. Harms, J. W. Thornton, Analyzing protein structure and function using ancestral gene reconstruction. *Curr. Opin. Struct. Biol.* **20**, 360–366 (2010). [Medline](#) [doi:10.1016/j.sbi.2010.03.005](#)
32. S. F. Field, M. V. Matz, Retracing evolution of red fluorescence in GFP-like proteins from Faviina corals. *Mol. Biol. Evol.* **27**, 225–233 (2010). [Medline](#) [doi:10.1093/molbev/msp230](#)
33. B. Nagar, W. G. Bornmann, P. Pellicena, T. Schindler, D. R. Veach, W. T. Miller, B. Clarkson, J. Kuriyan, Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res.* **62**, 4236–4243 (2002). [Medline](#)
34. B. D. Redelings, M. A. Suchard, Joint Bayesian estimation of alignment and phylogeny. *Syst. Biol.* **54**, 401–418 (2005). [Medline](#) [doi:10.1080/10635150590947041](#)
35. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). [Medline](#) [doi:10.1093/molbev/msm088](#)
36. M. A. Seeliger, M. Young, M. N. Henderson, P. Pellicena, D. S. King, A. M. Falick, J. Kuriyan, High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein Sci.* **14**, 3135–3139 (2005). [Medline](#) [doi:10.1110/ps.051750905](#)
37. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010). [Medline](#) [doi:10.1107/S0907444909047337](#)
38. P. Evans, Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006). [Medline](#) [doi:10.1107/S0907444905036693](#)
39. A. Vagin, A. Teplyakov, MOLREP: An automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997). [doi:10.1107/S0021889897006766](#)
40. P. D. Adams, P. V. Afonine, G. Bunkóczki, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010). [Medline](#) [doi:10.1107/S0907444909052925](#)

41. A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, G. N. Murshudov, REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184–2195 (2004). [Medline doi:10.1107/S0907444904023510](#)
42. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). [Medline doi:10.1107/S0907444910007493](#)
43. V. B. Chen, W. B. Arendall 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). [Medline doi:10.1107/S0907444909042073](#)
44. U. B. Ericsson, B. M. Hallberg, G. T. Detitta, N. Dekker, P. Nordlund, Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* **357**, 289–298 (2006). [Medline doi:10.1016/j.ab.2006.07.027](#)
45. D. Dalgarno, T. Stehle, S. Narula, P. Schelling, M. R. van Schravendijk, S. Adams, L. Andrade, J. Keats, M. Ram, L. Jin, T. Grossman, I. MacNeil, C. Metcalf 3rd, W. Shakespeare, Y. Wang, T. Keenan, R. Sundaramoorthi, R. Bohacek, M. Weigle, T. Sawyer, Structural basis of Src tyrosine kinase inhibition with a new class of potent and selective trisubstituted purine-based compounds. *Chem. Biol. Drug Des.* **67**, 46–57 (2006). [Medline doi:10.1111/j.1747-0285.2005.00316.x](#)
46. N. M. Levinson, K. Shen, M. A. Young, M. Koldobskiy, M. Karplus, P. A. Cole, J. Kuriyan, A Src-like inactive conformation in the Abl tyrosine kinase domain. *PLOS Biol.* **4**, e144 (2006). [doi:10.1371/journal.pbio.0040144](#)
47. B. Nagar, O. Hantschel, M. A. Young, K. Scheffzek, D. Veach, W. Bornmann, B. Clarkson, G. Superti-Furga, J. Kuriyan, Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell* **112**, 859–871 (2003). [Medline doi:10.1016/S0092-8674\(03\)00194-6](#)
48. R. Maiti, G. H. Van Domselaar, H. Zhang, D. S. Wishart, SuperPose: A simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, W590–W594 (2004). [doi:10.1093/nar/gkh477](#) [Medline](#)