Classification: Biophysics and Computational Biology

Halogen-aromatic π interactions modulate inhibitor residence time

Christina Heroven[a], Victoria Georgi[b], Gaurav K. Ganotra[e], Paul E. Brennan[a,f], Finn Wolfreys[a,f], Rebecca C. Wade[e,h,i], Amaury E. Fernández-Montalván[b], Apirat Chaikuad[a,c,d,1], Stefan Knapp[a,c,d,g,1]

[a]Nuffield Department of Clinical Medicine, Structural Genomics Consortium, University of Oxford, Oxford, OX3 7DQ, UK

[b]Bayer AG, Drug Discovery, Pharmaceuticals, Lead Discovery Berlin, 13353 Berlin, Germany

[c]Buchmann Institute for Molecular Life Sciences, Johann Wolfgang Goethe-University, D-60438 Frankfurt am Main, DE

[d]Institute for Pharmaceutical Chemistry, Johann Wolfgang Goethe-University, D-60438 Frankfurt am Main, DE

[e]Molecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS), 69118 Heidelberg, Germany

[f]Target Discovery Institute, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, OX3 7FZ, UK

[g]German Cancer Network (DKTK), Frankfurt/Mainz site, D-60438 Frankfurt am Main, DE

[h]Zentrum für Molekulare Biologie, DKFZ-ZMBH Alliance, Heidelberg University, 69120 Heidelberg, Germany;

[i]Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany

[1] Corresponding authors: Apirat Chaikuad (chaikuad@pharmchem.uni-frankfurt.de) or Stefan Knapp (knapp@pharmchem.uni-frankfurt.de), Institute for Pharmaceutical Chemistry, Johann Wolfgang Goethe-University, Max-von-Laue Str. 9, D-60438 Frankfurt am Main, Germany. Tel: +49(0)6979829871

Running title: Residence time modulation by halogens
Keywords: Halogen-π interaction, drug residence time, kinase

ABSTRACT

Prolonged drug residence times may result in longer lasting drug efficacy, improved pharmacodynamic properties and "kinetic selectivity" over off-targets with fast drug dissociation rates. However, few strategies have been elaborated to rationally modulate drug residence time and thereby to integrate this key property into the drug development process. Here, we show that the interaction between a halogen moiety on an inhibitor and an aromatic residue in the target protein can significantly increase inhibitor residence time. By using the interaction of the serine/threonine kinase haspin with 5-iodotubercidin (5-iTU) derivatives as a model for an archetypal active state (type I) kinase-inhibitor binding mode, we demonstrate that inhibitor residence times markedly increase with the size and polarizability of the halogen atom. This key interaction is dependent on the interactions with an aromatic residue in the gate keeper position and we observe this interaction in other kinases with an aromatic gate keeper residue. We provide a detailed mechanistic characterization of the halogen-aromatic π interactions in the haspin-inhibitor complexes by means of kinetic, thermodynamic, and structural measurements along with binding energy calculations. Since halogens are frequently used in drugs and aromatic residues are often present in the binding sites of proteins, our results provide a compelling rationale for introducing aromatic-halogen interactions to prolong drug-target residence times.

Significance

The time that a drug resides on its target protein has emerged as an important parameter for drug development. A longer drug-target residence time can lead to improved drug efficacy and reduced adverse side effects. However, general strategies for the rational design of drugs with long target residence times are lacking. Here, by experimental and computational characterization of a set of inhibitor-kinase complexes, we show how interactions between inhibitor halogen atoms and protein aromatic residues can increase target residence times. Our results provide a simple strategy for the rational design of inhibitors with prolonged target residence times.

Introduction

The kinetics of drug binding have emerged as important parameters in drug development. A long drug residence time will result in prolonged inhibition after the free drug concentration has dropped due to *in vivo* clearance, potentially leading to improved drug efficiency and reduced off-target mediated toxicity. In cases where slow off-rates are specific for the target, kinetic selectivity can be achieved over fast off-target dissociation despite similar binding constants(1). The kinetics of the interaction of a drug with its target are defined by the association rate (k_{on}) and the dissociation rate (k_{off}) constants. For bimolecular interactions, the ratio of these two parameters defines the equilibrium dissociation constant (K_D) of a drug, and hence the drug occupancy. Since on- and off-rates are coupled in simple rigid bimolecular interactions, the favorable scenario of fast "on" and slow "off"-rate interactions at a given K_D cannot be achieved using this minimalistic binding model. Instead more complex binding models such as induced conformational changes that may trap a ligand in an induced binding pocket are frequently evoked to explain slow binding kinetics.

Kinases are particularly dynamic proteins that provide multiple opportunities for the development of inhibitors that target induced or allosteric binding sites (2-4). One of the first kinase inhibitors for which slow dissociation rates have been described is the p38 MAP (mitogen activating protein) kinase inhibitor BIRB-796. This inhibitor binds to an inactive conformation in which the DFG motif is displaced in a so-called "DFG-out" conformation (5). Inhibitors that bind to this conformation are called type II inhibitors and often have prolonged residence times (τ). However, not all type II inhibitors show slow binding kinetics, suggesting that the DFG-out conformational change *per se* is not sufficient to explain the slow dissociation rates of BIRB-796 from p38 MAP kinase(6). Indeed, more recent studies attributed the slow binding kinetics to efficient hydrophobic contacts in the DFG-out pocket, rather than the kinetic dissociation barrier introduced by the DFG-out transition(7). However, conformational change has also been evoked as the main mechanism contrib-

uting to the slow off-rate of the breast cancer drug lapatinib, a type I inhibitor of the epidermal growth factor receptor (8).

In addition to protein conformational changes, the rearrangement of water molecules has been discussed as a potential mechanism influencing inhibitor residence time (9). An example for the influence of water molecules on ligand binding kinetics is the type I CDK inhibitor roniciclib whose slow off-rate is the result of changes in the hydration network coupled to conformational adaptation of the DFG motif (10). In some cases, the presence of water-shielded hydrogen bonds can also lead to slow dissociation behavior (11). In addition, reversible covalent inhibitors have recently emerged as an interesting strategy for prolonging target engagement of inhibitors by the interaction of a transient covalent bond between an electrophile and a cysteine residue present in the kinase active site (12-15).

Even though protein conformational changes and allosteric pockets can be specifically targeted, they do not offer a straightforward route for ligand design for the modulation of off-rates. Also, the design of reversible covalent interactions requires the presence of cysteine residues in the drug binding site. Many drug receptors, including a large number of kinases, contain cysteines in close proximity to their active sites (16), but the development of covalent inhibitors may not be feasible for all drug targets. In addition, the introduction of a reactive group into an inhibitor poses additional challenges.

Here, we present data that suggest that interactions mediated by halogens, that are common in drugs, and aromatic residues, that are also typically found in drug binding sites on proteins (17), can be utilized to design ligands with slow off-rates. We used 5-iodotubercidin (5-iTU), a close analogue of the kinase cofactor ATP, as a model inhibitor for a canonical active state kinase binding mode (type I) which does not induced any conformational changes that could contribute to slow inhibitor binding kinetics. Screening against more than 100 diverse kinases showed that an aromatic gatekeeper residue that interacts with the halogen moiety of this inhibitor is required for high affinity binding. We chose haspin, a serine/threonine kinase with known three-dimensional structure (18, 19), as a model system. Analysis of ligand binding kinetics surprisingly showed that 5-iTU had slow binding kinetics. Mutation of the gatekeeper residue as well as removal or substitution of the iodide with other halogens showed that the slow inhibitor off-rates are due to a π-stacking interaction of the halogen with the aromatic gatekeeper. Further studies on a different kinase, cdc2 like kinase (CLK1), showed strong inhibition by 5-iTU, support the generality of our findings. Here, we present structural, biophysical and computational data that strongly suggest that halogen interactions with aromatic residues can be exploited for the development of inhibitors with slow off-rates.

Results and Discussion

5-iTU exhibited slow dissociation rate from haspin

Comparative analysis of the high resolution haspin structures revealed high conservation of the binding modes of both 5-iTU and the nucleoside adenosine (Fig. 1A-B). In contrast to adenosine or ATP, which binds with a K_D of about 180 µM(19), 5-iTU showed high affinity for haspin and an unexpectedly long target engagement time. We then further assessed the thermodynamics and kinetics of the binding with ITC (Isothermal Titration Calorimetry), BLI (BioLayer Interferometry) and SPR (Surface Plasmon Resonance) experiments, which consistently confirmed the tight binding with a relatively slow binding kinetics, as measured for instance by BLI (Fig. 1C). Comparing the binding mode of 5-iTU with that of adenosine, the most striking structural difference between these highly similar molecules was the presence of the iodide moiety which was positioned in close proximity to the F605 gatekeeper forming a halogen-aromatic π interaction (Fig. 1D). We therefore hypothesized that this interaction might contribute most of the increase in binding free energy (ΔG) and be responsible for the slow dissociation rate of 5-iTU from haspin.

Preference of 5-iTU for kinases with aromatic gatekeeper

To address our hypothesis, we screened 5-iTU against 137 diverse kinases using temperature shift assays (20) and observed unexpected selectivity of the inhibitor with only 10 kinases exhibiting ΔT_m of >5 °C (Fig. 2A, Table S1). Interestingly, analyses of the gatekeeper residues of kinase targets that showed significant temperature shifts and therefore high affinity for 5-iTU, revealed a strong preference for kinases harboring a phenylalanine (Phe) at this position whereas kinases that showed weak interaction (ΔT_m 2-5 °C) revealed no preference for a certain residue. This analysis supported our hypothesis that the aromatic gatekeeper is important for high affinity binding (Fig. 2B). To confirm these results, we determined the structure of 5-iTU bound to another high affinity target that is structurally very diverse from haspin, CLK1 (ΔT_m of 8.6 °C). As expected, the interaction of 5-iTU with CLK1 remarkably resembled that observed in haspin, including the conserved interaction geometry of the iodide with the CLK1 gatekeeper F241. 5-iTU bound CLK1 with high affinity (K_d of ~7 nM by ITC) and slow off-rates estimated to be ~50 mins by BLI (Fig. 2C-E and Fig. S1).

Importance of halogen σ-hole property for halogen-π interaction for slow off-rate

Formation of a halogen-π interaction (C-X···π) is driven by the directional positive polarization along the halogen σ-bond with the π molecular orbital of the aromatic system (21). However, the main focus of halogen-protein interactions has been on halogen carbonyl/sulphonyl interaction (C-X···O,S bonds) with few examples on the analysis of halogen-

aromatic interactions (22). Because the partial positive charge along the halogen σ-bond diminishes with the size of the halogen, we next substituted the iodide by smaller halogens and characterized the affinities and binding kinetics of these 5-iTU derivatives. Indeed, ITC experiments showed that the affinities of 5-iTU halogen derivatives were reduced with decreasing size of the halogen (Fig. 3 and Table S2).

Removal of the halogen led to a 42-fold decrease in the potency of tubercidin (TU) when compared to that of 5-iTU, and similarly an 8-fold decrease was measured for 5-fluorotubercidin (5-fTU). Analyses of the binding kinetics of these five synthesized 5-tubercidin halogen derivatives with haspin were performed using three independent techniques: kinetic probe competition assays (kPCA) in solution(23), BLI and SPR using immobilized haspin. Binding affinities determined by these three independent methods correlated well with each other and also with the binding constants determined in solution using ITC (Fig. 3B). Dissociation rate constants from all experiments revealed the same behavior with 5-iodide substituted tubercidin displaying the slowest off-rate. The off-rates increased with decreasing halogen size from 5-iodo to 5-fluoro substituted tubercidin and the unsubstituted tubercidin showed the fastest off-rates of binding (Fig. 3C-D). However, the absolute values differed somewhat between the different experimental methods used with the residence time ranging from 60 mins (BLI) to 7 mins (SPR) for 5-iTU (SI Appendix, Table S2 and S3). We also observed slower on rates using BLI but not in the SPR experiments. While the general trends were the same in both technologies, the differences in on- and off-rates that have been observed might be due to differences in protein immobilization. As SPR is the more established technology, we used SPR kinetic data for quantitative analysis. The substitution from hydrogen to iodide at the 5 position of TU led to a 48-fold increase in the on-rates and 274-fold decrease in the off-rates (Fig. 3D and SI Appendix, Table S3).

Thus, the fast dissociation kinetics observed for 5-fTU coincided with the lack of a pronounced σ-hole in smaller halogens, and hence an inability to form a polar halogen-π interaction with the aromatic gatekeeper. The increase in enthalpically favorable polar interactions of larger halogens was also evident in the calorimetric data that showed a steadily decreasing (more negative) binding enthalpy change from tubercidin to the larger halogens (H < F < Cl < Br < I). This effect of the halogen moieties was supported by the crystal structures showing that despite the highly conserved binding mode of all derivatives, the presence of an additional water in the tubercidin complex within the space adjacent to the gatekeeper created by the removal of larger halogen substituents and the longer distance to the fluoro group, presented suboptimal geometry for direct contacts with F605 in tubercidin and 5-fTU, respectively (Fig. 3E and SI Appendix, Fig. S2).

Importance of aromatic phenylalanine for slow off-rate binding

We next investigated the contributions of the Phe gatekeeper to the binding affinity and the observed slow dissociation kinetics of 5-iTU. Six haspin gatekeeper mutants using amino acids commonly found in kinases were generated, and their affinities against 5-iTU and its halogen-derivatives were analyzed using DSF assays (Fig. 4A). As expected, the potency of 5-iTU with the aromatic tyrosine mutant was comparable to wild-type haspin.

For comparison, we analyzed the binding characteristics of 5-iTU with a representative mutant (F605T) in detail. Structural superimposition between the wild-type and the mutant structures (F605Y and F605T) showed that the gatekeeper mutation did not affect the binding mode of 5-iTU, yet led to slight variation of the environment in the pocket. The substitution of the bulky aromatic residue F605 with the small threonine resulted in extension of a water network that filled the expanded binding site in the mutant (Fig. 4B and SI Appendix, Fig. S3). No direct contact was observed between the threonine side chain and the iodide of 5-iTU, although interactions might be mediated through a water bridge. The absence of any strong contact was in agreement with the fast kinetics with ~16-fold lower affinity as demonstrated by SPR (Fig. 4C-D). In comparison to the wild-type, the loss of the halogen-π contact in the F605T mutant led to a 6-fold decrease and 8-fold increase in association and dissociation rates, respectively, with the estimated residence time of 5-iTU dramatically dropping to less than a minute (Fig. 4E and SI Appendix, Table S3).

Thermodynamics of halogen-π interactions in slow kinetic behavior

In order to assess the energetic contributions of the halogen-aromatic gatekeeper interaction, we calculated the energy of this interaction using *ab initio* quantum mechanics and classical methods. The second order Møller–Plesset interaction energies (E_{MP2}) between the inhibitor and the gatekeeper residue correlate well with dissociation rate constants and equilibrium dissociation constants determined experimentally (Fig. 5A and SI Appendix, Fig. S5 and Tables S6-S9). Partitioning of E_{MP2} into its constituent energy components using a many-body interaction energy decomposition scheme shows that the major contribution to E_{MP2} is the correlation energy (E_{CORR}) which describes second-order intermolecular dispersion interactions and the correlation corrections to the Hartree-Fock energy. E_{CORR} increases in magnitude with increasing size of the halogen, corresponding to the decreasing rate of dissociation measured experimentally (Fig. 5B), and indicating the importance of the halogen interaction with the aromatic gatekeeper for the prolongation of residence times as halogen size increases. The computed *ab initio* energies also correlate for the interaction of 5-iTU with the F605Y mutant but the magnitude of the interaction energy of

5-iTU with the threonine mutant was underestimated. This discrepancy could be explained by the additional bound water molecule present between 5-iTU and the threonine residue in the F605T structure which was not included in the calculation but would make an energetically favorable contribution to the binding. To account for the complete protein structure in the computation of the binding free energies of the haspin-ligand complexes, we used the classical MMGBSA approach with an implicit solvent model. The computed energies correlate well with calorimetric data measured by ITC, consistent with the increasingly favorable enthalpic contribution to binding as halogen size increases (Fig. 5C-D and SI Appendix, Table S10).

Halogens are frequently found in approved drugs. Interestingly, drug candidates with heavy halogens (Cl, Br, I) perform better in the clinical development pipeline increasing steadily from phase II (44%) to approved drugs (63%) whereas fluorine-containing drug candidates decrease from 56% in phase II to 36.6% in launched drugs (17). Unfortunately, information on binding kinetics is not available for these molecules but it is tempting to speculate that the presence of halogens may lead to longer lasting drug-protein interactions. A recent survey of the PDB for halogen-protein interactions reported that 33% of all non-bonded interactions (excluding C-X···H contacts) of the heavier halogens in the protein database form aromatic stacking interactions of the C-X···π type (24). A large number of halogens also interact with backbone carbonyls and thiols and it would be interesting to investigate if these interactions also result in slower ligand binding kinetics. For the selected case of protein kinases, aromatic gatekeeper residues are the most frequently found residue type. Our proposed strategy of incorporating heavier halogens into inhibitors with the goal of increasing target residence time by designing interactions with aromatic residues is a general approach that may lead to active compounds with improved pharmacological properties. The biophysical, structural and computational data presented here on 5-halogen substituted tubercidin derivatives provides a good basis for future studies on this exciting topic.

ACKNOWLEDGMENT

Author Contributions

C.H., V.G, G.K.G., F.W. performed experiments, R.C.W., A.E.F-M., P.B., A.C., S.K supervised the research, S.K., A.C. drafted the manuscript with contributions from all authors. All authors approved the final version.

Notes

The authors declare no competing financial interests.

REFERENCES

1.   Copeland RA, Pompliano DL, & Meek TD (2006) Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* 5(9):730-739.
2.   Muller S, Chaikuad A, Gray NS, & Knapp S (2015) The ins and outs of selective kinase inhibitor development. *Nat Chem Biol* 11(11):818-821.
3.   Chaikuad A*, et al.* (2014) A unique inhibitor binding site in ERK1/2 is associated with slow binding kinetics. *Nat Chem Biol* 10(10):853-860.
4.   Zhao Z*, et al.* (2014) Exploration of type II binding mode: A privileged approach for kinase inhibitor focused drug discovery? *ACS Chem Biol* 9(6):1230-1241.
5.   Pargellis C*, et al.* (2002) Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat Struct Biol* 9(4):268-272.
6.   Willemsen-Seegers N*, et al.* (2017) Compound Selectivity and Target Residence Time of Kinase Inhibitors Studied with Surface Plasmon Resonance. *J Mol Biol* 429(4):574-586.
7.   Schneider EV, Bottcher J, Huber R, Maskos K, & Neumann L (2013) Structure-kinetic relationship study of CDK8/CycC specific compounds. *Proc Natl Acad Sci U S A* 110(20):8081-8086.
8.   Wood ER*, et al.* (2004) A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res* 64(18):6652-6659.
9.   Englert L*, et al.* (2010) Displacement of disordered water molecules from hydrophobic pocket creates enthalpic signature: binding of phosphonamidate to the S(1)'-pocket of thermolysin. *Biochim Biophys Acta* 1800(11):1192-1202.
10.   Ayaz P*, et al.* (2016) Conformational Adaption May Explain the Slow Dissociation Kinetics of Roniciclib (BAY 1000394), a Type I CDK Inhibitor with Kinetic Selectivity for CDK2 and CDK9. *ACS Chem Biol* 11(6):1710-1719.
11.   Schmidtke P, Luque FJ, Murray JB, & Barril X (2011) Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J Am Chem Soc* 133(46):18903-18910.
12.   Bradshaw JM*, et al.* (2015) Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat Chem Biol* 11(7):525-531.

13.   Miller RM, Paavilainen VO, Krishnan S, Serafimova IM, & Taunton J (2013) Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J Am Chem Soc* 135(14):5298-5301.

14.   Forster M*, et al.* (2016) Selective JAK3 Inhibitors with a Covalent Reversible Binding Mode Targeting a New Induced Fit Binding Pocket. *Cell Chem Biol* 23(11):1335-1340.

15.   Chaikuad A, Koch P, Laufer S, & Knapp S (2017) Targeting the Protein Kinases Cysteinome. *Angew Chem Int Ed Engl*.

16.   Liu Q*, et al.* (2013) Developing irreversible inhibitors of the protein kinase cysteinome. *Chem Biol* 20(2):146-159.

17.   Xu Z*, et al.* (2014) Halogen bond: its role beyond drug-target binding affinity for drug discovery and development. *J Chem Inf Model* 54(1):69-78.

18.   Eswaran J*, et al.* (2009) Structure and functional characterization of the atypical human kinase haspin. *Proc Natl Acad Sci U S A* 106(48):20198-20203.

19.   Villa F*, et al.* (2009) Crystal structure of the catalytic domain of Haspin, an atypical kinase implicated in chromatin organization. *Proc Natl Acad Sci U S A* 106(48):20204-20209.

20.   Fedorov O, Niesen FH, & Knapp S (2012) Kinase inhibitor selectivity profiling using differential scanning fluorimetry. *Methods Mol Biol* 795:109-118.

21.   Ho PS (2015) Biomolecular halogen bonds. *Top Curr Chem* 358:241-276.

22.   Tatko CD & Waters ML (2004) Effect of halogenation on edge-face aromatic interactions in a beta-hairpin peptide: enhanced affinity with iodo-substituents. *Org Lett* 6(22):3969-3972.

23.   Schiele F, Ayaz P, & Fernandez-Montalvan A (2015) A universal homogeneous assay for high-throughput determination of binding kinetics. *Anal Biochem* 468:42-49.

24.   Lu Y, Wang Y, & Zhu W (2010) Nonbonding interactions of organic halogens in biological systems: implications for drug discovery and biomolecular design. *Phys Chem Chem Phys* 12(18):4543-4551.

25.   Bullock AN*, et al.* (2009) Kinase domain insertions define distinct roles of CLK kinases in SR protein phosphorylation. *Structure* 17(3):352-362.

Figure legends

FIGURE 1. The 5-iodotubercidin inhibitor (5-iTU) exhibits tight binding with slow dissociation kinetics from haspin. A) Chemical structures of 5-iTU and adenosine. B) Superimposition of haspin-5iTU and AMP (pdb id: 4ouc) reveals similar binding modes of the two compounds. C) BLI sensorgram suggests slow kinetic behavior of the 5-iTU-haspin interaction. D) The iodide and the benzene moieties of 5-iTU and F605, respectively, are located in close proximity with a favorable geometry for a halogen-π bond.

FIGURE 2. Interaction of 5-iTU with various human kinases. A) Temperature shift assays show high ΔT_m preferentially for kinases harboring an aromatic gatekeeper. B) Distribution of gatekeeper residues reveals that the majority of strong ΔT_m hits harbor a phenylalanine gatekeeper. C) The binding mode of 5-iTU in the off-target CLK1, including the iodide gatekeeper interaction, is highly conserved. CLK1 has a high affinity for 5-iTU as measured by ITC (D) and a slow off-rate (E) as assessed by BLI.

FIGURE 3 Binding kinetics of haspin with five tubercidin derivatives harboring halogen substituents at the 5-position. A) ITC thermodynamic binding parameters. B) Comparison of dissociation constants (K_D) measured by ITC, BLI, SPR and kPCA shows good correlation of the measured equilibrium data. C) SPR sensorgrams demonstrating increasingly slow dissociation rates with increasing size of the halogens. D) Rate plot with Isoaffinity Diagonal (RaPID) of k_{on} and k_{off} constants measured by BLI, SPR and kPCA. The red arrow indicates the trend to increasing k_{on} and decreasing k_{off} upon increasing the atomic radii of the halogens. E) Crystal structures reveal conserved binding modes of all five tubercidin derivatives, albeit with an additional water molecule adjacent to the inhibitor and F605 gatekeeper in tubercidin.

FIGURE 4 Effect of gatekeeper mutation on the binding kinetics of haspin with tubercidin derivatives. A) Tm shifts of six haspin mutants against five tubercidin derivatives. B) Superimposition of 5-iTU-complexed crystal structures of wild-type, F605Y and F605T haspin reveals conserved binding mode of the inhibitor, yet differences in bound water molecules within the binding site. C) SPR sensorgram demonstrates fast binding kinetics for the interaction between 5-iTU and the F605T mutant, and this is accompanied by a significant decrease in K_D (D) and decreased k_{on} and generally increased k_{off} constants (E).

FIGURE 5 Correlation of calculated binding free energies with experimental parameters for the halogen-gatekeeper interaction. A) Second-order Møller–Plesset interaction energy

(E_{MP2}) and B) second-order correlation correction energy term (E_{CORR}) between the TU derivatives and the gatekeeper residues *vs* BLI k_{off} values. The linear fits and correlation coefficients (R^2) were computed omitting the outlier F605T mutant. The experimental error bars are smaller than the size of the data plots. Comparison of (C) MMGBSA internal and solvation contributions ($\Delta E_{gas}+\Delta\Delta G_{solvation}$) *vs* the ITC enthalpies (ΔH_{ITC}) and (D) MMGBSA binding free energies (ΔG_{MMGBSA}) *vs* the ITC binding free energies (ΔG_{ITC}) of the interactions between haspin and TU derivatives. Some ΔG_{MMGBSA} values are positive as they only include translational and rotational entropic terms and do not include vibrational and conformational entropy contributions.

Figure 1
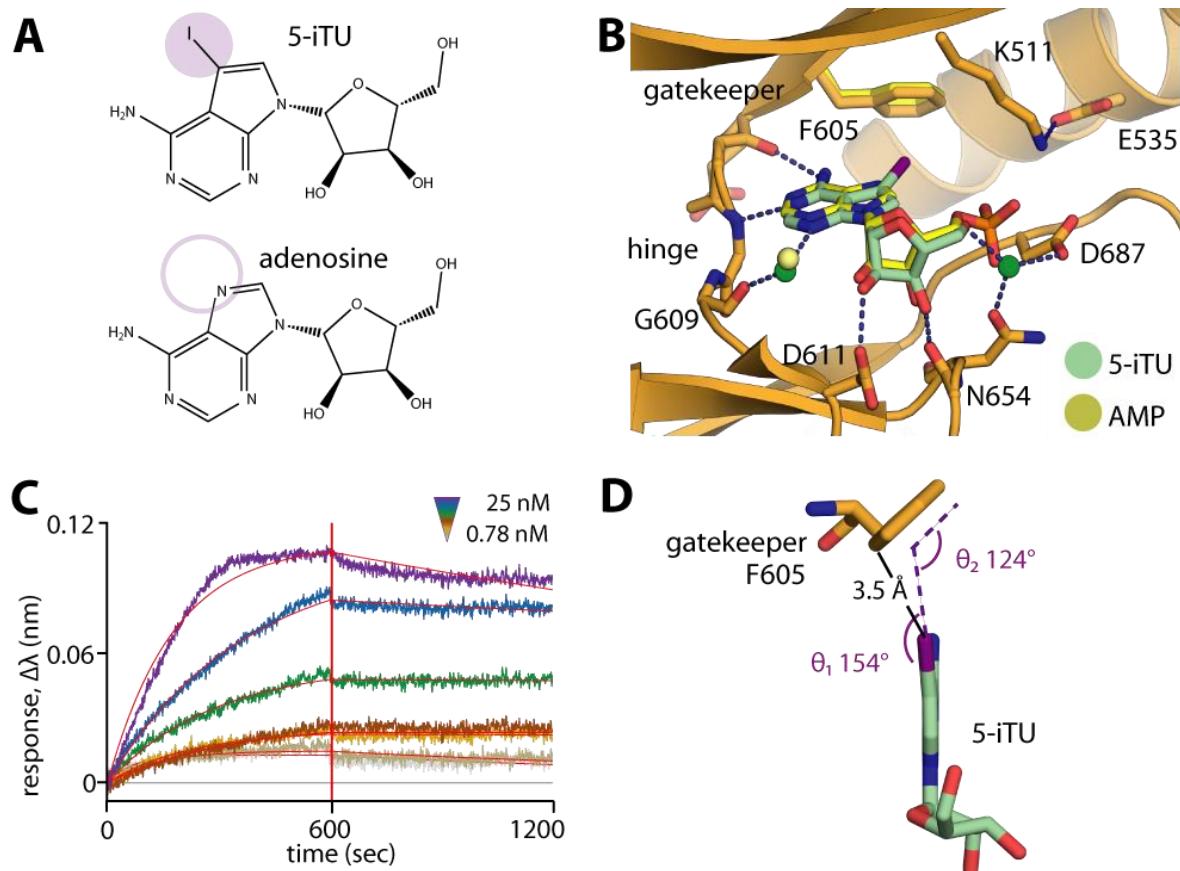
Figure 2
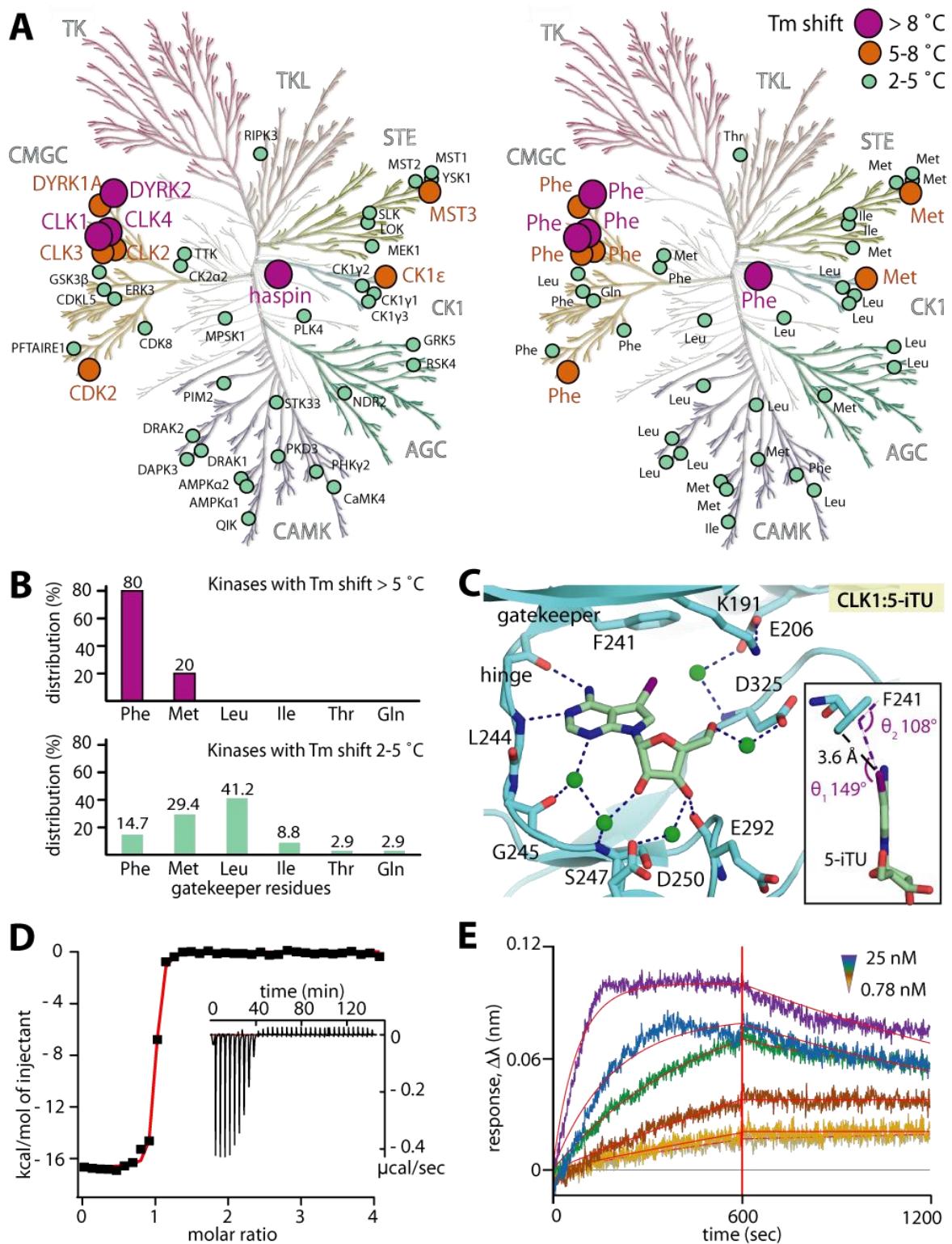
Figure 3
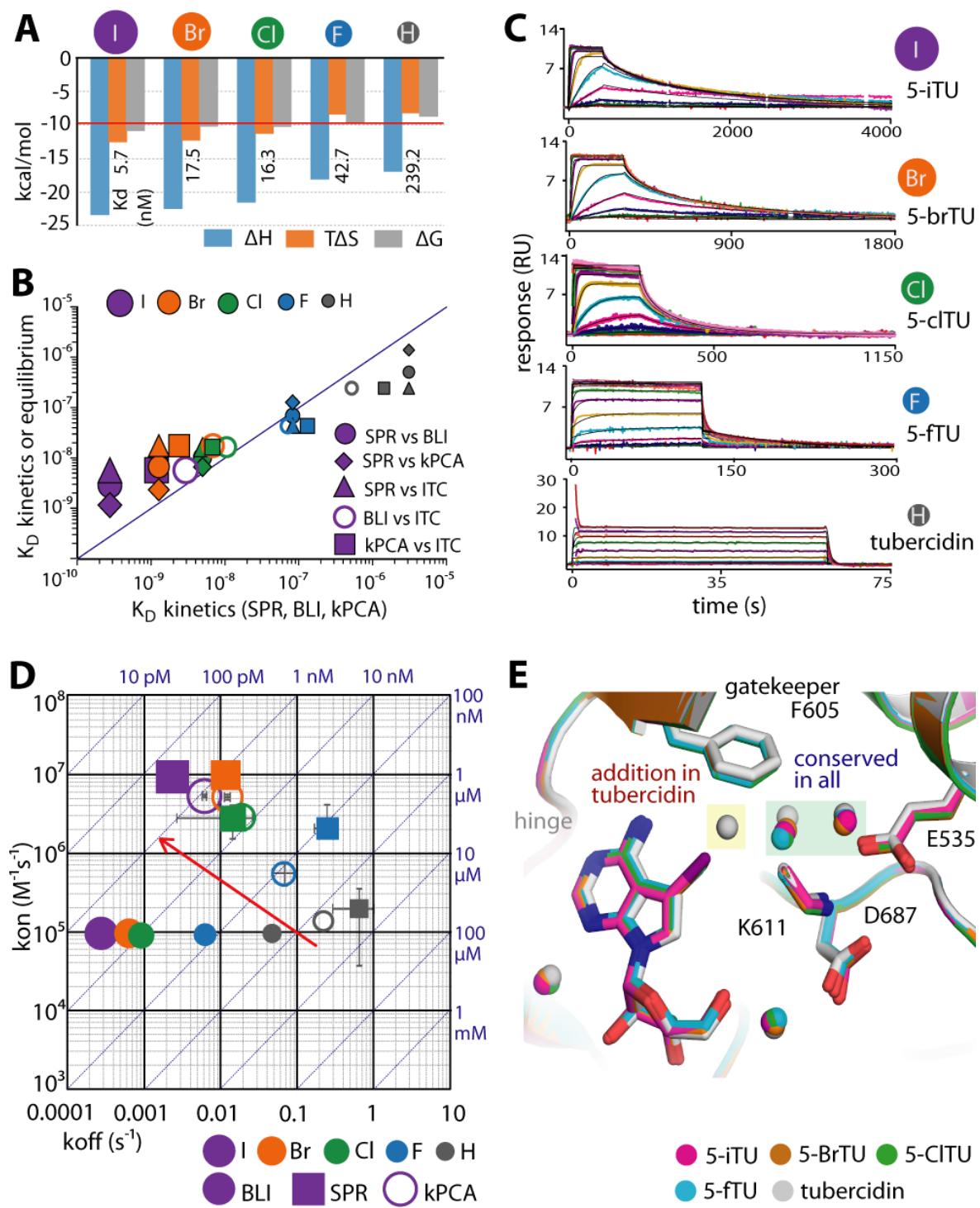
Figure 4
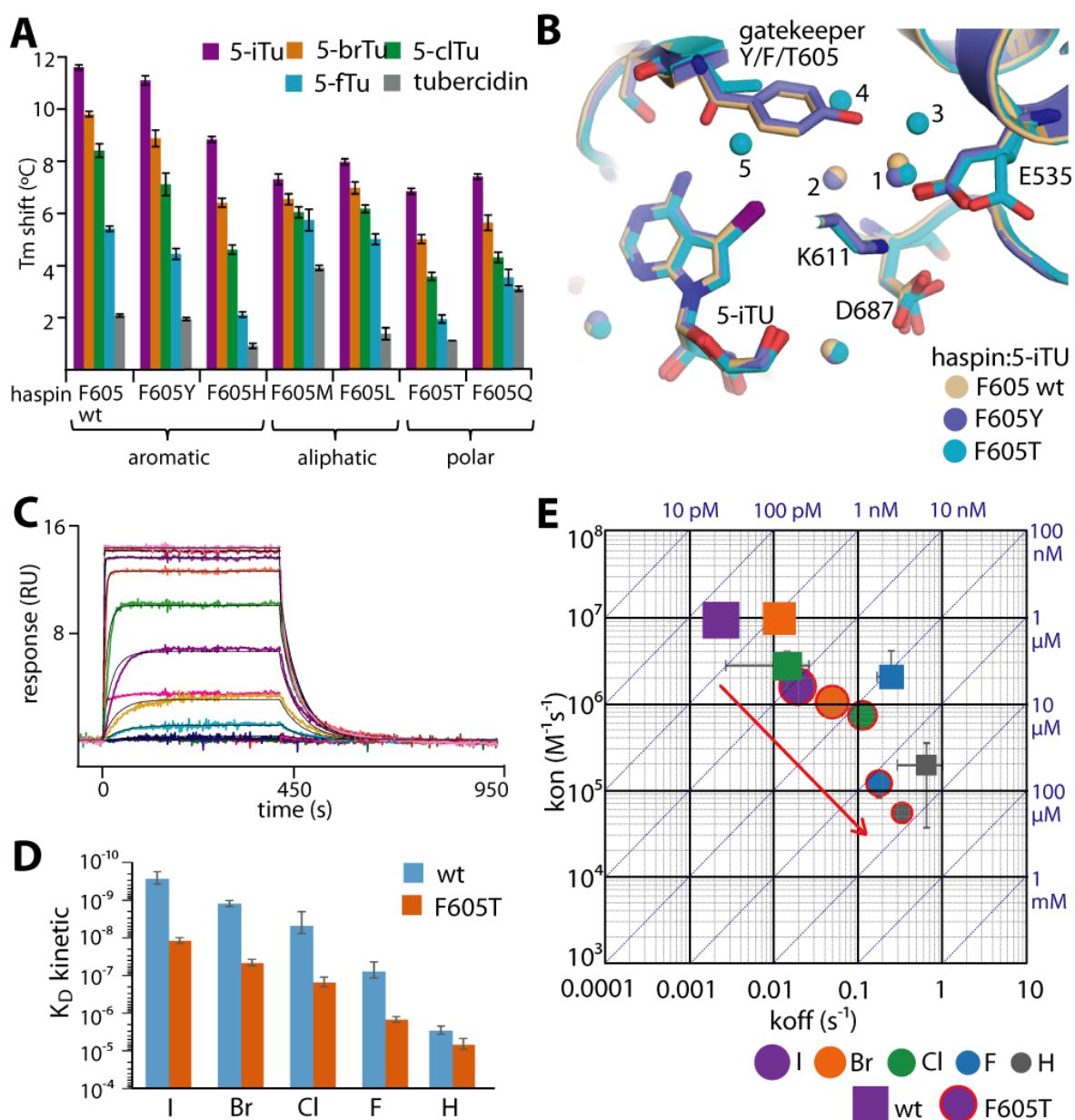
Figure 5
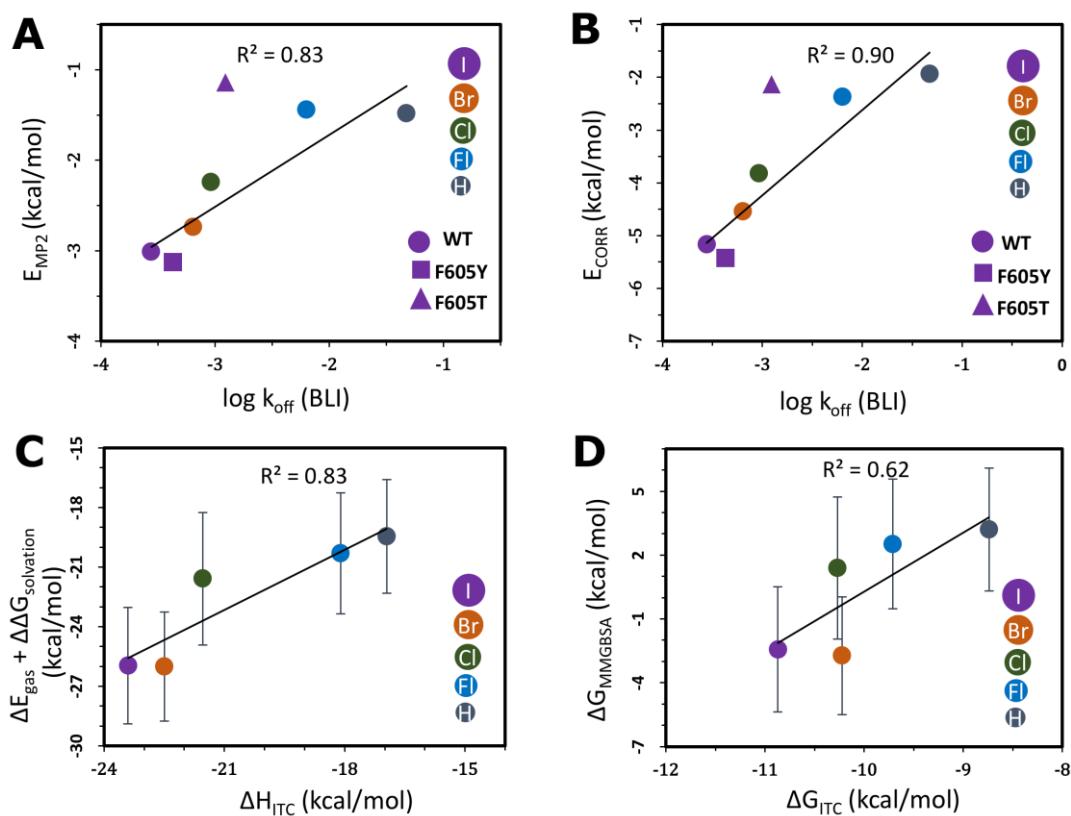
Supplemental Material

Halogen-aromatic π interactions modulate inhibitor residence time

Christina Heroven, Victoria Georgi, Gaurav K. Ganotra, Paul Brennan, Fynn Wolfreys, Rebecca C. Wade, Amaury, E. Fernández-Montalván, Apirat Chaikuad, Stefan Knapp

Content:

Material and Methods

Supplemental Figure 1 (Binding mode 5ITu CLK1)

Supplemental Figure 2 (Binding modes Tu derivatives)

Supplemental Figure 3 (Structure of haspin mutants)

Supplemental Figure 4 (Details of gatekeeper interactions)

Supplemental Figure 5 (Correlation plots of QM data & k_{off}/K_D values)

Supplemental Table 1 (ΔT_m data)

Supplemental Table 2 (BLI and ITC data)

Supplemental Table 3 (SPR data)

Supplemental Table 4 (Solvent accessible surface area calculations)

Supplemental Table 5 (Geometric parameters of gatekeeper X interaction)

Supplemental Table 6 - 9 (Total interaction energy calculations)

Supplemental Table 10 (MMGBSA Binding free energy calculations)

Supplemental Table 11 (Data collection and crystallographic refinement)

Material and Methods

Protein purification

Haspin and CLK1, both wild-type and all mutants, have been purified as described (18, 25). Briefly, the recombinant proteins were purified using Co^{2+}/Ni^{2+} affinity chromatography. For haspin, the His-tagged protein were subsequently purified by size-exclusion chromatography and the final protein was stored in 50 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP. For CLK1, the histidine tag was removed by incubating the protein with TEV protease overnight. The cleaved CLK1 protein was separated by reverse purification on Ni^{2+} affinity chromatography, and subsequently size exclusion chromatography. The final CLK1 protein was stored in 30 mM HEPES, pH 7.5, 300 mM NaCl, 50 mM L-Arginine/L-Glutamate mix, 10 mM DTT, 1% glycerol.

Protein crystallization

All crystallization experiments were performed using sitting-drop vapour-diffusion method at 4 °C. For Haspin, the protein at ~12 mg/ml was incubated with 1 mM inhibitors, and the complexed crystals were obtained using the crystallization condition containing 51-63% MPD and 0.1M SPG buffer, pH 6.0-6.5. To obtain the inhibitor-CLK1 complex, apo crystals grew in 20% 1,2-propanediol, 5% glycerol and 0.1 M $NaKPO_4$ were soaked with inhibitor overnight.

Data collection and structure determination

All diffraction data were collected at Diamond Light Source, and processed using MOSFLM(26). Scaling was performed using aimless from the CCP4 suite(27). All structures were solved by molecular replacement using Phaser (28) and the deposited CLK1 and haspin structures (PDB entries XYZ and 4OUC respectively) as models. All structures were subjected to one round of automated model building using *ARP/wARP(29)*, followed by iterative cycles of manual model building in COOT (30), alternated with refinement using REFMAC (31). TLS definitions used in the final refining rounds were calculated using the TLSMD server (31). The model quality and geometric correctness of all complexes was verified using MolProbity (32). Statistics for data collection and structure refinement are summarized in Supplemental Table 11.

Thermal shift assays

Protein were diluted to 2 µM in a buffer containing 10 mM HEPES, pH 7.5 and 500 mM NaCl, and mixed with SYPRO Orange at 1000-fold dilution of the dye. The inhibitors were added at 10 µM final concentration. The DSF assay was performed using a Real-Time PCR Mx3005p machine (Stratagene) according to the protocol described previously (33).

Isothermal Titration Calorimetry

All proteins were exchanged into a suitable storage buffer. CLK1 at 100 µM was stored in 20 mM HEPES, pH 7.5, 300 mM NaCl, 50 mM L-arginine/L-glutamate mix and 0.5 mM TCEP. For haspin at 80 µM, the buffer containing 20 mM HEPES, pH 7.5, 250 mM NaCl and 0.5 mM TCEP was used for the wild-type protein, while the gatekeeper mutants were buffer exchanged into 30 mM HEPES, pH 7.5, 400 mM NaCl and 0.5 mM TCEP to increase their stabilities. Calorimetric measurements were carried out using a VP-ITC calorimeter (MicroCal) at 15 °C. For all experiments, the proteins were titrated into the reaction cell containing the compound. Integrated heat of titrations were manually corrected and analysed in Origin. Using a single binding site model, the obtained curve was fitted following a nonlinear least-square minimization algorithm. The binding isotherms and the

measured binding enthalpy changes enabled the calculation of entropy changes (TΔS), Gibbs free energy (ΔG), the stoichiometry *n* and K_d.

Biolayer Interference

The binding kinetics were measured by Biolayer Interference method (BLI) using Octet RED384 system (*forté*BIO). For haspin, the experiments were performed in the buffer condition containing 20 mM HEPES, pH 7.5, 400 mM NaCl and 0.5 mM TCEP. For CLK1, the same buffer supplemented with 50 mM L-arginine/L-glutamate mix was used. Biotinylated proteins, prepared as previously described, were immobilized on streptavidin biosensors, which were subsequently quenched with L-biotin(34). The interference patterns of association and dissociation events were measured through a time course of 600 seconds. The binding data were corrected using a double-referencing method, and the kinetics analyses were performed according to the manufacture protocol (*forté*BIO).

Surface Plasmon Resonance (SPR)

SPR experiments were performed in HBS-PE+ buffer on a Biacore T200 System (GE Healthcare). Biotinlyated wt and mutant Haspin (50 μg/ml in HBS-P+ buffer) were captured to a SA sensor chip (GE Healthcare) at typical densities of 2-5 kRU using the protocols provided by the manufacturer. Compounds were serially diluted in DMSO and transferred to assay buffer in a 1:100 dilution step to achieve their final test concentrations at a [DMSO] = 1%. For binding analysis, contact times of 60, 120, 240, 300 or 420 seconds were used depending on the kinetics assessed in preliminary tests. Likewise, dissociation times were adjusted to180, 900, 1500 or 3600 seconds, to achieve return of the SPR signals to baseline levels. To obtain kinetic and affinity parameters, sensorgrams (acquired at 10 Hz) were fitted using the BIAevaluation Software (GE Healthcare) to a 1:1 Langmuir model accounting for mass transport limitations. Steady state analysis was performed with the same software using a single site equilibrium binding equation.

Equilibrium und Kinetic probe competition assays (ePCA and kPCA)

ePCA and kPCA experiments were performed in Tris-HCl pH7.5, 150 mM NaCl, 0.01% Tween, 0.01% BSA, 2 mM DTT buffer as previously described for CDK2 in Schiele et al (23). Biotinylated wt and mutant Haspin (4 nM in assay) were labelled at a molar ratio of 8:1 with SA-Terbium (Cisbio) as TR-FRET donor. Tracers 236 and 199 (Invitrogen) labelled with an Alexa 647 TR-FRET acceptor were respectively used as kinase specific probes at a final concentration of 100 nM.

Compounds were diluted and transferred to Greiner black small volume 384-well microtiter test plates as described(23). For ePCA, tracer and labeled proteins were dispensed to the ready-to-use compound plates to a final volume of 5 μL and the mixture was incubated for 2 h prior to acquisition of the steady state TR-FRET ratiometric signals (665/620 nm) upon excitation at 337 nm. Normalized values were fitted to a logistic 4-parameter model using the Genedata Screener^{TM} software, and Ki values calculated using the Cheng-Prusoff relationship. For kPCA, the tracer was dispensed to the ready-to-use compound plates prior to introducing them into the PHERAstar FS^{TM} microtiter plate reader. Then the labeled proteins were added to wells to a final volume of 10 μL using the injector system of the instrument, and kinetic TR-FRET readings were made at time zero and every 10 seconds. Blank-subtracted kinetic traces were analyzed with a competitive binding kinetics model using the GraphPad Prism^{TM} software as described (23).

Prior to compound testing, the steady state affinities of the probes were determined by equilibrium binding titrations (0 to 400 nM) on various Haspin concentrations (0 to 8 nM)

with end-point readings of the TR-FRET signals. The probes' association and dissociation kinetics were characterized by titrating them on 4 nM labeled Haspin (0.5 nM SA-Tb) and acquiring the TR-FRET signals in real time. Binding curves were fitted to the corresponding models with Graph Pad Prism^{TM} in order to obtain the affinity and kinetic constants used as parameters in the Cheng-Prusoff and Motulsky and Mahan models.

Quantum mechanical interaction energy calculations

The energy contributions of the inhibitor-aromatic gatekeeper interaction were calculated using *ab initio* Møller–Plesset perturbation theory to second order (MP2). The Moeller-Plesset perturbation theory improves on the Hartree-Fock method by adding electron-correlation effects by means of Rayleigh-Schrödinger perturbation theory to different orders (second order in our case). The *Protein Preparation* wizard of the *Maestro* program of the *Schrodinger suite* (*Version 2015.r3*) was used to pre-process the X-ray crystallographic structures of the haspin-inhibitor complexes, to add missing side chains and to optimize the H-bond network. The *impref* utility of the *Maestro* was used for energy minimization using the *OPLS3* force field. The *impref* utility(35) first optimizes position of hydrogen atoms followed by all-atom minimization where non-hydrogen atoms are restrained with a harmonic potential using a force constant of 25 kcal/mol.Å^2. The coordinates of the inhibitor and the gatekeeper phenylalanine residue were extracted from these energy-minimized structures of haspin-inhibitor complexes. The termini of the phenylalanine residue were blocked with hydrogen atoms and their positions were optimized using the *OPLS3* force field in the *Maestro* program of the *Schrödinger suite (36)*. In the case of the gatekeeper mutants, the corresponding gatekeeper residues (tyrosine and threonine) were prepared in the same way.

The *def2TZVP* basis set was used for all calculations and effective core potentials (ECPs) were used for the iodine atom. *Ab initi*o interaction energies at the MP2 level were calculated using the *GAMESS* software, and partitioned into their constituent interaction energy terms (see Equation 1) using the many body interaction energy decomposition scheme (EDS) described by *Góra et al.*(37). In this scheme, the total interaction energy is calculated in a super-molecular approach as the difference between the total energy of a complex (here, of the inhibitor and the gatekeeper residue) and the sum of the energies of its isolated constituents. In all calculations, the complex centered basis set (CCBS) was used consistently and the results are therefore basis set superposition error (BSSE) free due to the full counterpoise correction.

$$E_{MP2} = E_{EL,MTP} + E_{EL,PEN} + E_{EX} + E_{DEL} + E_{CORR}$$

$$E_{SCF} = E_{EL,MTP} + E_{EL,PEN} + E_{EX} + E_{DEL}$$

Equation (1)

$$E_{HL} = E_{EL,MTP} + E_{EL,PEN} + E_{EX}$$

$$E_{EL} = E_{EL,MTP} + E_{EL,PEN}$$

As shown in Equation 1, the total interaction energy at the MP2 level of theory (E_{MP2}) includes the components of the Hartree-Fock interaction energy (E_{SCF}) and the second order Coulomb correlation correction term (E_{CORR}). This correlation term (E_{CORR}) includes the

second order intermolecular dispersion energy and the correlation corrections to the SCF components. The Hartree-Fock interaction energy (E_{SCF}) was partitioned into a first order Heitler-London component (E_{HL}) and a higher order Hartree-Fock delocalization interaction energy component (E_{DEL}), which encompasses the induction and the associated exchange effects. Because their separation could lead to a non-physical charge transfer, this component was not partitioned any further. The Heitler-London interaction energy component (E_{HL}) can be separated into the first-order electrostatic interactions (E_{EL}) of monomers (the inhibitor and the gatekeeper residue in our case) and the associated Heitler-London exchange repulsion energy (E_{EX}) due to the Fermi electron correlation effects. The electrostatic interaction energy (E_{EL}) was obtained as a first-order term in the polarization perturbation theory and the exchange repulsion term (E_{EX}) was calculated by subtracting the electrostatic interaction energy from the Heitler-London energy ($E_{EX} = E_{HL} - E_{EL}$). $E_{EL,MTP}$ refers to the electrostatic multipole component estimated from an atomic multipole expansion, $E_{EL,PEN}$ is the electrostatic penetration energy, calculated from the following expression: $E_{EL,PEN} = E_{EL} - E_{EL,MTP}$.

Binding free energy calculations

The molecular mechanics-generalized Born surface area (MM/GBSA) method was used to estimate the binding free energy of the inhibitors to haspin kinase. The initial coordinates of the haspin-inhibitor complexes were obtained from the co-crystallized structures (see Supplementary Figure 2). The *Protein Preparation* wizard of the *Schrodinger suite* (*Version 2015.r3*) was used for pre-processing of the structures, formation of disulfide bonds, addition of hydrogen atoms and assigning protonation states at pH 7.0. The *pmemd* module of the *Amber14* software suite (38) was used to perform the molecular dynamics (MD) simulations with the *ff14SB*(39) force field for protein. The *LEap* module of *AmberTools14* was used to construct the topologies of the haspin-inhibitor complexes. The ligand parameters were generated based on the generalized Amber force field (*GAFF*). To improve the description of charge, dipole moment and geometry of halogenated compounds in molecular mechanics calculations, the positive region (σ hole) centered on the halogen atom was represented by an extra-point charge (EP). This inclusion of an EP results in improved modeling of halogen-bonding in MD simulations. The force field parameters for this EP were taken from *Ibrahim et al.(40)*. For generation of the partial atomic charges for the ligands, the *RESP(41)* program was used to fit the atom-centered charges to the molecular electrostatic potential (MEP) grid computed by the *GAMESS* program. The system was centred and aligned with the axes to minimize the volume. The system was then solvated using the *TIP3P* water model(42) by immersing the protein-ligand complex in a cubic box of water molecules, such that the shortest distance between the edge of the solvation box and the complex is 10 Å. The net charge (-2e) of the system was then neutralized by adding Na^+ counter ions. For each system, energy minimization was performed in three 1500-cycle consecutive runs using the steepest descent minimization method followed by switching to the conjugate gradient method after 500 cycles. Gradually decreasing harmonic restraints with force constants of 500, 1 and 0 kcal/mol.Å^2 were used for non-hydrogen atoms in three consecutive runs. Energy minimization was followed by 1 nanosecond (*ns*)

of gradual heating from 10 K to 300 K with harmonic restraints with a force constant of 50 kcal/mol.Å^2 acting on non-hydrogen atoms. Then the system was equilibrated for 1 *ns* under *NPT* conditions at 300K, with heavy atoms (except solvent ions) harmonically restrained with a force constant of 50 kcal/mol.Å^2. This was followed by an *NPT* equilibration of 2 *ns* without any positional restraints. The potential energy function and atomic coordinates were calculated using a 2 femtoseconds (*fs*) time step. The *SHAKE*(43) algorithm was used to constrain all the bonds involving hydrogen atoms. The *Particle Mesh Ewald* (PME) method was used to calculate the electrostatic interactions. A cut-off of 10 Å was set for generating the non-bonded pair list and this pair list was updated every 100 steps. After equilibration, data were collected over a 6 *ns* simulation run for binding free energy calculations and 3000 sets of atomic coordinates were saved every 2 picoseconds (*ps*). MM/GBSA calculations of the binding free energy were performed using the *MMPBSA.py* module implemented in the *Amber14* analysis tools. A single-trajectory approach was used in which receptor, ligand and complex geometries were extracted from a single MD trajectory. All the ions and water molecules were stripped from the trajectory snapshots. A salt concentration of *0.15 M* and the Born implicit solvent model (*igb = 2*) were used. Each binding free energy was computed as the sum of a molecular mechanics term ($\triangle E_{gas}$), a Gibbs solvation term ($\triangle\triangle G_{solvation}$) and an entropic contribution ($T\triangle S_{solute}$). For the entropic contribution to binding free energy, we computed translational and rotational entropies with a rigid rotor model using the *MMPBSA.py* module. The calculation of vibrational entropies using normal-mode analysis with MMPBSA.py failed due to the inclusion of the EP in the force field. The free energy of binding for some of the derivatives is positive since vibrational and conformational entropy terms are neglected.

References

1. Copeland RA, Pompliano DL, & Meek TD (2006) Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* 5(9):730-739.
2. Muller S, Chaikuad A, Gray NS, & Knapp S (2015) The ins and outs of selective kinase inhibitor development. *Nat Chem Biol* 11(11):818-821.
3. Chaikuad A*, et al.* (2014) A unique inhibitor binding site in ERK1/2 is associated with slow binding kinetics. *Nat Chem Biol* 10(10):853-860.
4. Zhao Z*, et al.* (2014) Exploration of type II binding mode: A privileged approach for kinase inhibitor focused drug discovery? *ACS Chem Biol* 9(6):1230-1241.
5. Pargellis C*, et al.* (2002) Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat Struct Biol* 9(4):268-272.
6. Willemsen-Seegers N*, et al.* (2017) Compound Selectivity and Target Residence Time of Kinase Inhibitors Studied with Surface Plasmon Resonance. *J Mol Biol* 429(4):574-586.
7. Schneider EV, Bottcher J, Huber R, Maskos K, & Neumann L (2013) Structure-kinetic relationship study of CDK8/CycC specific compounds. *Proc Natl Acad Sci U S A* 110(20):8081-8086.
8. Wood ER*, et al.* (2004) A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res* 64(18):6652-6659.
9. Englert L*, et al.* (2010) Displacement of disordered water molecules from hydrophobic pocket creates enthalpic signature: binding of phosphonamidate to the S(1)'-pocket of thermolysin. *Biochim Biophys Acta* 1800(11):1192-1202.

10.   Ayaz P*, et al.* (2016) Conformational Adaption May Explain the Slow Dissociation Kinetics of Roniciclib (BAY 1000394), a Type I CDK Inhibitor with Kinetic Selectivity for CDK2 and CDK9. *ACS Chem Biol* 11(6):1710-1719.

11.   Schmidtke P, Luque FJ, Murray JB, & Barril X (2011) Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J Am Chem Soc* 133(46):18903-18910.

12.   Bradshaw JM*, et al.* (2015) Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat Chem Biol* 11(7):525-531.

13.   Miller RM, Paavilainen VO, Krishnan S, Serafimova IM, & Taunton J (2013) Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J Am Chem Soc* 135(14):5298-5301.

14.   Forster M*, et al.* (2016) Selective JAK3 Inhibitors with a Covalent Reversible Binding Mode Targeting a New Induced Fit Binding Pocket. *Cell Chem Biol* 23(11):1335-1340.

15.   Chaikuad A, Koch P, Laufer S, & Knapp S (2017) Targeting the Protein Kinases Cysteinome. *Angew Chem Int Ed Engl*.

16.   Liu Q*, et al.* (2013) Developing irreversible inhibitors of the protein kinase cysteinome. *Chem Biol* 20(2):146-159.

17.   Xu Z*, et al.* (2014) Halogen bond: its role beyond drug-target binding affinity for drug discovery and development. *J Chem Inf Model* 54(1):69-78.

18.   Eswaran J*, et al.* (2009) Structure and functional characterization of the atypical human kinase haspin. *Proc Natl Acad Sci U S A* 106(48):20198-20203.

19.   Villa F*, et al.* (2009) Crystal structure of the catalytic domain of Haspin, an atypical kinase implicated in chromatin organization. *Proc Natl Acad Sci U S A* 106(48):20204-20209.

20.   Fedorov O, Niesen FH, & Knapp S (2012) Kinase inhibitor selectivity profiling using differential scanning fluorimetry. *Methods Mol Biol* 795:109-118.

21.   Ho PS (2015) Biomolecular halogen bonds. *Top Curr Chem* 358:241-276.

22.   Tatko CD & Waters ML (2004) Effect of halogenation on edge-face aromatic interactions in a beta-hairpin peptide: enhanced affinity with iodo-substituents. *Org Lett* 6(22):3969-3972.

23.   Schiele F, Ayaz P, & Fernandez-Montalvan A (2015) A universal homogeneous assay for high-throughput determination of binding kinetics. *Anal Biochem* 468:42-49.

24.   Lu Y, Wang Y, & Zhu W (2010) Nonbonding interactions of organic halogens in biological systems: implications for drug discovery and biomolecular design. *Phys Chem Chem Phys* 12(18):4543-4551.

25.   Bullock AN*, et al.* (2009) Kinase domain insertions define distinct roles of CLK kinases in SR protein phosphorylation. *Structure* 17(3):352-362.

26.   Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):48-57.

27.   Winn MD*, et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235-242.

28.   McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.

29.   Langer G, Cohen SX, Lamzin VS, & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7):1171-1179.

30.   Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.

31. Painter J & Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62(Pt 4):439-450.

32. Davis IW*, et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375-383.

33. Fedorov O*, et al.* (2007) A systematic interaction map of validated kinase inhibitors with Ser/Thr kinases. *Proc Natl Acad Sci U S A* 104(51):20523-20528.

34. Keates T*, et al.* (2012) Expressing the human proteome for affinity proteomics: optimising expression of soluble protein domains and in vivo biotinylation. *N Biotechnol* 29(5):515-525.

35. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, & Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 27(3):221-234.

36. Schmidt MW*, et al.* (1993) General atomic and molecular electronic structure system. *Journal of Computational Chemistry* 14(11):1347-1363.

37. Gora RW, Sokalski WA, Leszczynski J, & Pett VB (2005) The nature of interactions in the ionic crystal of 3-pentenenitrile, 2-nitro-5-oxo, ion(-1), sodium. *J Phys Chem B* 109(5):2027-2033.

38. Lee J*, et al.* (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J Chem Theory Comput* 12(1):405-413.

39. Maier JA*, et al.* (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* 11(8):3696-3713.

40. Ibrahim MA (2011) Molecular mechanical study of halogen bonding in drug discovery. *J Comput Chem* 32(12):2564-2574.

41. Dupradeau FY*, et al.* (2010) The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys Chem Chem Phys* 12(28):7821-7839.

42. Jorgensen WLC, J.; Madura, J.; Impey, R. W.; Klein, M.L., (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 79:926–935.

43. Ryckaert JP, Ciccotti G, & Berendsen H (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints – Molecular-Dynamics of N-Alkanes. *J Comput Phys* 23:327–341.

44. Cavallo L, Kleinjung J, & Fraternali F (2003) POPS: A fast algorithm for solvent accessible surface areas at atomic and residue level. *Nucleic Acids Res* 31(13):3364-3366.

Supplemental Figure S1
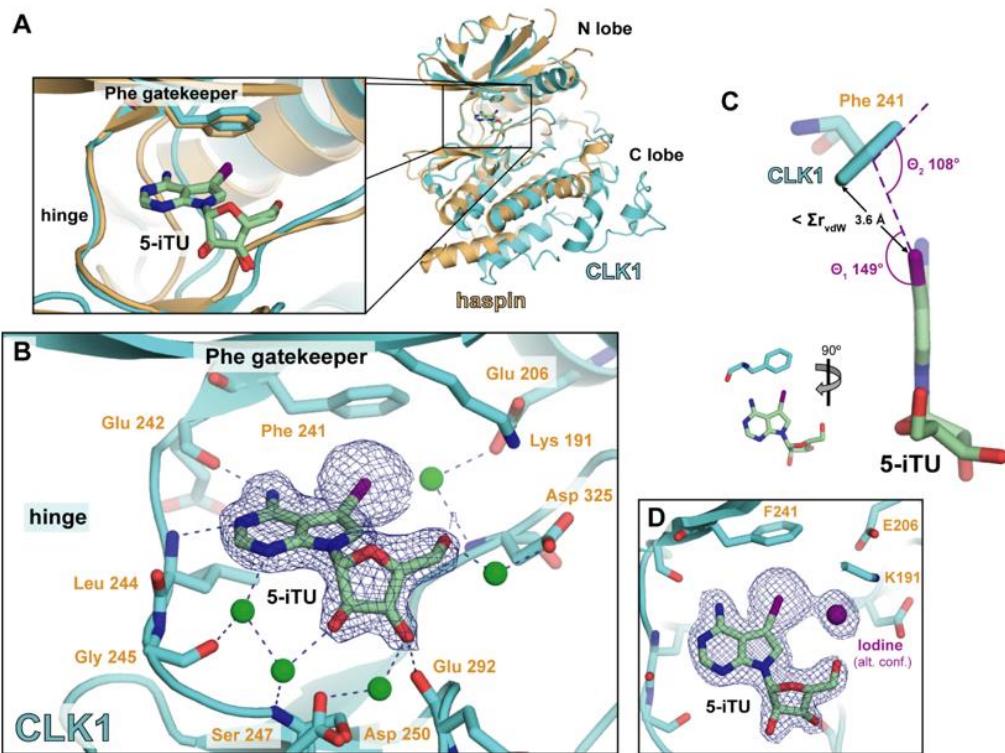
Supplemental Figure S1. The binding mode of 5-iTU is maintained in CLK1. (**A**) Superimposition of the structures of CLK1 and haspin (PDB ID: 4OUC) in complex with 5-iTU, showing that the binding mode is conserved. (**B**) Co-crystal structure of CLK1 and 5-iTU, highlighting interacting amino acid residues and the Phe gatekeeper. Water molecules are represented as green spheres, hydrogen bonds within 3 Å as blue dashed lines. Relevant atoms are colored as follows: oxygen – red, nitrogen – blue and iodine – purple. |2Fo|-|Fc | omitted electron density map is contoured at 3σ level. (**C**) Geometric measures of the putative π-X bond. (**D**) |2Fo|-|Fc | omitted electron density map contoured at 1σ level, showing alternative   conformation   of   delocalized   iodine   in   binding   pocket.

Supplemental Figure S2
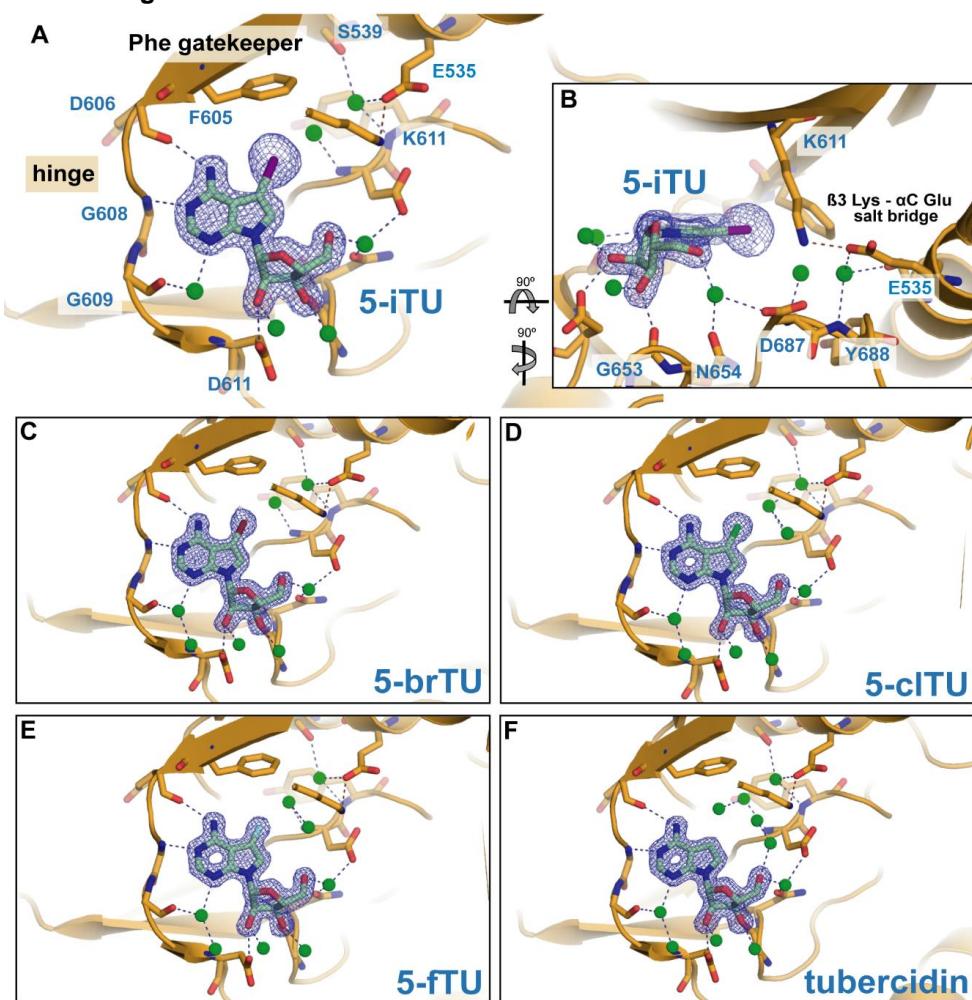
Supplemental Figure S2 - Structural studies of haspin in complex with the halogenated derivatives. (**A-B**) Co-crystal structure of haspin and 5-iTU, showing the ATP-competitive binding mode for the inhibitor. Interacting residues and the Phe gatekeeper are highlighted. Water molecules are represented as green spheres, hydrogen bonds within 3 Å as blue dashed lines. Relevant atoms are colored as follows: oxygen – red, nitrogen – blue and iodine – purple. |2Fo|-|Fc | omitted electron density map is contoured at 3σ level. (**C-F**) Co-crystal structures of haspin and different derivatives (see blue label), representation equivalent to (A). Bromine atom is shown in dark red, chlorine atom in green and fluorine atom in cyan.

Supplemental Figure S3
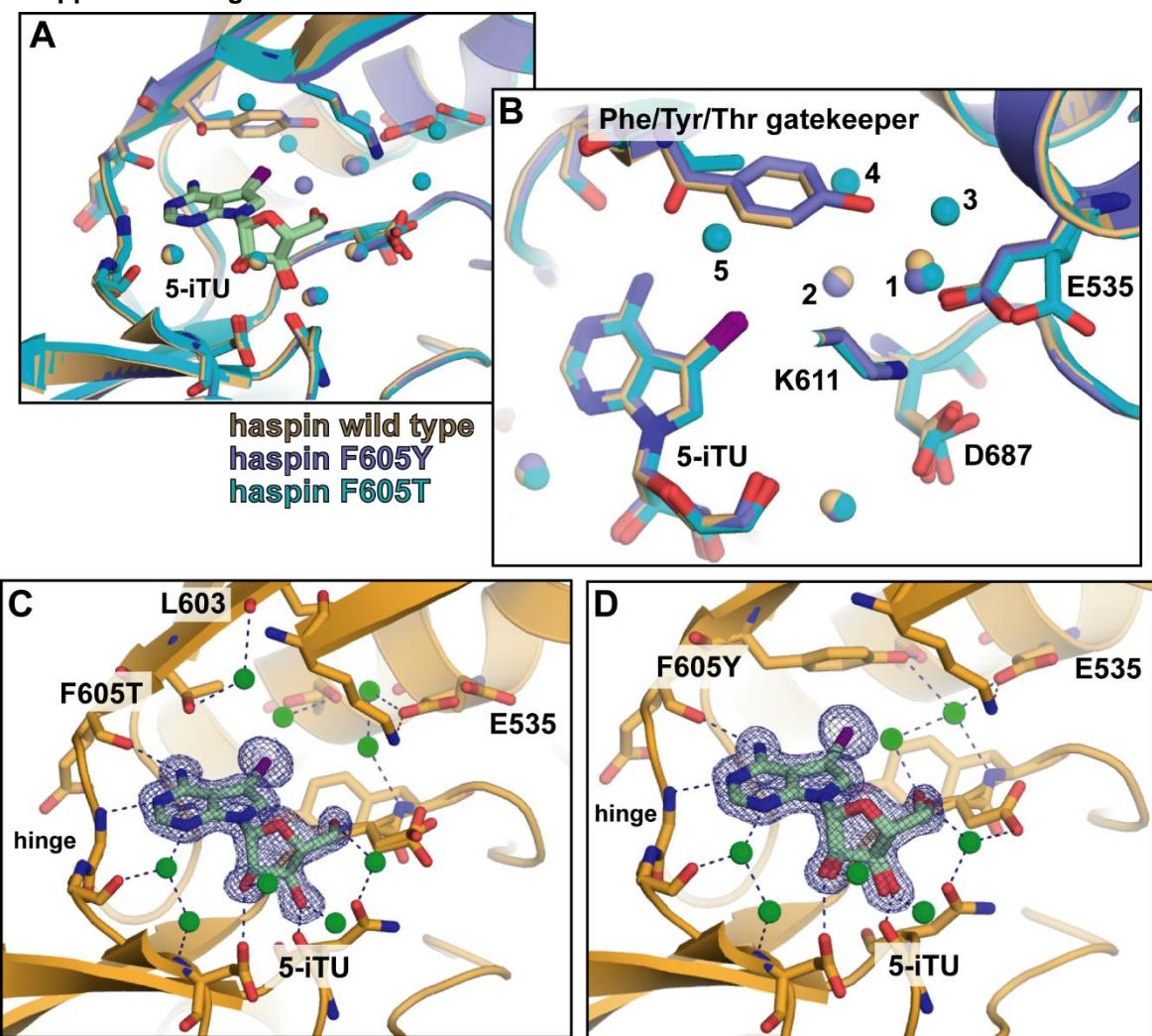
Supplemental Figure S3 - 5-iTU-haspin binding site is maintained in the absence of an aromatic gatekeeper residue. (**A**) Superimposition of co-crystal structures of wild-type haspin (orange), haspin[F605Y] (slate) and haspin[F605T] (cyan) in complex with 5-iTU, showing that the binding modus is maintained in gatekeeper mutants. (**B**) Similar to (A), highlighting relevant residues in the binding pocket. Water molecules are shown as spheres and numbered from 1-5. (**C**) Co-crystal structure of haspin[F605T] in complex with 5-iTU. Water molecules are represented as green spheres and hydrogen bonds within 3 Å as blue dashed lines. Relevant atoms are coloured as follows: oxygen – red, nitrogen – blue and iodine – purple. |2Fo|-|Fc | omitted electron density map is contoured at 3σ level. (**D**) Same as (C), showing the co-crystal structure of haspin[F605Y] in complex with 5-iTU.

Supplemental Figure S4
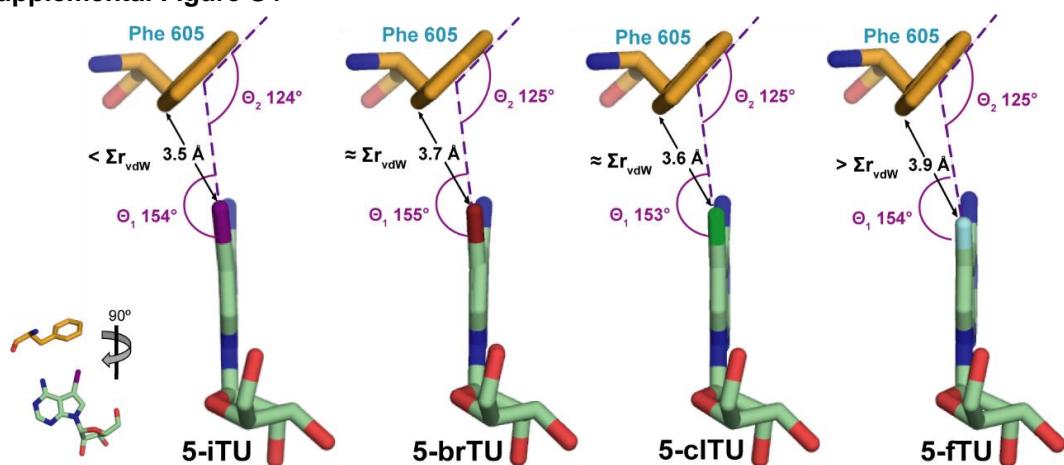
Supplemental Figure S4. Geometric measures of putative halogen-π-bonds. Overview of the halogenated derivatives approaching the aromatic ring of the Phe gatekeeper residue in haspin. The distance between the halogen and the closest carbon atom of the aromatic ring was measured and compared with the sum of the van der Waals radii ($\sum r_{vdW}$). Θ_1 is the angle between C-X bond to the centre of the phenylalanine aromatic group (C-X···π), while Θ_2 is the angle of the halogen to the plane of phenylalanine aromatic ring (X···π-C).

Supplemental Figure S5
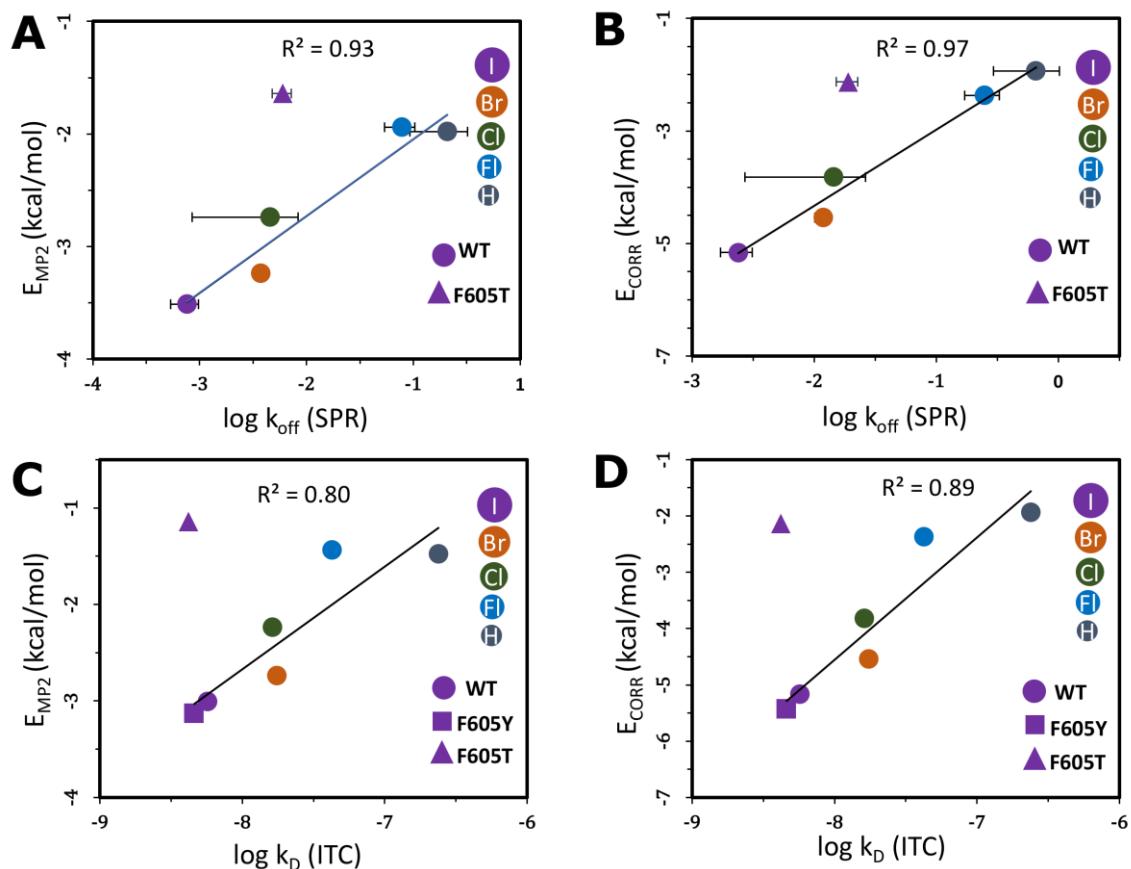
Supplemental Figure S5 – Correlation plots of computed quantum mechanical energies against experimental binding parameters. A) Second-order Møller–Plesset interaction energy (E_{MP2}) between the tubercidin derivatives and the gatekeeper residue versus the experimental (SPR) dissociation rate constants (k_{off}) of the tubercidin derivatives. B) The second-order correlation correction energy term (E_{CORR}) for the interaction between the tubercidin derivatives and the gatekeeper residue versus the experimental (SPR) dissociation rate constants (k_{off}) of the tubercidin derivatives This correlation energy (E_{CORR}) includes second-order intermolecular dispersion interactions and the correlation corrections to the Hartree-Fock (HF) energy. C) Second-order Møller–Plesset interaction energy (E_{MP2}) between tubercidin derivatives and gatekeeper residue versus the experimental (ITC) binding affinities (k_D) of the tubercidin derivatives. D) Second-order correlation correction energy term (E_{CORR}) for the interaction between the tubercidin derivatives and the gatekeeper residue versus the experimental (ITC) binding affinities (k_D) of the tubercidin derivatives. The correlation coefficients (R^2) and the linear fits were computed omitting the outlier data points for the F605T mutant. The error bars for the K_D (ITC) values are smaller than the size of the data point symbols.

Supplemental Table S1. DSF results for 5-iTU against 137 kinases. All results are listed with decreasing melting temperature shifts (ΔT_m). The gatekeeper residues (GK) of the kinases were determined by sequence alignment and analysis of the structures for hits with $\Delta T_m > 2^0$.

Kinase	ΔT_m	GK	Kinase	ΔT_m	GK	Kinase	ΔT_m	Kinase	ΔT_m
Haspin	11.40	Phe	TTK	2.20	Met	MAPK2TG	0.79	CAMK1G	0.20
DYRK2	10.38	Phe	RSK4	2.20	Leu	GRK1	0.76	CAMK1D	0.20
CLK4	9.59	Phe	AMPK1	2.19	Met	MYLK	0.73	CDPK1PF	0.12
CLK1	8.60	Phe	PIM2	2.04	Leu	CHEK2	0.70	VRK3	0.10
CLK3	7.50	Phe	GSK3β	2.00	Leu	CAMK2A	0.70	VRK2	0.10
CK1ε	5.90	Met	CK2α2	1.97	Phe	RIPK2	0.69	TOPK	0.10
DYRK1A	5.80	Phe	MAP2K2	1.90		MAPK9A	0.60	CAMK2B	0.10
CLK2	5.20	Phe	PRKCL1	1.83		CAMK2G	0.60	AAK1	0.09
CDK2	4.98	Phe	ITK	1.78		BMP2K	0.60	ADRBK2	0.03
MST3	4.96	Met	STK38	1.70		RPS6KA1	0.50	BMX	0.02
DRAK2	4.84	Leu	SNF1LK	1.70		AMPKA2	0.50	PRKG2	0.00
LOK	4.70	Ile	RPS6KA2	1.70		LOC340156	0.48	PDK1	0.00
YSK1	4.30	Met	RIOK2	1.70		ADRBK1	0.48	NLK	-0.10
SLK	4.20	Ile	PRKCL2	1.70		MAPK13	0.46	MAP2K6	-0.10
CAMK4	3.75	Leu	MERTK	1.70		TEC	0.45	NEK6	-0.20
AMPKα2	3.58	Met	PDPK1	1.69		CDKL2	0.43	CDK6	-0.20
MPSK1	3.40	Leu	RPS6KA3	1.66		NEK7	0.40	PKMYT1	-0.22
CK1γ2	3.30	Leu	PRKCZ	1.56		NEK2	0.40	DDR1	-0.38
PHKγ2	3.24	Phe	MAPK3A	1.50		MAP3K5	0.40	CDC42BPG	-0.38
RIPK3	3.21	Thr	TLK1	1.45		JAK1	0.40	MAPK11	-0.40
CDKL5	3.16	Phe	PIM1	1.40		SRPK2	0.39	CAMK2D	-0.40
MST1	3.10	Met	NEK11	1.31		CDPK1PV	0.39	DMPK1	-0.58
STK33	3.06	Leu	PRKACA	1.30		BRK1	0.39	CDC42BPAB	-0.74
PLK4	2.90	Leu	PIM3	1.30		MAPK2PF	0.35		
DAPK3	2.90	Leu	GAK	1.30		CDC42BPB	0.35		
DRAK1	2.80	Leu	CDKL3	1.30		ZAK	0.34		
PFTAIRE1	2.71	Phe	MYLK2	1.28		DCAMKL	0.33		
CDK8	2.71	Phe	IKBKB	1.25		YANK3	0.30		
MST2	2.70	Met	PRKD2	1.24		PRKX	0.28		
ERK3	2.70	Gln	PCTK1	1.11		TYK2	0.27		
CSNK1γ3	2.70	Leu	NEK9	1.06		MSSK1	0.26		
PRKCN	2.62	Met	CSNK2A1	1.05		LIMK1	0.26		
NDR2	2.60	Met	TNIK	1.00		SRPK1	0.25		
CK1γ1	2.50	Leu	PRKG1	1.00		VRK1	0.20		
QIK	2.41	Ile	NEK1	0.96		TYRO3	0.20		
STK39	2.40	Met	PCTK2	0.83		PAK6	0.20		
MEK1	2.29	Met	MST4	0.80		PAK5	0.20		
GRK5	2.22	Leu	YANK1	0.79		CDKL1	0.20		

Supplemental Table 2. BLI and ITC data.

Supplemental Table S2A - Summary of ITC and BLI data for 5-iTU in complex with CLK1 and CLK3

5-iTU vs	ITC data						BLI data		
	n	ΔH kcal/mol	$T\Delta S$ kcal/mol	ΔG kcal/mol	K M^{-1}	K_d nM	k_{on} $M^{-1} s^{-1}$	k_{off} s^{-1}	Resi-denceTime
CLK1	0.96	-16.66 (± 0.09)	-5.94	-10.72	1.39e+8 (± 1.99e+7)	**7.2** (±1.0)	10.9e+4 (±8.48e+3)	3.14e-4 (±1.06e-5)	**53.1 min** (±6.3 min)
CLK3	1.00	-12.04 (± 0.05)	-2.44	-9.60	1.91e+7 (± 1.28e+6)	**52.4** (±3.5)	-	-	-

Supplemental Table S2B- Summary of ITC and BLI data for halogenated derivatives against haspin

haspin vs	ITC data						BLI data		
	n	ΔH kcal/mol	$T\Delta S$ kcal/mol	ΔG kcal/mol	K M^{-1}	K_d nM	k_{on} $M^{-1} s^{-1}$	k_{off} s^{-1}	Resi-denceTime
tubercidin	0.95	-16.95 (± 0.10)	-8.21	-8.74	4.18e+6 (±1.74e+5)	**239.2** (±10.0)	9.56e+4 (±1.45e+3)	4.74e-2 (±6.72e-4)	**21.1 sec** (±0.3 sec)
5-FTu	1.00	-18.10 (± 0.07)	-8.39	-9.71	2.34e+7 (±1.56e+6)	**42.7** (±2.8)	9.26e+4 (±9.71e+2)	6.32e-3 (±4.44e-5)	**2.6 min** (±1.1 sec)
5-ClTu	0.98	-21.54 (± 0.06)	-11.27	-10.27	6.14e+7 (±3.29e+6)	**16.3** (±0.9)	9.00e+4 (±5.94e+2)	9.22e-4 (±8.00e-6)	**18.1 min** (±9.4 sec)
5-BrTu	1.00	-22.50 (±0.07)	-12.28	-10.22	5.70e+7 (±4.26e+6)	**17.5** (±1.3)	9.57e+4 (±3.36e+2)	6.42e-4 (±5.76e-5)	**25.9 min** (±2.3 min)
5-ITu	0.96	-23.40 (±0.07)	-12.53	-10.87	1.75e+8 (±1.66e+7)	**5.7** (±0.5)	9.59e+4 (±5.61e+2)	2.76e-4 (±5.96e-6)	**60.4 min** (±1.3 min)

Supplemental Table S2C- Summary of ITC and BLI data for haspin gatekeeper mutants against 5-iTU

5-iTU vs	ITC data						BLI data		
	n	ΔH kcal/mol	$T\Delta S$ kcal/mol	ΔG kcal/mol	K M^{-1}	K_d nM	k_{on} $M^{-1} s^{-1}$	k_{off} s^{-1}	Resi-denceTime
F605Y	0.96	-22.26 (± 0.05)	-11.27	-10.99	2.34e+7 (± 1.56e+6)	**4.6** (±0.3)	11.2e+4 (±1.45e+3)	4.29e-4 (±8.87e-6)	**38.9 min** (±48 sec)
F605H	0.98	-20.35 (± 0.05)	-9.42	-10.93	6.14e+7 (± 3.29e+6)	**5.1** (±0.3)	-	-	-
F605M	0.94	-22.09 (±0.04)	-10.89	-11.20	5.70e+7 (± 4.26e+6)	**3.2** (±0.2)	-	-	-
F605L	0.97	-23.35 (±0.04)	-12.36	-10.99	1.75e+8 (± 1.66e+7)	**4.6** (±0.4)	-	-	-
F605T	0.93	-20.48 (±0.03)	-9.42	-11.06	1.75e+8 (± 1.66e+7)	**4.2** (±0.4)	12.4e+4 (±6.22e+2)	1.23e-3 (±6.54e-6)	**13.5 min** (±4.3 sec)
F605Q	0.98	-23.23 (±0.04)	-12.56	-10.67	1.75e+8 (± 1.66e+7)	**8.0** (±0.8)	-	-	-

Supplemental Table S3. Summary of SPR data for haspin wild-type and F605T mutant against tubercidin derivatives

A wild-type (F605)

	K_D eq (nM)	K_D kin (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})
tubercidin (hydrogen)	2560 ±1230	2910 ±707	1.96×10^5 ±1.59×10^5	6.57×10^{-1} ±3.62×10^{-1}
5-fTU (fluoride)	155 ±91	79.3 ±35.9	2.06×10^6 ±2.07×10^6	2.49×10^{-1} ±0.79×10^{-1}
5-clTU (Chloride)	6.6 ±2.8	4.9 ±2.8	2.79×10^6 ±1.26×10^6	1.45×10^{-2} ±1.18×10^{-2}
5-brTU (Bromide)	2.6 ±0.3	1.3 ±0.3	9.71×10^6 ±2.67×10^6	1.19×10^{-2} ±0.18×10^{-2}
5-iTU (iodide)	0.8 ±0.2	0.3 ±0.1	9.39×10^6 ±2.03×10^6	2.40×10^{-3} ±0.70×10^{-3}
fold difference				
H to I	↓ 3200	↓ 9700	↑ 48	↓ 274
F to I	↓ 194	↓ 264	↑ 5	↓ 104

B mutant F605T

	K_D eq (nM)	K_D kin (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})
tubercidin (hydrogen)	5410 ±2180	6870	5.53×10^4 ±1.25×10^4	3.29×10^{-1} ±0.68×10^{-1}
5-fTU (fluoride)	1660 ±661	1490 ±258	1.22×10^5 ±0.40×10^5	1.76×10^{-1} ±0.33×10^{-1}
5-clTU (Chloride)	166 ±81	154 ±44	7.47×10^5 ±1.00×10^5	1.12×10^{-1} ±0.25×10^{-1}
5-brTU (Bromide)	53.3 ±24.6	46.3 ±9.1	1.07×10^6 ±0.12×10^6	4.88×10^{-2} ±0.77×10^{-2}
5-iTU (iodide)	12.8 ±5.1	12.0 ±2.2	1.61×10^6 ±0.44×10^6	1.89×10^{-2} ±0.38×10^{-2}
fold difference				
H to I	↓ 422	↓ 573	↑ 29	↓ 17
F to I	↓ 129	↓ 124	↑ 13	↓ 9

C wild-type (F605) vs mutant F605T on 5-iTU

fold difference	K_D eq (nM)	K_D kin (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})
5-iTU	↓ 16	↓ 40	↑ 6	↓ 8

Supplemental Table S4. Solvent accessible surface area of ordered water molecules in the binding pocket. The Solvent Accessible Surface Area (SASA) of water molecules W1-W5 (as shown in Supplemental Figure 3), the nitrogen atom of the ε-amino group of K511 (NZ) and the closest C atom of the aromatic ring of F605 (CD1), was obtained by the POPS server, using a probe with 1.4 Å radius(44). The structures with the halogenated derivatives are compared with the haspin apo-structure (PDB ID: 2WB8).

SASA (Å^2)	haspin/ 5-iTU	haspin/ 5-brTU	haspin/ 5-clTU	haspin/ 5-fTU	haspin/ tubercidin	Apo haspin
W1	2.13	2.20	1.88	1.84	1.74	1.34
W2	3.14	2.69	2.64	2.64	1.70	2.06
W3	-	-	4.03	3.94	3.50	5.07
W4	-	-	-	-	3.47	-
W5	-	-	-	-	3.21	-
NZ (K511)	4.56	4.88	3.80	3.46	4.02	4.23
CD$_1$ (F605)	1.49	1.39	1.44	1.41	1.37	1.31

Supplemental Table S5. Measured geometric parameters between the halogens and the Phe gatekeeper

	CLK1 / 5-iTU	haspin / 5-iTU	haspin / 5-brTU	haspin / 5-clTU	haspin / 5-fTU	Preferred geometry of X-bonds
Distance between halogen and closest C atom of gatekeeper	3.58 Å ± 0.28 Å ($< \sum r_{vdW}$)	3.52 Å ± 0.18 Å ($< \sum r_{vdW}$)	3.68 Å ± 0.19 Å ($\approx \sum r_{vdW}$)	3.61 Å ± 0.17 Å ($\approx \sum r_{vdW}$)	3.85 Å ± 0.17 Å ($> \sum r_{vdW}$)	$\leq \sum r_{vdW}$
Sum of van der Waals radii	3.68 Å	3.68 Å	3.60 Å	3.45 Å	3.20 Å	-
Angle Θ_1	149.5º	153.7º	154.7 º	153.3º	154.0º	$\approx$ 160º - 165º
Angle Θ_2	108.0º	124.0º	125.1º	124.6º	125.3º	$\approx$ 120º ($\approx$ 90º for π-x-bond)

Supplemental Table S6

Inhibitor	log k_{off} (SPR)	log k_{off} (BLI)	log K_D (SPR)	log K_D (ITC)	E_{EL}	E_{HL}	E_{SCF}	E_{MP2}
5-iTU	-2.62	-3.56	-9.10	-8.24	-1.85	2.79	2.15	-3.01
5-brTU	-1.92	-3.19	-8.59	-7.76	-1.09	2.22	1.78	-2.74
5-clTU	-1.84	-3.04	-8.18	-7.79	-0.52	1.91	1.58	-2.24
5-fTU	-0.60	-2.20	-6.81	-7.37	0.17	1.15	0.94	-1.44
tubercidin	-0.18	-1.32	-5.59	-6.62	-0.09	0.65	0.47	-1.48
Pearson Correlation coefficient (R^2) with log k_{off} (SPR)					0.82	-0.98	-0.98	0.93
Pearson Correlation coefficient (R^2) with log k_{off} (BLI)					0.82	-0.98	-0.99	0.91
Pearson Correlation coefficient (R^2) with log K_D (SPR)					0.71	-0.97	-0.99	0.87
Pearson Correlation coefficient (R^2) with log K_D (ITC)					0.79	-0.96	-0.97	0.86

Supplemental Table S6 - Total interaction energy [$kcal. \; mol^{-1}$] between tubercidin derivatives and gatekeeper Phe 605 residue at consecutively increasing levels of quantum mechanical theory, see equation 1. E_{EL} is the electrostatic energy only, E_{HL} includes the Heitler-London energy, E_{SCF} includes the Hartree-Fock energy as well, and E_{MP2} is the full Moeller-Plesset second order energy. k_{off} values were measured by SPR and BLI, and K_D values were measured by SPR and ITC.

Supplemental Table S7

Inhibitor	log k_{off} (SPR)	log k_{off} (BLI)	log K_D (SPR)	log K_D (ITC)	$E_{EL,MTP}$	$E_{EL,PEN}$	E_{EX}	E_{DEL}	E_{CORR}
5-iTU	-2.62	-3.56	-9.10	-8.24	2.48	-4.33	4.63	-0.64	-5.16
5-brTU	-1.92	-3.19	-8.59	-7.76	1.03	-2.12	3.30	-0.42	-4.54
5-clTU	-1.84	-3.04	-8.18	-7.79	-0.21	-0.32	2.44	-0.34	-3.82
5-fTU	-0.60	-2.20	-6.81	-7.37	0.13	0.04	0.97	-0.21	-2.37
tubercidin	-0.18	-1.32	-5.59	-6.62	-0.07	-0.01	0.74	-0.18	-1.94
Pearson Correlation coefficient (R^2) with log k_{off} (SPR)					-0.55	0.69	-0.94	0.89	0.98
Pearson Correlation coefficient (R^2) with log k_{off} (BLI)					-0.68	0.76	-0.92	0.88	0.96
Pearson Correlation coefficient (R^2) with log K_D (SPR)					-0.49	0.61	-0.88	0.80	0.95
Pearson Correlation coefficient (R^2) with log K_D (ITC)					-0.70	0.76	-0.90	0.88	0.92

Supplemental Table S7 - Contribution of the different interaction energy terms to the total interaction energy, E_{MP2} [$kcal. \; mol^{-1}$], between tubercidin derivatives and the gatekeeper Phe 605 residue. See equation 1 for the definition of the terms. $E_{EL,MTP}$ is the electrostatic multipole term, $E_{EL,PEN}$ is the penetration electrostatic term, E_{EX} is the exchange term, E_{DEL} is the delocalization term, and E_{CORR} is the correlation energy term. k_{off} values were measured by SPR and BLI, and K_D values were measured by SPR and ITC.

Supplemental Table S8

System	log k_{off} (SPR)	log k_{off} (BLI)	log K_D (SPR)	log K_D (ITC)	E_{EL}	E_{HL}	E_{SCF}	E_{MP2}
Wild type	-2.62	-3.56	-9.10	-8.24	-1.85	2.79	2.15	-3.01
F605Y	ND	-3.37	ND	-8.34	-2.27	3.01	2.30	-3.13
F605T	-1.72	-2.91	-7.89	-8.38	-0.78	1.38	0.99	-1.14
Pearson Correlation coefficient (R^2) with log k_{off} (BLI)					0.84	-0.91	-0.92	0.94
Pearson Correlation coefficient (R^2) with log K_D (ITC)					-0.52	0.64	0.65	-0.69

Supplemental Table S8 - Total interaction energy [*kcal. mol^{-1}*] between tubercidin derivatives and gatekeeper Phe 605 residue at consecutively increasing levels of quantum mechanical theory, see equation 1. E_{EL} is the electrostatic energy only, E_{HL} includes the Heitler-London energy, E_{SCF} includes the Hartree-Fock energy as well, and E_{MP2} is the full Moeller-Plesset second order energy. k_{off} values were measured by SPR and BLI, and K_D values were measured by SPR and ITC. The correlation coefficient was calculated only for k_{off} and K_D values measured by BLI and ITC, respectively, as there were no SPR data available for the F605Y mutant.

Supplemental Table S9

System	log k_{off} (SPR)	log k_{off} (BLI)	log K_D (SPR)	log K_D (ITC)	$E_{EL,MTP}$	$E_{EL,PEN}$	E_{EX}	E_{DEL}	E_{CORR}
Wild type	-2.62	-3.56	-9.10	-8.24	2.48	-4.33	4.63	-0.64	-5.16
F605Y	ND	-3.37	ND	-8.34	0.41	-2.68	5.28	-0.71	-5.43
F605T	-1.72	-2.91	-7.89	-8.38	-6.97	6.19	2.17	-0.39	-2.13
Pearson Correlation coefficient (R^2) with log k_{off} (BLI)					-1.00	0.99	-0.88	0.88	0.93
Pearson Correlation coefficient (R^2) with log K_D (ITC)					0.86	-0.82	0.58	-0.57	-0.68

Supplemental Table S9 - Contribution of the different interaction energy terms to the total interaction energy, E_{MP2} [*kcal. mol^{-1}*] between 5-iTU and the gatekeeper residue for the wild type and the two mutants. See equation 1 for the definition of the terms. $E_{EL,MTP}$ is the electrostatic multipole term, $E_{EL,PEN}$ is the penetration electrostatic term, E_{EX} is the exchange term, E_{DEL} is the delocalization term, and E_{CORR} is the correlation energy term. k_{off} values were measured by SPR and BLI, and K_D values were measured by SPR and ITC. The correlation coefficient was calculated only for k_{off} and K_D values measured by BLI and ITC, respectively, as there were no SPR data available for the F605Y mutant.

Supplemental Table S10

Inhibitor	$\Delta E_{gas} + \Delta\Delta G_{solvation}$ (Kcal/mol)	$T\Delta S_{MMGBSA}$ (Kcal/mol)	ΔG_{MMGBSA} (Kcal/mol)
5-iTU	-25.95 ± 2.93	-23.52 ± 0.02	-2.43 ± 2.95
5-brTU	-26.00 ± 2.75	-23.27 ± 0.02	-2.73 ± 2.77
5-clTU	-21.58 ± 3.33	-22.97 ± 0.02	1.39 ± 3.35
5-fTU	-20.31 ± 3.04	-22.83 ± 0.01	2.52 ± 3.05
tubercidin	-19.45 ± 2.87	-22.65 ± 0.02	3.20 ± 2.89

Supplemental Table S10 - Binding free energies calculated using the MMGBSA approach for the binding of tubercidin derivatives with haspin.

Supplemental Table SS11 - Data collection and refinement statistics

	haspin-5-iTU	haspin-5-brTU	haspin-5-clTU	haspin-5-fTU	haspin-tubercidin	haspin[F605Y]-5-iTu	haspin[F605T]-5-iTu	CLK1-5-iTu
Data collection								
Space group	$P\,2_12_12_1$	$P\,2_12_12_1$	$P\,2_12_12_1$	$P\,2_12_12_1$	$P\,2_12_12_1$	$P\,2_12_12_1$	$P\,2_12_12_1$	$P\,121$
Wavelength (Å)	0.9794	0.9795	0.9794	0.9795	0.9794	0.9795	0.9795	1.0282
Cell dimensions								
a, b, c (Å)	78.6, 78.7, 79.7	78.3, 78.8, 79.4	77.7, 78.9, 79.8	78.7, 78.9, 79.6	78.6, 78.8, 80.0	78.8, 79.0, 79.5	78.4, 78.7, 79.8	56.2, 116.2, 91.0
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 99.0, 90.0
Resolution (Å) [a]	23.98-1.76	23.87-1.55	23.78-1.46	23.80-1.48	24.05-1.43	22.01-1.45	22.01-1.43	12.96-2.00
	(1.79-1.76)	(1.63-1.55)	(1.48-1.46)	(1.51-1.48)	(1.45-1.43)	(1.53-1.45)	(1.45-1.43)	(2.04-2.00)
R_{merge} [a]	0.087 (0.224)	0.073 (0.643)	0.075 (0.629)	0.149 (2.022)	0.102 (1.135)	0.060 (0.570)	0.122 (2.039)	0.104 (0.949)
$I / \sigma I$ [a]	13.9 (7.2)	10.8 (2.3)	12.3 (2.5)	9.9 (2.5)	8.6 (1.2)	14.2 (3.0)	11.3 (2.7)	7.6 (1.6)
Completeness (%)	98.1 (96.8)	99.8 (99.8)	99.8 (98.1)	100.0 (100.0)	99.9 (98.3)	98.3 (96.9)	99.8 (99.5)	99.6 (100)
CC1/2 [a]	0.994 (0.968)	0.996 (0.758)	0.996 (0.741)	0.994 (0.648)	0.994 (0.379)	0.997 (0.801)	0.995 (0.751)	0.997 (0.680)
Redundancy [a]	6.2 (6.5)	5.2 (5.2)	6.0 (5.5)	9.3 (9.2)	5.9 (4.2)	5.4 (5.4)	9.3 (9.5)	5.2 (5.2)
Unique reflections	48,665 (2,699)	71,557 (10,321)	85,664 (4,145)	83,206 (4,087)	92,144 (4,423)	86,860 (12,366)	91,424 (4,471)	106,488 (5242)
Refinement								
R_{work}/R_{free}	0.138/0.174	0.160/0.183	0.149/0.176	0.148/0.167	0.143/0.165	0.158/0.175	0.150/0.175	0.188/0.218
No. atoms								
Protein	2,602	2,778	2,742	2,612	2,595	2,726	2,655	8,251
Inhibitor	21	42	53	20	19	42	25	120
Others[b]	249	336	253	220	293	356	236	346
B-factors								
Protein	21.45	27.38	21.82	19.59	21.52	23.86	20.65	26.74
5-ITu/derivative	16.85	18.87	26.56	13.66	14.84	16.47	15.37	27.75
Water/solvents	27.38	38.58	27.87	25.69	30.48	35.89	25.40	45.94
r.m.s. deviations[c]								
bond length (Å)	0.017	0.016	0.013	0.013	0.012	0.015	0.013	0.015
bond angles (°)	1.8	1.6	1.6	1.6	1.5	1.6	1.6	1.5
Ramachandran								
outliers (%)	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0
favored (%)	98.48	98.85	98.55	98.48	98.78	98.84	98.51	97.7
Coordinate error (Luzzati plot)	0.18 A	0.19 A	0.17	0.17	0.17	0.17	0.18	0.28

[a] values in parentheses refer to the highest resolution shell.
[b] Others indicate water and solvent molecules

[c] r.m.s. deviations indicates root mean square deviations