Support Information

High-throughput Virtual Screening using Quantum Mechanical Probes: Discovery of Selective Kinase Inhibitors

Ting Zhou and Amedeo Caflisch*

June 22, 2010

Department of Biochemistry, University of Zurich,

Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Phone: (+41 44) 635 55 21, Fax: (+41 44) 635 68 62

Email: caflisch@bioc.uzh.ch

*To whom correspondence should be addressed.

Assessment of QM probe method on protein kinase CDK2

Before applying the quantum mechanical (QM) probe method to EphB4, it was tested on the cyclin-dependent kinase 2 (CDK2) for which a large number of ATP-binding site inhibitors have been published.[1–8] The structure of CDK2 was downloaded from the PDB (PDB entry 1KE5), and hydrogen atoms were added by CHARMM[9] according to the protonation states of side chains and termini at pH 7. Then the structure was minimized with CHARMM using the CHARMm[10] force field and MPEOE partial charges.

The catalytic domain in protein kinases is composed of two lobes connected by a segment termed "hinge loop". The majority of ATP-competitive inhibitors are involved in at least one hydrogen bond with the hinge loop.[11] There are two hydrogen bond (HB) acceptors and one donor in the backbone of the hinge loop so that two water probes and one hydrogen fluoride (HF), respectively, were used (Figure S-I). About 1,000 compounds, randomly selected from the ZINC library, were docked into the ATP-binding site of cyclin-dependent kinase 2 (CDK2) using version 4 of AutoDock[12]. A total of about 100,000 poses were generated by docking, then minimized using the CHARMm force field, and filtered by van der Waals (vdW) interaction energy (IE) and vdW efficiency as mentioned in ref 7. To study the effectiveness of the probe method, we selected poses based on probe energies and visually inspected whether there is a particular interaction at the expected position.

Figure S-II shows the QM probe energy is an effective detector of canonical HBs. The structures of four putatively inactive compounds (compound **27**–**30** termed decoys hereafter) and the cocrystallized ligand in PDB entry 1KE5, and their interactions with the protein are schematically shown in Figure S-II. From the vdW point of view, the four decoys match the binding pocket, since they all passed the filters of vdW and vdW efficiency.[7] However their unfavorable polar interactions with the hinge region are not detected. A main reason of failure in detecting the unfavorable polar contacts is that the E_{ele} averages out the electrostatic interaction

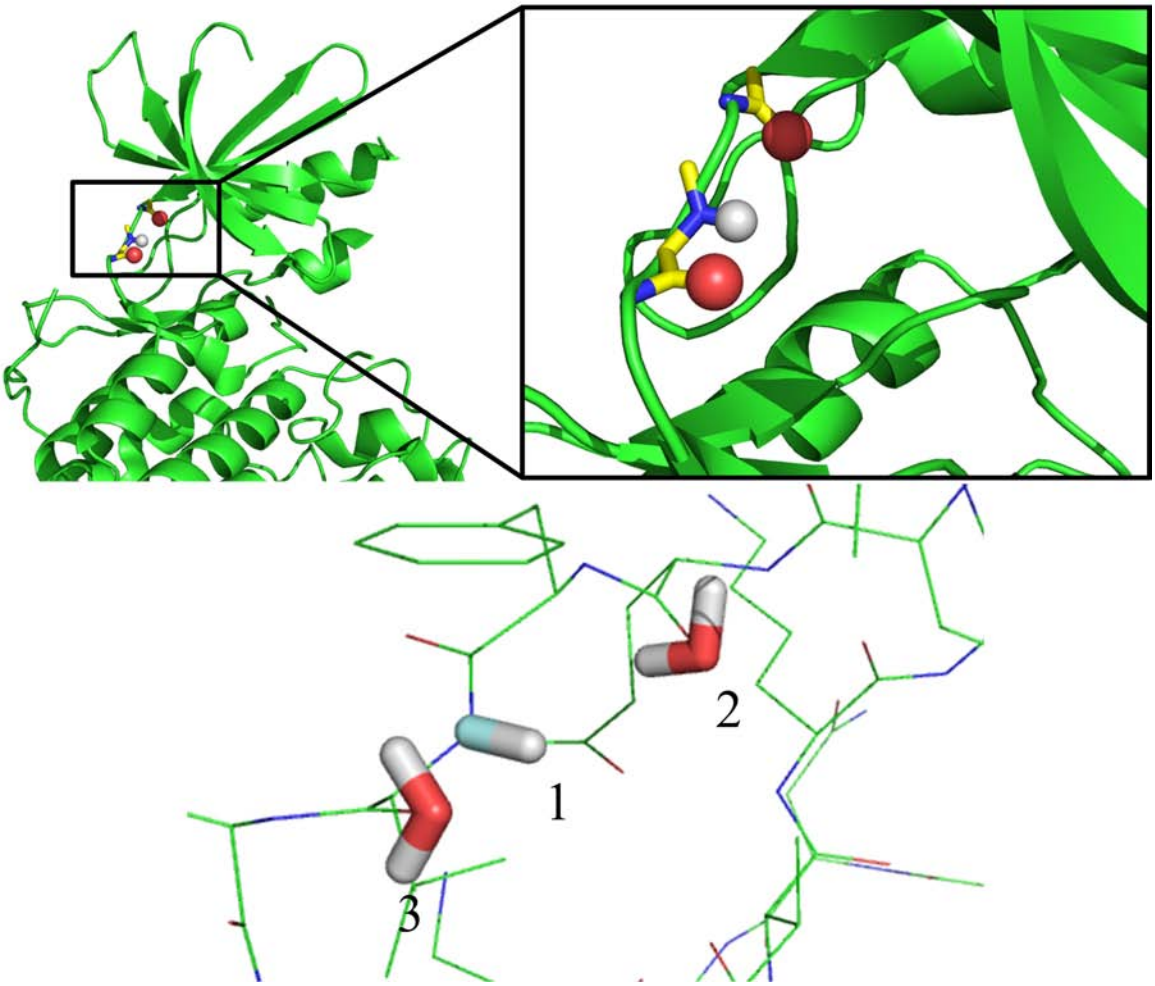
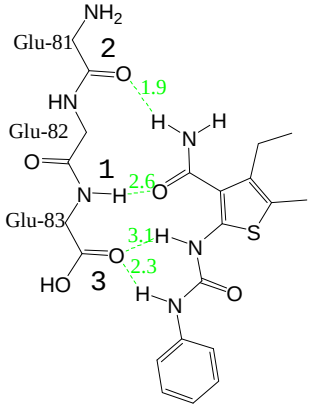


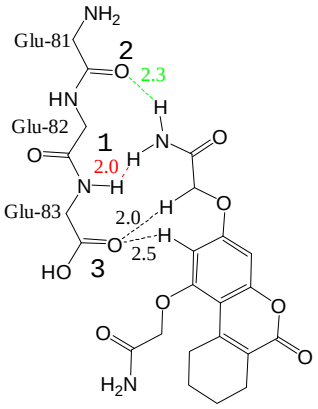
Figure S-I: The three QM probes at the hinge loop of the protein kinase CDK2. Positions 1, 2, and 3 are the backbone –NH– of Leu83, –CO– of Glu81, and –CO– of Leu83, respectively.

between the decoys and the protein, hence the resolution or sensitivity is not high enough. This averaging effect is also present for the known inhibitor (Figure S-II) in the structure of PDB entry 1KE5 which was optimized and evaluated using the same protocol, and it has $E_{\text{ele}} = -5.02$ kcal·mol^{-1}, and $E_{\text{ele}}/\text{MW} = -0.0153$ kcal·g^{-1}, where MW is the molecular weight. Electrostatic IE and electrostatic efficiency of this cocrystallized inhibitor are quite similar to some inactive compounds (Figure S-II). Nevertheless, with QM probe method, an adverse contact always shows a positive (unfavorable) probe energy, and can be easily distinguished. The three probe energies of the cocrystallized inhibitor are -3.94 kcal·mol^{-1}, -3.27 kcal·mol^{-1}, and -1.63 kcal·mol^{-1}, while each of the compound **28**–**30** has one positive probe energy, which is consistent with the unfavorable polar contact in the pose (Figure S-II). Note that compound **27** is not eliminated by the filters of the three probes at the hinge region, but was not selected for experimental testing, since the hydrophobic pocket is not satisfied (see subsection 3.3).

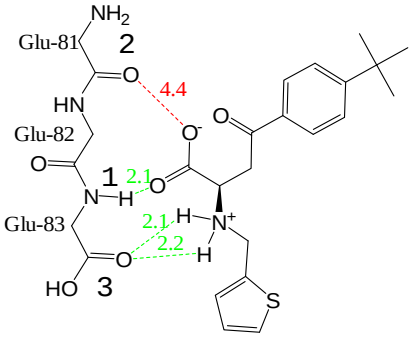
The QM probe method is able to detect non-classical HBs. Most of the force fields use fix-charge approximation to describe electrostatic interactions. QM gains an advantage in evaluation of complex charge–charge interactions, e.g., anion–cation interaction,[13] metal–ligand interaction,[14] and HB.[15] The non-classical HBs exist in protein–ligand complex extensively,[16] not only in kinase cases,[17] but also in other targets.[18] The compound in Figure S-II(b) is involved in a pair of C$-$H$\cdots$O non-classical HBs with the protein (colored in black).[17] The probe energy of Probe 3 is -3.22 kcal·mol^{-1}, which expresses a distinct sign of a favorable interaction there. Figure S-III shows another compound (compound **31**) interacting with the hinge loop at Probe 2 by a non-classical HB, whose probe energy (-2.02 kcal·mol^{-1}) is also comparable with that of a classical HB. The partial charge of the hydrogen atom bonded to the C_6 (red in Figure S-III) is identical with that of the hydrogen atom in unsubstituted phenyl ring if the partial charges are assigned using MPEOE approach. This is not accurate because the $-NO_2$ group is strongly electron-withdrawing, and the $=$CH$=$ at the para position of the $-NO_2$ is more positive (atomic charges calculated by QM at MP2 6-31+G(d,p) level is shown in Figure S-VI) than the analogue



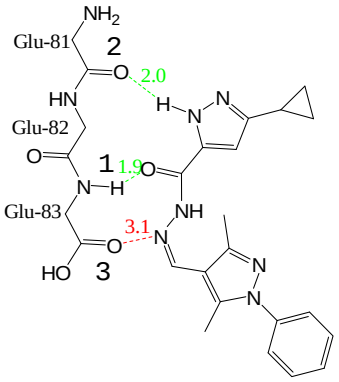
(a) Three probes match the decoy well.



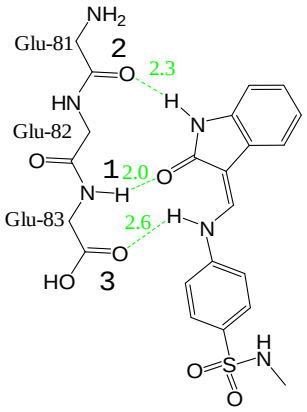
(b) Probe 1 mismatches the decoy.



(c) Probe 2 mismatches the decoy.



(d) Probe 3 mismatches the decoy.



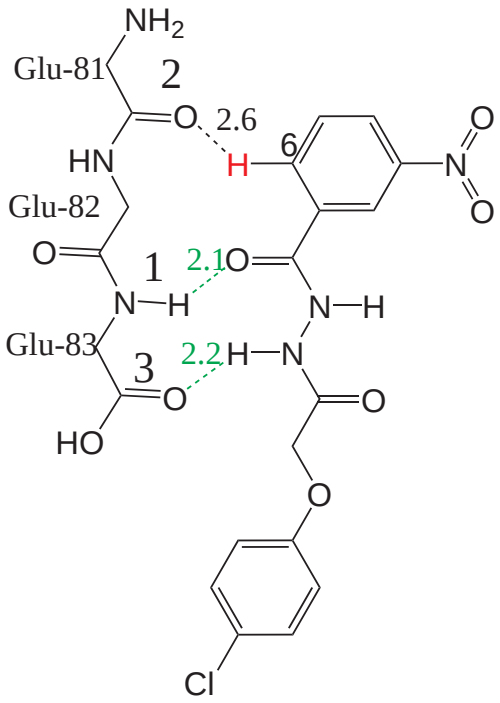
(e) Ligand in the structure of PDB entry 1KE5

Compound	ZINC ID	E_{ele} [a]	E_{vdW}	E_{ele}/MW	E_{vdW}/MW	Probe 1	Probe 2	Probe 3
27	476434	−1.6	−40.1	−0.005	−0.132	−2.38	−2.36	−3.65
28	43974	−4.0	−42.5	−0.012	−0.123	2.67	−3.43	−3.11
29	49031	−5.7	−35.5	−0.016	−0.103	−5.61	3.41	−7.22
30	55643	−0.9	−44.3	−0.003	−0.127	−4.63	−3.22	4.32
1KE5 Ligand		−5.0	−41.5	−0.015	−0.126	−3.94	−3.27	−1.63

(f) Probe energy of compounds.[b]

[a] Electrostatic interaction energy. [b] All energy values are in kcal·mol^{-1}. MW is in g·mol^{-1}.

Figure S-II: Assessment of QM probe method on CDK2 using four putatively inactive compounds (a)–(d), and a co-crystalized ligand (e). The distances between the critical atoms are noted with the digits above the dashed lines. The unit of length is Å. The green color denotes favorable HB interactions, the red indicates unfavorable interactions, and the black means favorable interactions but forming non-classical HBs.



31

Compound	ZINC ID	E_{ele}	E_{vdW}	E_{ele}/MW	E_{vdW}/MW	Probe 1	Probe 2	Probe 3
31	298885	-2.1	-39.5	-0.006	-0.113	-3.36	-2.02	-4.26

Figure S-III: Non-classical HB. All energy values are in kcal·mol^{-1}. The colors of the dashed lines and the digits have the same meanings as in the Figure S-II. The hydrogen atom in red is discussed in the main text.

without a $-NO_2$ group. Note that the experimental pK_a of nitrobenzene is 3.98 (at 0 °C).[19] Therefore, this hydrogen atom becomes a potential HB donor, and will form a HB when there is a HB acceptor nearby gaining a favorable interaction.

References

[1] L. Meijer, A. Borgne, O. Mulner, J. P. J. Chong, J. J. Blow, N. Inagaki, M. Inagaki, J. G. Delcros, J. P. Moulinoux, *Eur. J. Biochem.* **1997**, *243*, 527–536.

[2] S. H. Kim, *Pure Appl. Chem.* **1998**, *70*, 555–565.

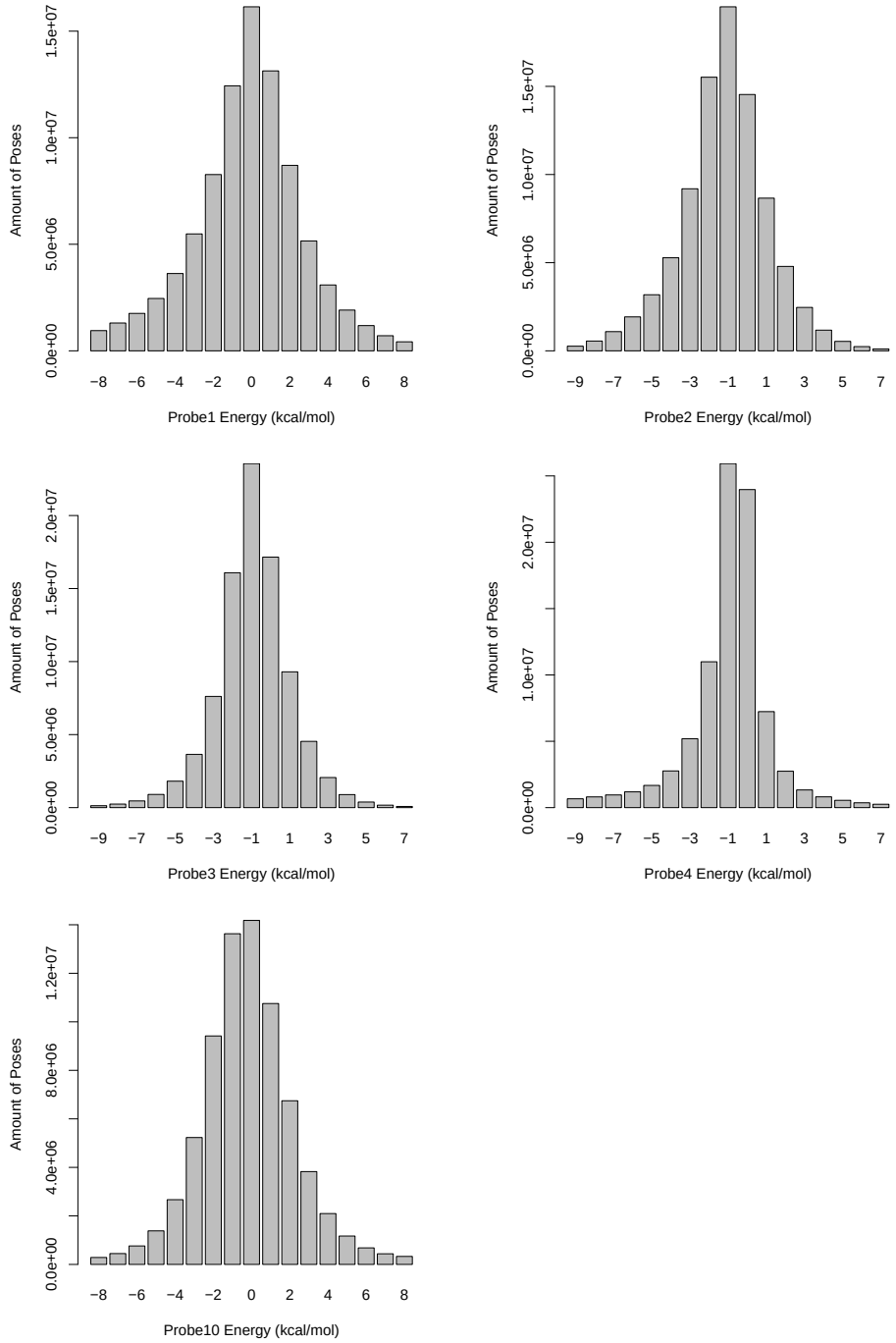
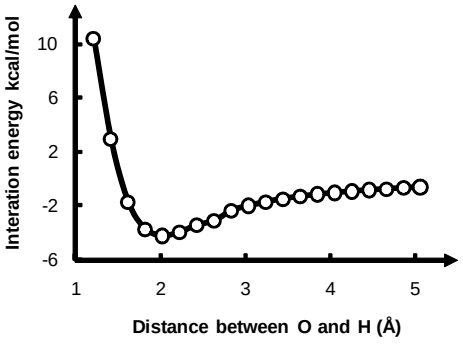
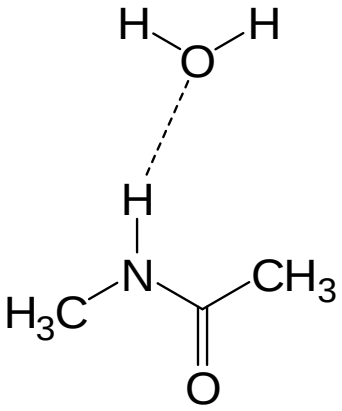


Figure S-IV: Distribution of energies of 5 probes (Probe 1–4, and 10) across 89,350,018 poses of neutral molecules.



(a)



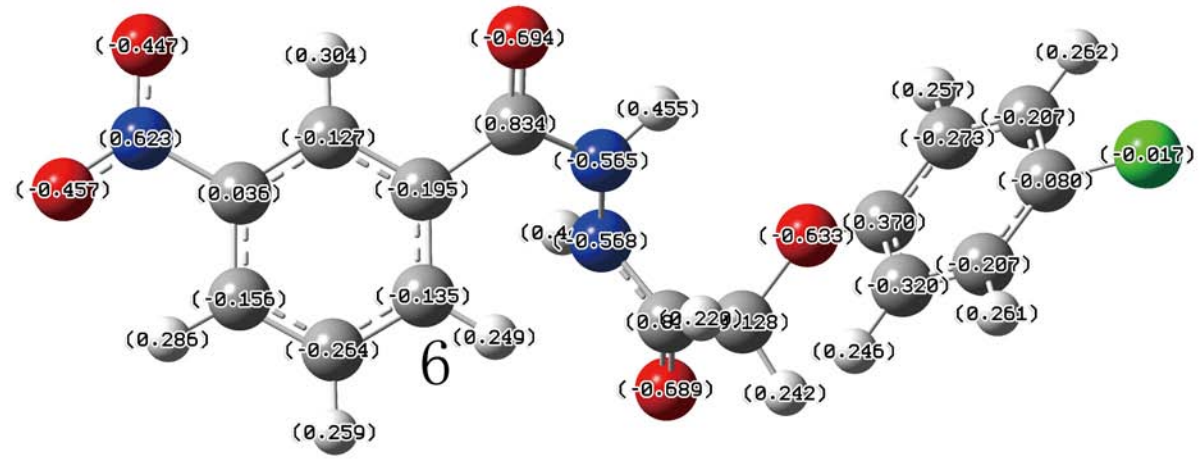
(b)

Distance(Å)	PM6	PM3	RM1	AM1
1.22	10.44	25.07	11.93	27.76
1.42	2.92	14.02	4.80	16.33
1.62	−1.74	1.92	0.38	5.69
1.83	−3.75	−1.97	−0.73	−1.09
2.03	−4.25	−1.18	−0.82	−3.53
2.23	−3.98	−1.02	−1.39	−3.73
2.43	−3.45	−1.45	−2.03	−3.24
2.64	−3.15	−1.52	−1.97	−2.74
2.84	−2.39	−1.41	−1.76	−2.43
3.04	−2.02	−1.26	−1.53	−1.70
3.24	−1.72	−1.11	−1.33	−1.39
3.45	−1.50	−0.97	−1.17	−1.17
3.65	−1.32	−0.86	−1.04	−1.02
3.85	−1.17	−0.77	−0.92	−0.90
4.06	−1.05	−0.69	−0.83	−0.81
4.26	−0.94	−0.62	−0.75	−0.73
4.46	−0.85	−0.56	−0.67	−0.66
4.66	−0.77	−0.51	−0.61	−0.60
4.87	−0.69	−0.46	−0.55	−0.55
5.07	−0.63	−0.42	−0.50	−0.50

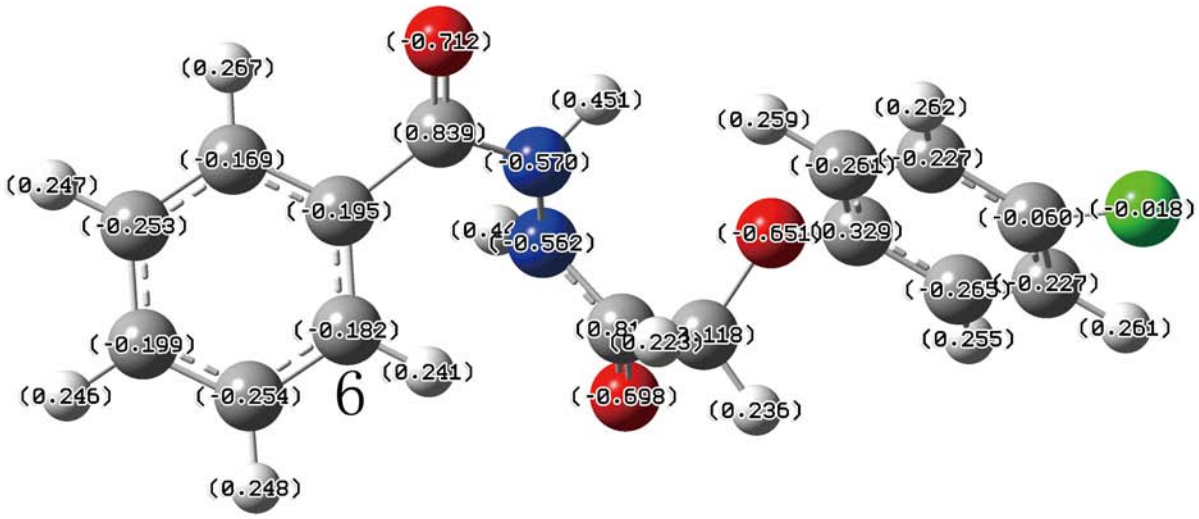
(c)

Figure S-V: (a) The PM6 interaction energies[a] between a water molecule and a N-methylacetamide are plotted against the distances from atom H to O connected with the dashed line in (b). The movement of the water molecule is along the vector defined by the atom H and O in the fully PM6-optimized conformation, i.e., the conformation when the distance equals to 2.03 Å in the second column of (c). After moving the water molecule to a new position, the conformation is partially optimized. The coordinates of atoms except for the H and the O atom are optimized using four Hamiltonians (PM6, PM3, RM1, and AM1) in MOPAC. The unit of all energy terms is kcal·mol^{-1}. The energy terms do not contain basis set superposition error correction.

[a] IE $= H_{\text{complex}} - H_{\text{water}} - H_{\text{N-methylacetamide}}$, where H is the formation enthalpy.



(1) Compound **31**



(2) Analogue of compound **31** without $-NO_2$ group

Figure S-VI: The quantum mechanical atomic charges of compound **31** and its analogue without $-NO_2$ group. The structures were minimized by B3LYP 6-31+G(d,p) starting with the docking structures. The digits in the parenthesis are the partial charges calculated by natural bond orbital theory at MP2 6-31+G(d,p) level. The group charge of C_6H changes from 0.059 electronic unit to 0.114 electronic unit, when a $-NO_2$ group substitutes its para hydrogen.

Table S-I: Probe energies of minor-different conformers of four known inhibitors of EphB4. The conformer in the first row of each block is the minimized conformer of each inhibitors which is identical to those listed in Table 2 of the main text. The other conformers are snapshots taken every 10 fs from a short molecular dynamics run (100 fs at 50 K) without minimization.

Conformer	Probe 1	Probe 2	Probe 3	Probe 4	Probe 10	RMSD of Coordinates(Å)
ALTA	−4.31	0.37	0.14	0.34	0.41	0.000
ALTA_1	−4.27	0.33	−0.04	0.33	0.34	0.025
ALTA_2	−4.32	0.34	0.20	0.35	0.34	0.026
ALTA_3	−4.16	0.39	0.05	0.30	0.43	0.030
ALTA_4	−4.18	0.34	0.25	0.36	0.38	0.031
ALTA_5	−4.24	0.34	0.08	0.30	0.40	0.035
ALTA_6	−4.39	0.40	0.05	0.37	0.43	0.031
ALTA_7	−4.18	0.38	−0.03	0.33	0.42	0.030
ALTA_8	−4.23	0.39	−0.03	0.35	0.41	0.031
ALTA_9	−4.16	0.39	0.31	0.35	0.46	0.034
ALTA_10	−4.12	0.35	−0.19	0.28	0.39	0.034
MIYA9f	−3.04	−2.25	−1.16	−5.53	−3.47	0.000
MIYA9f_1	−3.05	−2.17	−1.27	−5.49	−3.43	0.020
MIYA9f_2	−3.03	−2.26	−1.18	−5.45	−3.31	0.026
MIYA9f_3	−3.02	−2.39	−1.14	−5.64	−3.62	0.028
MIYA9f_4	−3.03	−1.96	−1.21	−5.27	−3.21	0.029
MIYA9f_5	−2.97	−2.25	−1.16	−5.50	−3.50	0.029
MIYA9f_6	−3.01	−2.22	−1.14	−5.25	−3.40	0.027
MIYA9f_7	−2.97	−2.09	−1.16	−5.35	−3.41	0.030
MIYA9f_8	−3.12	−2.15	−1.23	−5.52	−3.51	0.028
MIYA9f_9	−3.03	−2.03	−1.07	−5.15	−3.42	0.030
MIYA9f_10	−3.02	−1.99	−1.34	−5.23	−3.38	0.036
ONC102	−2.46	−2.43	−1.38	−0.22	−2.32	0.000
ONC102_1	−2.36	−2.40	−1.43	−0.20	−2.25	0.023
ONC102_2	−2.54	−2.43	−1.36	−0.22	−2.24	0.025
ONC102_3	−2.55	−2.49	−1.41	−0.20	−2.36	0.026
ONC102_4	−2.57	−2.54	−1.54	−0.22	−2.47	0.032
ONC102_5	−2.56	−2.34	−1.35	−0.20	−2.27	0.036
ONC102_6	−2.61	−2.38	−1.43	−0.21	−2.31	0.038
ONC102_7	−2.50	−2.33	−1.49	−0.21	−2.28	0.040
ONC102_8	−2.60	−2.33	−1.53	−0.20	−2.32	0.036
ONC102_9	−2.54	−2.50	−1.47	−0.22	−2.44	0.031
ONC102_10	−2.59	−2.51	−1.57	−0.21	−2.36	0.036
PP2	−3.71	−3.18	−1.24	0.12	−2.98	0.000
PP2_1	−3.73	−3.13	−1.34	0.12	−2.88	0.021
PP2_2	−3.62	−3.15	−1.26	0.12	−3.00	0.026
PP2_3	−3.72	−3.34	−1.24	0.11	−2.92	0.028
PP2_4	−3.63	−3.08	−1.23	0.15	−2.97	0.029
PP2_5	−3.63	−3.25	−1.09	0.14	−3.10	0.028
PP2_6	−3.57	−3.20	−1.21	0.09	−2.91	0.033
PP2_7	−3.71	−3.15	−1.13	0.15	−3.12	0.032
PP2_8	−3.65	−3.19	−1.17	0.10	−2.94	0.029
PP2_9	−3.60	−3.41	−1.06	0.13	−3.13	0.033
PP2_10	−3.65	−3.34	−1.14	0.08	−2.97	0.036

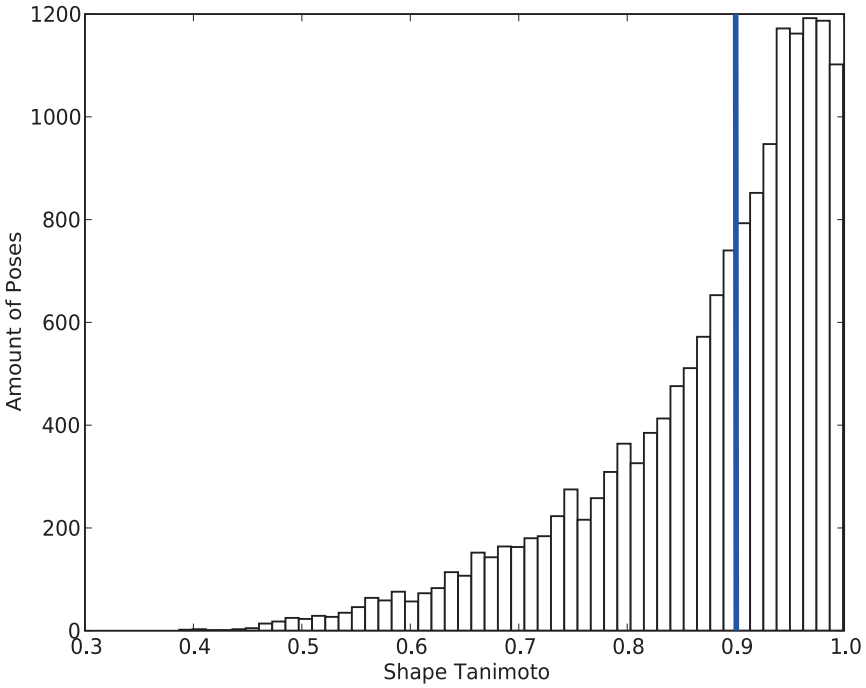


Figure S-VII: The distribution of shape Tanimoto of 15,979 poses. The blue vertical line at 0.9 emphasizes the threshold for filtering out the large-strain ligands.

Table S-II: List of 85 kinases for selectivity profile. The ATP concentration in each assay is denoted in the head row.

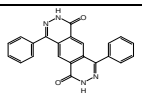
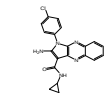
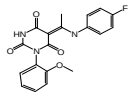
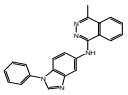
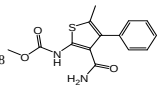
5μM*	20μM*	50μM*
ΔPH-PKBα (S473D)	Aurora B	ΔPH-PKBß (S474D)
CK2α	CaMKKß	AMPK
DYRK3	CDK2/cyclin A	BRSK2
EF2K	CHK1	BTK
EPH-B3	CHK2	CaMK1
ERK1	CK1δ	DYRK1a
ERK8	CSK	DYRK2
GSK3ß	FGF-R1	EPH-A2
HER4	GCK	IKKε
HIPK2	IR-HIS	LCK
IGF1R	IRAK4	MAPK2/ERK2
IKKß	JNK1α1	MAPKAP-K1a/RSK1
IRR	JNK2	MAPKAP-K1b/RSK2
MARK3	LKB1	MELK
MKK1	MAPKAP-K2	MINK1
p38γ MAPK	MLK1	MNK1
p38δ MAPK	MLK3	MNK2α
PAK4	MSK1	NEK2a
PIM2	MST2	NEK6
PKCζ	MST4	p38α MAPK
PLK1	NUAK1	PhKγ1
PRK2	p38ß MAPK	PKD1
	PAK5	smMLCK
	PAK6	Src
	PDK1	SRPK-1
	PIM1	TBK1
	PIM3	
	PKA	
	PKCα	
	PRAK	
	ROCKII	
	S6K1 (T412E)	
	SGK1	
	SYK	
	TTK	
	VEG-FR	
	YES1	

* The ATP concentrations are at or below the calculated Km for ATP for that kinase.

Table S-III: The sequence information and the gatekeeper residues of 85 kinases.

Protein Kinase	Accession No.	GI	Gatekeeper
AMPK[26-268]	NM_006252	46877068	M
Aurora B [1-344]	NM_004217	83776600	L
BRSK2 [2-674]	AF533878	33187742	L
BTK [2-659]	NP_000052.1	4557377	T
CaMK1 [2-369]	NM_003656	4502553	M
CaMKKβ [1-541]	NM_153499	27437017	F
CDK2[4-286]	NM_001798	16936528	F
CHK1 [1-476]	AF016582	2367669	L
CHK2 [5-543]	NM_007194	6005850	L
CK1δ [1-294]	AB063114	14422451	M
CK2α [2-391]	NM_001895	4503095	F
CSK [1-450]	NM_004383	4758078	T
DYRK1a [1-499]	NM_130437.2	18765754	F
DYRK2 [3-528]	NM_003583	4503427	F
DYRK3 [1-588]	AY590695	46909167	F
EF2K [2-725]	AAH32665	21618568	E
EPH-A2 [591-976]	NM_004431	32967311	T
EPH-B3 [561-998]	NM_004443	17975768	T
ERK1 [2-379]	BC013992	15559271	Q
ERK2 [1-358]	X58712	53002	Q
ERK8 [2-544]	AY065978	19263187	F
FGF-R1 [400-820]	M34641	182530	V
GCK [2 - 812]	BC047865	28839779	M
GSK3β [2-420]	L33801	529237	L
HER4 [706 - 991]	NM_005235	4885215	T
HIPK2 [165-564]	AF326592	17225377	F
IGF1R [954-1367]	NM_000875	4557665	M
IKKβ [1-736]	XM_032491	20538863	M
IKKε [1-716]	NM_014002	7661946	M
IR [1001-1382]	NM_000208.2	119395736	M
IRAK4 [140-460]	BC013316.1	15426432	Y
IRR [944-1236]	NM_014215	31657140	M
JNK1α1 [1-384]	L26318	474901	M
JNK2α2 [1-424]	L31951	598183	M
LCK [2-509]	X03533	244791455	T
LKB1 [1-433]	NP_000446	4507271	M
MAPKAP-K2 [46-400]	NM_032960	32481209	M
MARK3 [2-729]	U64205	3089349	M
MELK [2-651]	NM_014791	7661974	L
MINK1[1-320]	NM_015716	7657335	M
MKK1 [1-393]	Z30163	456202	M
MLCK [475-838]	NM_005965	16950601	L
MLK1 [132 - 413]	NM_033141	52421790	M
MLK3 [96 - 386]	NM_002419	4505195	M
MNK1 [2-424]	AB000409	2077825	F
MNK2α [2-465]	AF237775	11023170	F
MSK1 [2-802]	AF074393	3411157	L
MST2 [2-491]	U60206	1477789	M
MST4 [1-416]	NM_016542	15011880	M
NEK2A [1-445]	NM_002497	4505373	M
NEK6 [8-313]	NM_014397	19923407	L
NUAK1 [2-660]	NM_014840	7662170	M
p38β MAPK [1-364]	Y14440	2326554	T
p38α MAPK [1-360]	L35264	603919	T
p38γ MAPK [1-367]	Y10487	1785656	T
p38δ MAPK [1-365]	Y10488	2266640	M
PAK4 [2-591]	O96013	12585288	M
PAK5 [2-719]	Q9P286	12585290	M
PAK6 [2-681]	Q9NQU5	23396789	M
PDK1 [52-556]	NM_002613	4505695	L
PhKγ1[2-297]	X80590	1147567	F
PIM1 [2-313]	NM_002648	4505811	L
PIM2 [2-334]	U77735	1750276	L
PIM3 [2-326]	Q86V86	215274221	L
PKA [2-351]	NM_002730	4506055	M
PKBβ (S474D) [120-481]	NM_001626	4502023	M
PKBα (S473D) [118-480]	BC000479	12653417	M
PKCα [1-672]	NM_002737	4506067	M
PKCζ [2-592]	NM_002744	52486327	I
PKD1 [2-912]	NM_002742	115529463	M
PLK1 [1-603]	NM_005030	21359873	L
PRAK [1-471]	AF032437	3133291	M
PRK2 [501-984]	S75548	914100	M
ROCKII [2-543]	U38481	1384133	M
RSK1 [1-735]	M99169	206772	L
RSK2 [2-740]	NM_004586	4759050	T
S6K1 (T412E) [1-421]	NM_003161	4506737	L
SGK1 (S422D) [60-431]	NM_005627	25168263	L
Src [2-533]	NM_005417.3	4885609	T
SRPK1 [2-654]	NM_003137	47419936	F
SYK [1-635]	AAH01645.1	12804475	M
TBK1 [1-729]	NM_013254	7019547	M
TTK [1 - 857]	NM_003318	23308722	M
VEGFR [784-1338]	NM_002019.3	156104876	V
YES1 [1-543]	NM_005433	4885661	T

Table S-IV: Energy values in kcal·mol^{-1} of experimentally tested compounds[a].

ZINC ID	Structure	Probe1	Probe2	Probe3	Probe4	Probe5	Probe6	Probe7	Probe8	Probe9	Probe10	Probe11	Probe12	Probe13	hydrophobic matching	% inihibition at 50 μM
842896		−3.01	−2.51	0.21	−0.37	−0.03	0.12	0.28	0.12	−0.05	−0.97	−6.76	−8.52	−0.33	−9.27	19, 8
2361207		−4.49	−2.08	0.25	0.09	0.03	−0.04	0.40	0.11	−0.22	−2.90	−5.82	−3.78	−0.61	−7.96	~0
1406465		−3.84	−1.25	−2.25	−1.38	−0.15	0.07	0.29	0.06	0.05	−0.46	−3.96	−2.98	0.14	−5.02	0, 54
1213337		−2.94	−1.08	0.43	−4.04	−0.01	−0.08	0.43	0.25	0.06	−1.69	−11.19	2.57	−0.29	−8.19	42, 70
1053478		−3.90	−2.34	−2.03	−1.05	0.10	−0.08	0.35	0.18	−0.23	−1.17	−7.04	0.61	0.23	−5.65	-NA-
838240		−3.16	−1.46	−0.27	−1.90	−0.37	−0.14	0.24	0.13	−0.07	−2.05	−7.22	−2.08	−0.33	−5.43	-NA-

[a] We do not have energy values of the 23 compounds mentioned in the main text because most of them are derivatives of ZINC compounds, since the original compounds were not available. The compounds **1** to **26** in the main text are derivatives of ZINC compound 1053478.

[3] B. B. McConnell, F. J. Gregory, F. J. Stott, E. Hara, G. Peters, *Mol. Cell. Biol.* **1999**, *19*, 1981–1989.

[4] T. G. Davies, P. Tunnah, L. Meijer, D. Marko, G. Eisenbrand, J. A. Endicott, M. E. M. Noble, *Structure* **2001**, *9*, 389–397.

[5] M. Nesi, D. Borghi, M. G. Brasca, F. Florentini, P. Pevarello, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3205–3208.

[6] C. M. Richardson, C. L. Nunns, D. S. Williamson, M. J. Parratt, P. Dokurno, R. Howes, J. Borgognoni, M. J. Drysdale, H. Finch, R. E. Hubbard, P. S. Jackson, P. Kierstan, G. Lentzen, J. D. Moore, J. B. Murray, H. Simmonite, A. E. Surgenor, C. J. Torrance, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 3880–3885.

[7] P. Kolb, D. Huang, F. Dey, A. Caflisch, *J. Med. Chem.* **2008**, *51*, 1179–1188.

[8] Y. H. Jiang, R. C. H. Zhao, C. M. Verfaillie, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 10538–10543.

[9] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, *J. Comput. Chem.* **2009**, *30*, 1545–1614.

[10] F. Momany, R. Rone, *J. Comput. Chem.* **1992**, *13*, 888–900.

[11] J. J. Liao, *J. Med. Chem.* **2007**, *50*, 409–424.

[12] D. S. Goodsell, A. J. Olson, *Proteins* **1990**, *8*, 195–202.

[13] T. Zhou, D. Huang, A. Caflisch, *J. Med. Chem.* **2008**, *51*, 4280–4288.

[14] K. Raha, K. M. Merz, *J. Am. Chem. Soc.* **2004**, *126*, 1020–1021.

[15] J. Sponer, J. Leszczynski, P. Hobza, *Biopolymers* **2001**, *61*, 3–31.

[16] S. Sarkhel, G. R. Desiraju, *Proteins* **2004**, *54*, 247–259.

[17] A. C. Pierce, K. L. Sandretto, G. W. Bemis, *Proteins* **2002**, *49*, 567–576.

[18] B. Klaholz, D. Moras, *Structure* **2002**, *10*, 1197–1204.

[19] E. P. Serjeant, B. Dempsey, *Ionization Constants of Organic Acids in Aqueous Solution*, Pergamon, Oxford, **1979**.