Supporting Information

Oruganty et al. 10.1073/pnas.1207104110

SI Materials and Methods

The Aurora wild-type and mutants D311l/T288E and D311N/T288E were purified as described in *Materials and Methods*. All of the proteins were concentrated to contain 1 mg/mL of the mutant/wild-type and Omnia peptide kit 16 (Invitrogen) was used for the assay. The Omnia kit is a sox-based chelation-enhanced florescence assay that measures the fluorescence intensity of the phosphorylated peptide (Ex 360 nm/Em 485 nm). A synergy H4 microplate reader in the monochromator configuration was used to monitor the reaction. Each well had a final concentration of 50-μM sox-based peptide, 1 mM ATP, and 1 mM DTT, along with a proprietary buffer formulation premixed into a master mix. The reaction as initiated by adding 30 μL master mix to 10 μL of the protein in the well, and fluorescence measurements were taken at 1-min intervals for 60 min at 30 °C. The reactions were carried out in duplicate from two different purification attempts with similar results.

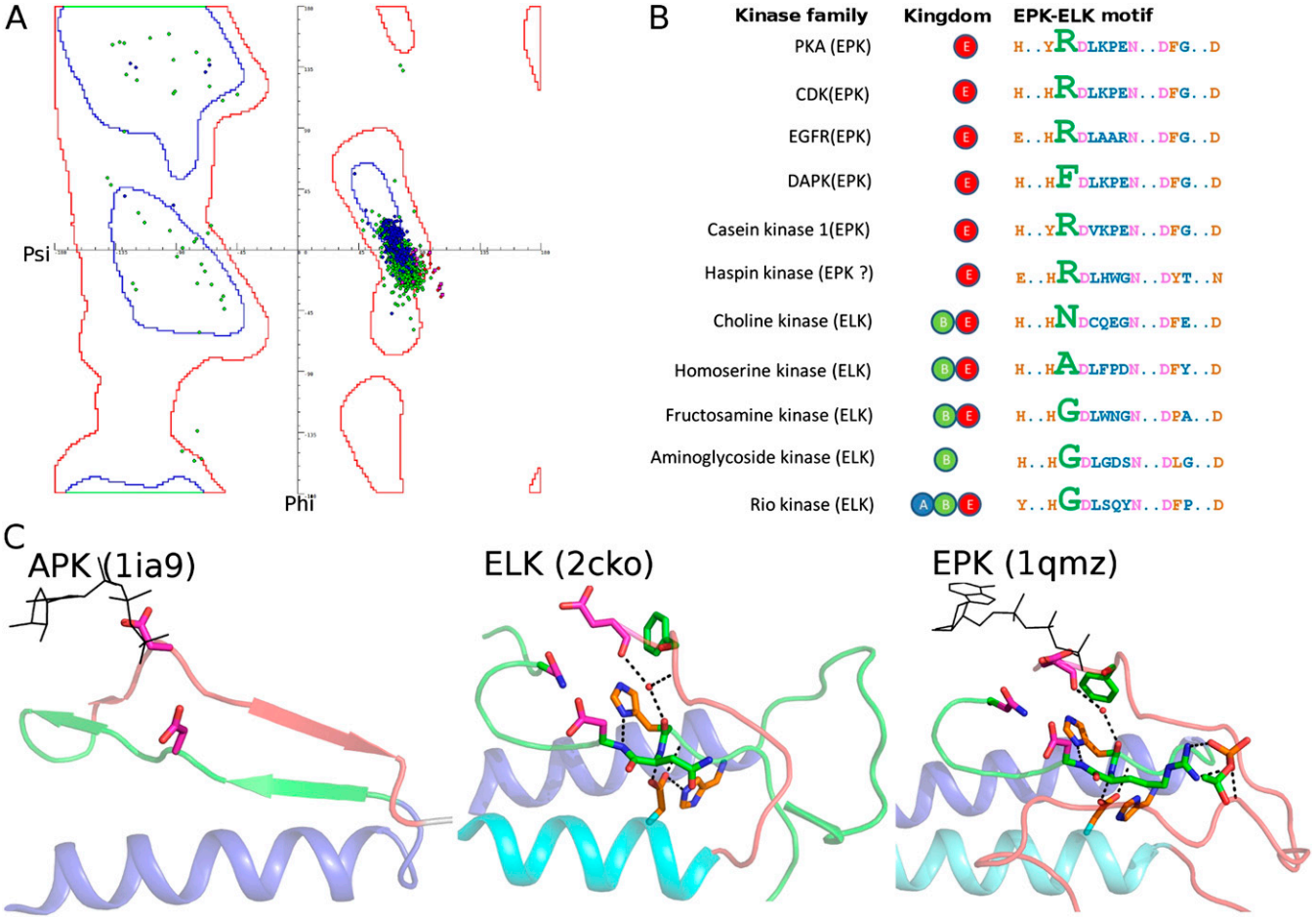

Fig. S1. Figure showing the evolution of "strain" with eukaryotic protein kinase/eukaryotic-like kinase (EPK-ELK) component. (*A*) Ramachandran map of arginine residues in ~1,400 protein kinase (EPK) and ELK structures, with resolution less than 2.0 Å . The panel shows the ϕ/ψ values for the HRD-Arg position in EPKs (circles, 1,468 structures) and ELKs (squares, 12 structures). The EPK residues are colored green if the HRD-Arg position is occupied by Arginine (1,368 structures), and are colored red if a non-Arginine residue occupies that position (100 structures). The ELK residues with glycine at the HRD-Arg position are shown as magenta squares (12 structures). The structures with phosphorylated activation loop are shown in blue (189 structures). (*B*) The EPK-ELK component residues in multiple EPK and ELK families are shown. The residues are colored according to the scheme given in Fig. 1. For clarity, the residue at the HRD-arginine position in EPKs is shown with a larger font size. The presence of each kinase in Archea, Bacteria, and Eukaryota is indicated by colored circles with A, B, or E. (*C*) The active site organization in kinases with the catalytic residues colored pink and ATP colored black. The catalytic loop is shown in green and the DFG containing loop is shown in red. The E-helix is colored blue and the F-helix is colored cyan. The APKs, such as Trp-Ca channel kinase (PDB ID1IA9) shown here do not have an F-helix (*Left*); (*Center*) an ELK, Choline kinase (PDB ID 2CKO); (*Right*) an EPK, CDK (PDB ID 1QMZ). The figure shows the EPK-ELK component residues in colors according to scheme in Fig. 1. The figure was generated using PyMOL.

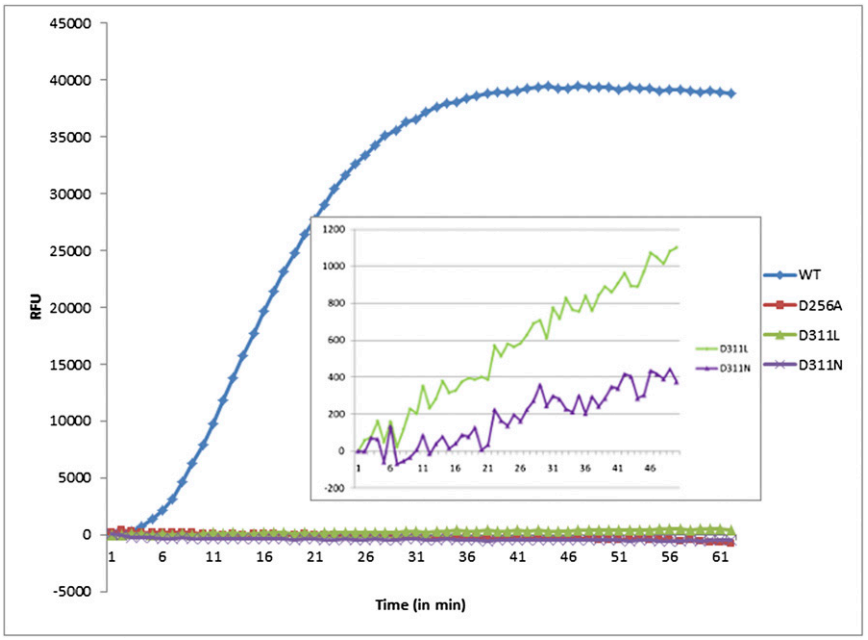

Fig. S2. The rate of reaction probed using the Omnia Aurora kinase kit from Invitrogen. The fluorescence intensity increases upon phosphorylation and this is given as relative fluorescence units (RFU). The time course of the reaction is plotted for 60 min, with datapoints taken every 1 min. The substrate peptide concentration was 50 μM, ATP concentration was 10 mM, and each well had 5 μg of either wild-type protein or mutant protein. A zoomed view of only the mutants D311L/T288E (shown as D311L) and D311N/T288E (shown as D311N) is shown (*Inset*) to show the difference in rates between the two mutants (the RFU value difference between the mutants and kinase dead mutant, D256A is plotted). The rates in terms of RFU/min are ~2,000/min for wild-type, 43.75/min for the D311L/T288E mutant, and 12.5/min for the D311N/T288E mutant. The T288E mutant by itself was also characterized. The T288E mutant showed a slope of activity around 0.8-times of the wild-type without a lag phase.

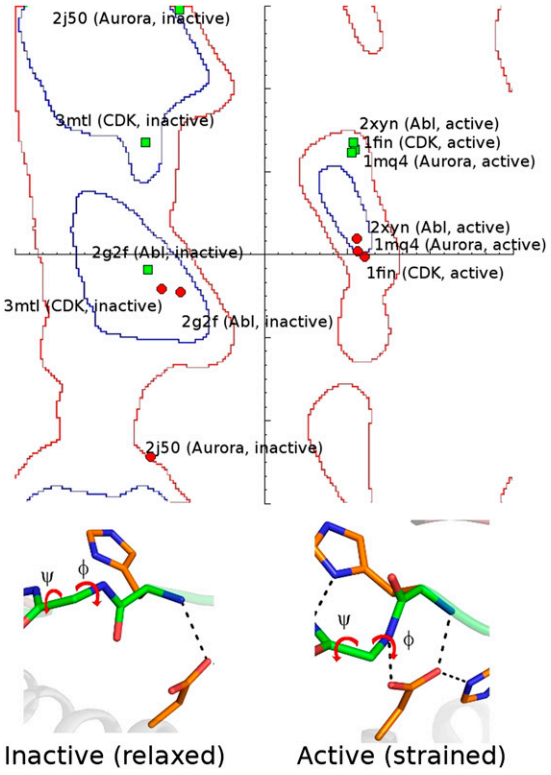

Fig. S3. The correlated changes occurring in the DFG-aspartate and HRD-arginine residues occurring when the EPK switches from an active DFG-in conformation to an inactive DFG-out conformation in three families of kinases. The HRD-arginine (Arg) residues are indicated by red circles and the DFG-aspartate (Asp) are indicated by green squares. The PDB structure and the active/inactive state of the kinase is indicated on the Ramachandran map for each point. (*Lower*) the two structures, one with a strained HRD-Arg and DFG-in conformation (PDB ID 1FIN, CDK), and the other with DFG-out conformation and a relaxed HRD-Arg conformation (PDB ID 2G2F, Abl kinase).

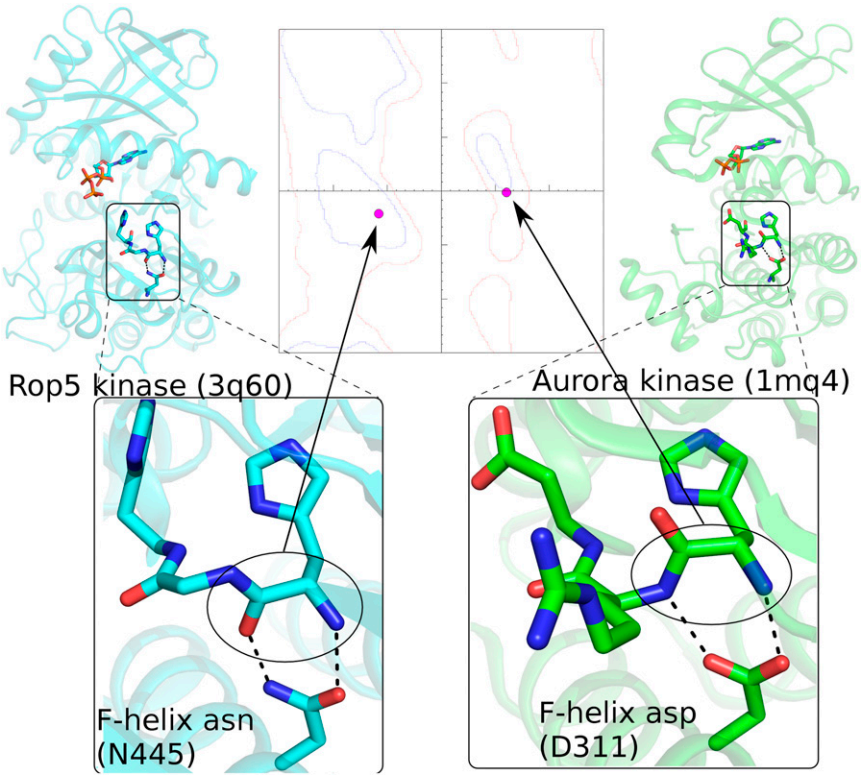

Fig. S4. Figure showing strain status of the HRD-Arg residue in ROP5 pseudokinase (PDB ID 3Q60) and active Aurora A kinase (PDB ID 1MQ4). A close-up of the active site is shown for clarity and the arrows point to the region of Ramachandran map occupied by the residue at HRD-Arg position (glycine in ROP5). Note the similarity of the ϕ/ψ values of ROP5 to inactive kinase conformations shown in Fig. 5. A similar conformational change (peptide flip) could occur in the Aurora D311N mutant, making it inactive.

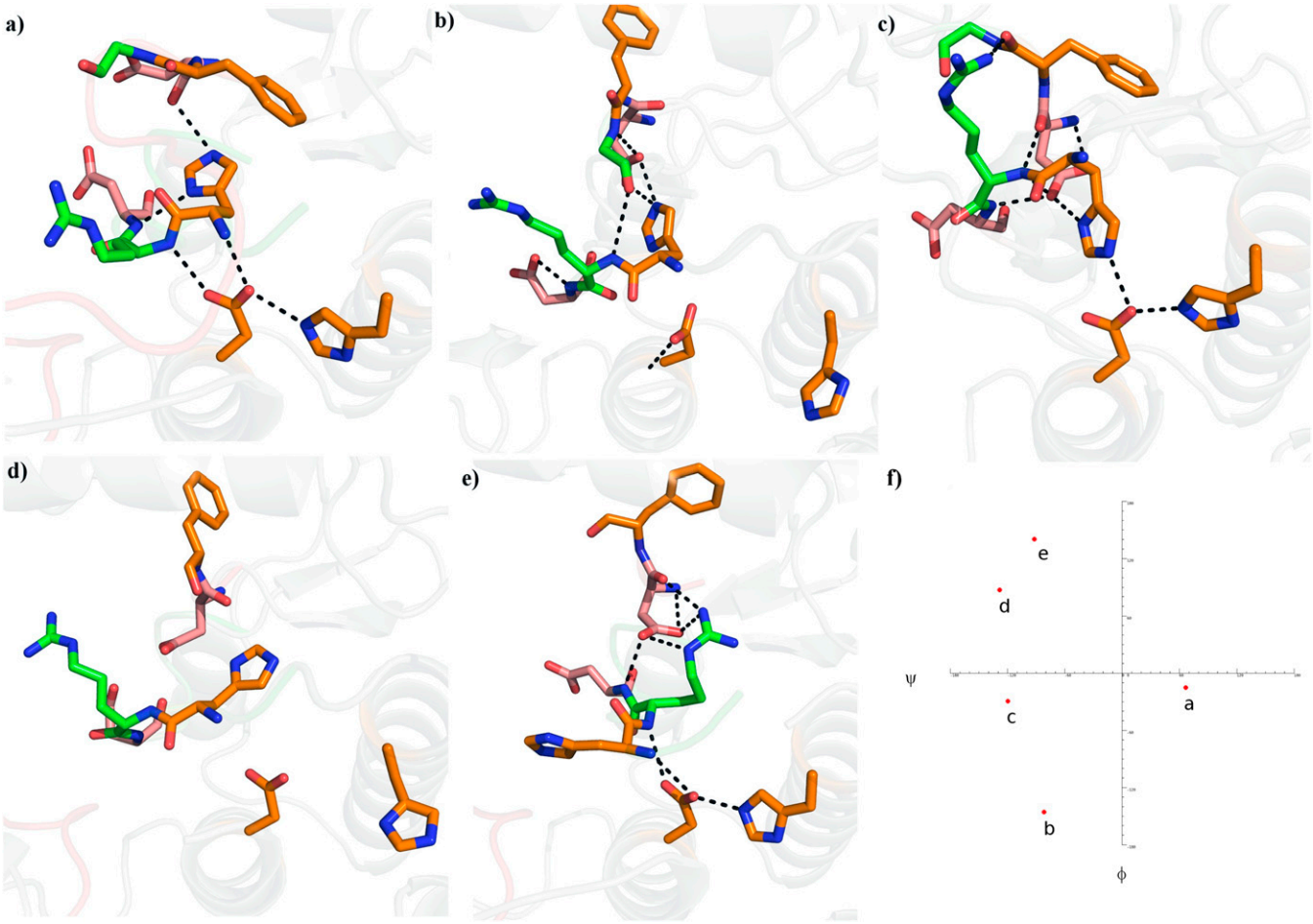

Fig. S5. Inactive structures in the transition from strain present state to strain lost state. (*A*) The Aurora kinase is shown in an active conformation with the strain intact. (*B*) Destabilization of the DFG loop is correlated with loss of strain in the HRD motif. The DFG motif in these panels is still in a canonical arrangement. (*C*) The DFG motif is seen "flipping up" in a noncanonical fashion. The HRD-Arg has lost the strain, but the catalytic loop is not perturbed. (*D*) The catalytic loop undergoes a dramatic change with the F-helix-Asp binding to side-chain of HRD-His. The catalytic loop HRD-Arg hydrogen bonds to the C-helix-Glu. (*E*) The rotation of the catalytic loop is complete and F-helix-Asp is seen to be bonding to backbone of HRD motif, although HRD-Arg is not under strain. (*F*) The ϕ/ψ values for each of the structures in the preceding panels. The PDB structure IDs used for generating this figure in order are: 3MYG, 2J50, 3DJ6, 2C6D, and 3LAU.

Table S1. List of high-resolution kinases (resolution <1.7 Å), which were used to study the strain in the catalytic loop

PDB ID	HRD-Arg residue	ϕ angle	ψ angle	N-Ca-C angle (dev)	Subclass/class
3hmi_A	Arg	76.2	−2.6	117.87 (2.43)	Abl/TK
3kfa_A	Arg	82.1	−13.2	115.37 (1.43)	Abl/TK
3kfa_B	Arg	81.9	−13	115.60 (1.52)	Abl/TK
1o6l_A	**Arg**	**62.9**	**−5.6**	**121.76 (3.99)**	**Akt/AGC**
2wei_A	Arg	80	−30.5	111.84 (0.02)	CAMK1/CAMK
3ot3_A	Arg	74.8	−4.2	118.09 (2.51)	CAMKL/CAMK
3ot8_A	Arg	74.2	−4.3	117.11 (2.12)	CAMKL/CAMK
3pa3_A	Arg	74.4	−1.4	117.20 (2.16)	CAMKL/CAMK
3pa4_A	Arg	73.5	−3	118.24 (2.57)	CAMKL/CAMK
1gz8_A	Arg	79.5	−11.9	116.60 (1.92)	CDK/CMGC
1h00_A	Arg	80.5	−19	117.61 (2.33)	CDK/CMGC
1jvp_P	Arg	82.7	−21.8	116.25 (1.78)	CDK/CMGC
1oit_A	Arg	83.2	−16.9	115.95 (1.66)	CDK/CMGC
1urw_A	Arg	83.5	−20.3	116.97 (2.07)	CDK/CMGC
2r3f_A	Arg	79.8	−15.2	118.19 (2.56)	CDK/CMGC
2r3g_A	Arg	83.2	−20.6	117.34 (2.22)	CDK/CMGC
2r3h_A	Arg	85	−19.1	117.09 (2.12)	CDK/CMGC
2r3i_A	Arg	82.1	−15.3	116.29 (1.80)	CDK/CMGC
2r3j_A	Arg	83.6	−20.3	117.44 (2.25)	CDK/CMGC
2r3l_A	Arg	84	−21	116.63 (1.93)	CDK/CMGC
2r3n_A	Arg	76.8	−22.3	117.55 (2.30)	CDK/CMGC
2r3p_A	Arg	81.2	−14.4	117.44 (2.25)	CDK/CMGC
2r3q_A	Arg	81.2	−13.5	115.95 (1.66)	CDK/CMGC
2r3r_A	Arg	78.1	−12.5	117.43 (2.25)	CDK/CMGC
2vtt_A	Arg	81.5	−27.2	117.41 (2.25)	CDK/CMGC
2izr_A	Arg	71	13.5	115.49 (1.47)	CK1/CK1
1om1_A	Arg	65.5	13.7	118.18 (2.55)	CK2/CMGC
2pvr_A	**Arg**	**70.2**	**6.4**	**115.84 (1.62)**	**CK2/CMGC**
3bqc_A	Arg	73.5	8	115.60 (1.52)	CK2/CMGC
3h30_A	Arg	70.9	12.2	115.71 (1.56)	CK2/CMGC
3h30_B	Arg	70.7	6.3	113.83 (0.81)	CK2/CMGC
3juh_A	**Arg**	**64.1**	**21.5**	**111.05 (−0.30)**	**CK2/CMGC**
3juh_B	**Arg**	**78.8**	**−1.1**	**115.48 (1.47)**	**CK2/CMGC**
3mb7_A	Arg	69	7.1	115.85 (1.62)	CK2/CMGC
3nsz_A	**Arg**	**68.1**	**13**	**115.16 (1.34)**	**CK2/CMGC**
3pvg_A	Arg	68.5	16.8	114.89 (1.24)	CK2/CMGC
2eu9_A	Thr	77.5	−9.4	115.78 (1.59)	CLK/CMGC
1jkl_A	**Phe**	**66.5**	**4.2**	**118.01 (2.49)**	**DAPK/CAMK**
1jks_A	Phe	66	5	117.56 (2.31)	DAPK/CAMK
2a2a_A	Phe	68	6.7	115.65 (1.54)	DAPK/CAMK
2a2a_B	Phe	70.3	4.7	115.56 (1.51)	DAPK/CAMK
2a2a_C	Phe	64.7	10.3	116.29 (1.79)	DAPK/CAMK
2a2a_D	Phe	69.7	10.1	115.66 (1.54)	DAPK/CAMK
2w4j_A	Phe	66.7	11.5	116.06 (1.71)	DAPK/CAMK
3bhy_A	Phe	71	1.5	116.12 (1.73)	DAPK/CAMK
3eha_A	**Phe**	**68.2**	**6.1**	**115.63 (1.53)**	**DAPK/CAMK**
3gu4_A	**Phe**	**70.1**	**2.9**	**116.54 (1.90)**	**DAPK/CAMK**
3gu5_A	**Phe**	**67.4**	**6.2**	**116.32 (1.81)**	**DAPK/CAMK**
3gu6_A	Phe	70.7	4.8	114.85 (1.22)	DAPK/CAMK
3gu8_A	Phe	65.9	4.2	117.02 (2.09)	DAPK/CAMK
2qo2_A	Arg	75	−15.1	115.59 (1.51)	Eph/TK
2qo7_A	**Arg**	**73.8**	**−12.7**	**115.76 (1.58)**	**Eph/TK**
2qo9_A	**Arg**	**73.2**	**−10.7**	**115.70 (1.56)**	**Eph/TK**
2qob_A	Arg	72.9	−7.7	116.41 (1.84)	Eph/TK
2qoc_A	Arg	74	−8.7	115.78 (1.59)	Eph/TK
2qod_A	Arg	75	−14.4	115.56 (1.50)	Eph/TK
2qof_A	Arg	75.6	−15.2	116.42 (1.85)	Eph/TK
2qoi_A	Arg	76.1	−13.8	115.80 (1.60)	Eph/TK
2qok_A	Arg	73.7	−11.2	115.66 (1.54)	Eph/TK
2qol_A	Arg	74.8	−14.1	116.13 (1.73)	Eph/TK
2qoo_A	Arg	72	−7	116.51 (1.88)	Eph/TK
2qoq_A	**Arg**	**73.7**	**−6.7**	**115.75 (1.58)**	**Eph/TK**

Table S1. Cont.

PDB ID	HRD-Arg residue	ϕ angle	ψ angle	N-Ca-C angle (dev)	Subclass/class
2rei_A	Arg	79.7	−27.5	115.77 (1.59)	Eph/TK
2vwx_A	Arg	76.5	−14	116.86 (2.02)	Eph/TK
2vwy_A	Arg	79.5	−13.1	115.44 (1.45)	Eph/TK
2vwz_A	Arg	76.9	−12.6	115.08 (1.31)	Eph/TK
2vx1_A	Arg	78	−12.8	116.03 (1.69)	Eph/TK
1mp8_A	Arg	70.4	−1.3	117.28 (2.19)	Fak/TK
2jko_A	Arg	77.3	−4.3	115.14 (1.34)	Fak/TK
3cc6_A	Arg	76.9	−2	116.26 (1.78)	Fak/TK
1p4o_A	Arg	81.1	−18.9	113.59 (0.72)	InsR/TK
1p4o_B	Arg	81.7	−17.5	113.97 (0.87)	InsR/TK
3bu3_A	Arg	74	−9.2	117.10 (2.12)	InsR/TK
3lxp_A	Arg	74.5	2.8	115.55 (1.50)	Jak/TK
2b9h_A	Arg	74.1	−6.9	116.34 (1.82)	MAPK/CMGC
2fst_X	Arg	81.6	−6.3	113.85 (0.82)	MAPK/CMGC
3dkc_A	Arg	82.5	−22.8	115.88 (1.63)	Met/TK
3f66_A	Arg	78.3	−14.2	116.01 (1.68)	Met/TK
3f66_B	Arg	82.1	−19.9	116.10 (1.72)	Met/TK
2w5a_A	Arg	−137	151.5	111.02 (−0.31)	NEK/Other
3ork_A	Arg	82.9	−12.9	115.55 (1.50)	NEK/Other
1t46_A	Arg	82.1	−17.1	116.27 (1.79)	PDGFR/TK
3g0e_A	Arg	86.5	−14.1	115.47 (1.47)	PDGFR/TK
3a99_A	Arg	66.1	8.3	117.34 (2.22)	PIM/CAMK
1rdq_E	**Arg**	**73.2**	**12.4**	**116.47 (1.87)**	**PKA/AGC**
1xh8_A	Arg	69.2	1.7	118.30 (2.60)	PKA/AGC
1xh9_A	Arg	70.3	10.2	118.08 (2.51)	PKA/AGC
3fjq_E	**Arg**	**67.4**	**13.9**	**119.23 (2.97)**	**PKA/AGC**
3idb_A	**Arg**	**69.2**	**11.1**	**123.62 (4.73)**	**PKA/AGC**
1fmk_A	Arg	83.9	−11.4	115.59 (1.52)	Src/TK
1qpc_A	**Arg**	**74.3**	**−4.4**	**117.71 (2.37)**	**Src/TK**
2src_A	Arg	81.6	−7.9	117.23 (2.17)	Src/TK
2c30_A	Arg	73.2	0.7	116.49 (1.88)	STE20/STE
2j0i_A	Arg	71.1	2.7	117.57 (2.31)	STE20/STE
3a7f_A	Arg	75	−7.9	117.57 (2.31)	STE20/STE
3a7i_A	Arg	70.8	−0.7	118.49 (2.68)	STE20/STE
3a7j_A	Arg	67.3	−1.1	117.93 (2.45)	STE20/STE
3fxz_A	Arg	75.9	−1.5	117.77 (2.39)	STE20/STE
1xbb_A	Arg	74.7	4.6	114.81 (1.20)	Syk/TK
3gen_A	Arg	78.4	−14.2	116.18 (1.75)	Tec/TK
3miy_A	Arg	80.9	−14.9	114.79 (1.19)	Tec/TK
3miy_B	Arg	80.1	−14.3	115.78 (1.59)	Tec/TK
3ewh_A	Arg	78.1	−10.7	116.17 (1.75)	VEGFR/TK

The N-Ca-C angle is given and in parentheses, its deviation from ideal angle is given as SDs above mean. The subclass/class classification was taken from the Kinbase database. The "active" conformations according to definition given *Materials and Methods* are in boldface.

Table S2. Contingency tables for the features studied in this article

Relax/strain	Feature occurring	Feature not occurring	Column total
DFG-in	DFG-in (R-spine formed)	DFG-out (R-spine not formed)	
Relax	157	16	173
Strain	2,145	124	2269
Row total	2,302	140	2442
F-helix-Asp	F-helix-Asp makes 2 canonical hbonds with HRD backbone	F-helix-Asp does not make 2 canonical hbonds with HRD backbone	
Relax	0	173	173
Strain	2,198	71	2269
Row total	2,243	199	2442
HRD-His	HRD-His makes 2 H bonds with HRD and DFG backbone	HRD-His not making 2 H bonds with HRD and DFG backbone	
Relax	45	128	173
Strain	1,357	912	2269
Row total	1,402	1040	2442
EPK-ELK	Four canonical EPK-ELK component H bonds present	Four canonical EPK-ELK component H bonds not present	
Relax	0	173	173
Strain	1,357	912	2269
Row total	1,357	1,085	2442
E-helix-His	E-helix-His interacting with F-helix aspartate	E-helix-His not interacting with F-helix aspartate	
Relax	4	169	173
Strain	1,113	1,156	2269
Row total	1,117	1,325	2442

In each case, the feature studied is compared against the presence/loss of strain (strain and relax, respectively, in the table). The row and column totals are given for each of the contingency tables.

Table S3. Statistics of the strained and relaxed form HRD-Arg containing structures with various ligands

Structure	Strain (2,269)	Relaxed (173)
ATP*	63	7
ANP†	127	9
ADP	64	5
Lys-Glu	1,606	73
TPO/SEP/PTR	652	66
TPO/SEP/PTR + ATP/ANP	69	6
Lys-Glu + ATP/ANP	119	3
Lys-Glu + ADP	47	3
Lys-Glu + ATP/ANP + TPO/SEP/PTR	50	0

Note that there are structures from different resolutions in the statistics and the list includes mutant forms. Manual assessment of each of the cases in the relaxed form was carried out along with analysis of electron density visualization to assess the significance of the strain status for activity.

*The relaxed forms consist of PDB ID 1UA2 (resolution 3.02), PDB ID 1S9I (ATP and inhibitor-bound structure), and PDB ID 3KMW (pseudoactive site of ILK in inactive conformation).

†The relaxed forms consist of one chain of PDB ID 2A19 (chain C), which shows a relaxed form, whereas chain B shows a strained form. PDB ID 3COK is an unusual inactive conformation of PDB ID PLK4, even when bound to ANP. PDB ID 1AD5 is an Hck structure with similarities to inactive CDK structure. PDB ID 2EB3, EGFR L858R mutant (resolution 2.84), although a higher resolution structure (PDB ID 2ITV) shows HRD-Arg in strained conformation. PDB ID 2C6D is an unusual Aurora A inactive conformation bound to ANP.