Table S1: List of catalytically primed structures

Uniprot	PDB	Chain	Reso	X_Phi	X_Psi	Asp_Phi	Asp_Psi	Asp_Chi1	Asp_Chi2	Phe_Phi	Phe_Psi	Phe_Chi1	Ligand
AKT2	1O6K	A	1.7	-120.8	-176.0	56.8	82.9	186.7	12.6	-98.9	24.5	299.4	ANP
AKT2	1O6L	A	1.6	-121.5	-178.2	58.4	83.3	187.4	7.0	-98.8	24.7	297.6	ANP
AURKA	5DN3	A	2.0	-130.9	-174.4	55.3	74.8	189.2	0.9	-89.0	14.2	275.5	ATP
AURKA	5DNR	A	1.9	-132.2	-179.7	60.8	76.7	188.4	2.7	-90.1	18.7	275.9	ATP
CDK2	1QMZ	A	2.2	-146.3	-172.6	47.5	78.4	192.7	-2.0	-93.9	24.3	280.2	ATP
CDK2	1QMZ	C	2.2	-147.5	-169.9	41.1	80.5	192.6	-4.4	-95.5	25.6	283.1	ATP
CDK2	4EOJ	A	1.6	-134.1	177.6	53.9	78.1	191.5	9.6	-94.3	17.8	277.9	ATP
CDK2	4EOM	A	2.1	-135.9	-180.0	56.8	76.4	194.3	-4.5	-93.7	18.3	281.7	ATP
CDK2	4EOM	C	2.1	-134.3	179.5	56.9	73.6	194.6	8.9	-89.3	27.1	288.3	ATP
CDK2	4EOO	A	2.1	-137.3	-178.3	51.9	74.5	189.5	-2.6	-88.8	23.7	283.6	ATP
CDK2	4EOO	C	2.1	-132.9	-176.4	50.3	74.1	194.7	1.1	-84.3	27.4	283.8	ATP
CDK2	4EOQ	A	2.1	-136.2	-179.0	55.1	77.3	197.2	1.6	-91.4	19.3	282.3	ATP
FGFR2	2PVF	A	1.8	-138.9	-175.2	54.7	80.5	190.7	-12.4	-90.1	24.3	291.1	ACP
FGFR2	3CLY	A	2.0	-138.8	-172.7	51.3	79.8	187.4	-3.1	-91.3	32.5	290.5	ACP
INSR	1IR3	A	1.9	-133.7	-168.2	47.5	85.0	199.9	-12.6	-100.3	24.6	303.4	ANP
INSR	3BU5	A	2.1	-128.7	-166.6	44.5	89.6	189.5	-14.9	-102.8	23.9	295.9	ATP
KAPCA	4WB5	A	1.6	-123.1	176.8	69.0	80.0	197.4	-7.7	-91.5	24.9	270.2	ATP
KAPCA	4WB6	A	2.1	-118.4	-169.5	51.7	82.5	197.0	-10.7	-95.6	17.7	271.2	ATP
KAPCA	4WB6	B	2.1	-117.8	-168.3	53.2	78.4	192.7	-12.2	-91.3	29.4	275.0	ATP
KAPCA	4WB8	A	1.5	-121.5	-177.2	62.7	77.8	196.8	-6.9	-92.0	24.6	268.9	ATP
KPCI	5LI1	A	2.0	-129.2	177.0	73.2	90.2	187.2	4.7	-103.9	11.5	308.2	ANP
MK01	5V60	A	2.1	-124.3	-175.3	60.5	75.4	193.8	5.9	-91.1	25.4	286.1	ACP
PAK1	3Q53	A	2.0	-131.4	-173.5	59.2	73.6	202.0	-16.8	-91.7	13.0	288.7	ATP
PAK4	4JDI	A	1.8	-134.3	177.9	63.7	71.6	185.9	8.0	-94.6	13.8	289.8	ANP
PDPK1	4A07	A	1.8	-135.4	-178.9	76.0	72.4	205.6	22.7	-109.0	8.7	286.3	ATP
PDPK1	4AW0	A	1.4	-134.8	177.3	78.7	76.2	199.3	18.1	-106.3	8.6	283.5	ATP
PDPK1	4AW1	A	1.6	-137.2	-178.3	69.0	84.7	188.0	9.9	-105.4	17.0	286.1	ATP
STK24	4QML	A	1.8	-130.4	172.3	70.8	79.3	196.8	-0.6	-104.6	20.0	285.9	ANP

Table S2: Cluster centroids

Spatial group	Cluster	X-DFG Φ,ψ	DFG-Asp Φ,ψ	DFG-Phe Φ,ψ	†DFG-Phe χ1	Example
DFGin	BLAminus	-129, 179	61, 81	-97, 20	-71	3BU5_A (INSR)
	BLAplus	-119, 168	59, 34	-89, -8	56	4OTH_A (PKN1)
	ABAminus	-112, -8	-141, 148	-128, 23	-64	1U46_A (ACK1)
	BLBminus	-135, 175	60, 65	-79, 145	-73	1R3C_A (MK14)
	BLBplus	-125, 172	60, 33	-85, 145	49	3PIX_A (BTK)
	BLBtrans	-106, 157	69, 21	-62, 134	-145	3VRZ_A (HCK)
	Noise	-	-	-	-	3BZ3_A (FAK1)
DFGout	BBAminus	-138, -176	-144, 104	-82, -9	-70	3BE2_A (VGFR2)
	Noise	-	-	-	-	3LW0_A (IGF1R)
DFGinter	BABtrans	-80, 128	-117, 24	-85, 133	-179	3SXR_A (BMX)
	Noise	-	-	-	-	5EW9_A (AURKA)

Angles in degrees.
†Phe side chain rotamers: plus: $0° < \chi_1 \leq 120°$; trans: $120° < \chi_1 \leq 240°$; minus: $240^o < \chi_1 \leq 360^o$

Table S3: Beta turns in activation loop across different conformations (% chains) and DFG-Asp χ_1 rotamer populations of each cluster (% of chains in each cluster)

Spatial groups	Clusters	Beta turn beginning with DFG-Asp (turn type)	Beta turn beginning with DFG-Phe (turn type)	Plus	Trans	Minus
DFGin	BLAminus	-	96 (I)	-	85	15
	BLAplus	49 (I)	11 (II)	-	25	75
	ABAminus	-	55 (I)	27	71	2
	BLBminus	7 (II)	-	-	67	31
	BLBplus	60 (II)	11 (II')	-	25	75
	BLBtrans	95 (II)	-		5	95
DFGout	BBAminus	7 (I)	10 (II')	31	66	3
DFGinter	BABtrans	95 (II)	-	52	39	10

Aspartic acid χ_1 rotamers are given as Plus ($0° < \chi_1 \leq 120°$), Trans ($120° < \chi_1 \leq 240°$), Minus ($240° < \chi_1 \leq 360°$).

Table S4: Structures with intact regulatory spines across different clusters (% of chains in each cluster)

Spatial group	Clusters	†S1 (HRD-His / DFG-Phe)	S2 (DFG-Phe / αC-Glu+4)	*S3 (αC-Glu+4 / β4-X)	All three
DFGin	BLAminus	100	99	99	98
	BLAplus	100	100	100	100
	ABAminus	67	98	98	67
	BLBminus	94	78	85	70
	BLBplus	99	74	94	73
	BLBtrans	99	27	95	27
DFGout	BBAminus	-	-	93	-
DFGinter	BABtrans	-	100	100	-

†A contact exists if the minimum distance between the side-chain atoms of the two residues is less than or equal to 4.5 Å.
*β4-X is typically a hydrophobic residue from β4 constituting the regulatory spine (Leu196 in Aurora A).

Table S5: Comparison of conformational labels with Möbitz's classification (number of chains)

		Möbitz DFGin							Möbitz DFGout				
Spatial group	**Clusters**	**Active**	**DFGact Chelixout**	**FG down**	**FGdown Chelixout**	**Gdown**	**Gdown Chelixout**	**AuP Met**	**DFG Flipped**	**DFGout Type2**	**AuP BRAF**	**AuP FMS**	**AuP IGF1R**
DFGin	BLAminus	1805	5	1	1	1	7	-	-	-	-	-	-
	BLAplus	1	13	38	7	-	-	-	-	-	-	-	-
	ABAminus	198	-	-	-	69	1	-	-	-	-	-	-
	BLBminus	5	-	5	6	31	49	-	-	-	-	-	-
	BLBplus	-	1	78	125	-	48	34	-	-	-	-	-
	BLBtrans	-	-	-	179	-	-	-	-	-	-	-	-
	Noise	43	2	2	16	2	6	1	3	-	-	-	-
DFGout	BBAminus	-	-	-	-	-	-	-	-	73	31	20	1
	Noise	-	-	-	-	-	-	-	12	2	7	6	4
DFGinter	BABtrans	-	-	-	-	-	-	-	13	-	-	-	-
	Noise	-	-	-	-	-	-	-	6	-	-	-	-

Table S6: Comparison of conformational labels with Ung et al.'s classification (number of chains)

Spatial groups	Clusters	*CIDI	†CIDO	‡CODI	§CODO	¶ωCD
DFGin	BLAminus	1434	-	18	-	-
	BLAplus	31	-	30	-	-
	ABAminus	265	-	6	-	-
	BLBminus	67	-	68	-	-
	BLBplus	87	-	212	-	-
	BLBtrans	-	-	173	-	-
	Noise	55	-	41	-	28
DFGout	BBAminus	-	135	-	17	1
	Noise	-	34	-	14	32
DFGinter	BABtrans	-	-	-	-	8
	Noise	-	-	5	-	50

* CIDI: C-helix-in - DFGin

† CIDO: C-helix-in - DFGout

‡ CODI: C-helix-out - DFGin

§ CODO: C-helix-out - DFGout

¶ ωCD - DFGintermediate

Table S7: FDA approved inhibitors in complex with structures from different clusters.

Inhibitor	PDB code	Clusters	Kinases (PDBid and chain(s))
Afatenib	0WN	BLAminus	EGFR (4G5JA,4G5PA)
Alectinib	EMH	ABAminus	ALK (3AOXA)
Axitinib	AXI	BLAminus	ABL1 (4TWPA,B)
		BBAminus	VGFR2 (4AG8A)
		DFGout-noise	ABL1 (4WA9B)
Bosutinib	DB8	BLAminus	PMYT1 (5VCYA), SRC (4MXOA,B, 4MXYA,B, 4MXZA,B), STK24 (4QMNA), WEE1 (5VC3A)
		BLBminus	EPHA2 (5I9XA)
		BLBplus	ERBB3 (4OTWA)
		DFGin-noise	STK10 (5AJQA,B)
		DFGinter-noise	ABL1 (3UE4A,B), KCC2(3SOAA)
Cobimetinib	EUI	BLBplus	MP2K1 (4AN2A, 4LMNA)
Crizotinib	VGH	ABAminus	ALK (2XP2A, 2YFXA, 5AAAA, 5AABA, 5AACA), ROS1 (3ZBFA)
		BLBplus	MET (2WGJA)
Dabrafenib	P06	BLBminus	BRAF (4XV2A,B, 5CSWB, 5HIEA,B,C,D)
Dacomitinib	1C9	BLAminus	EGFR (4I23A)
		BLBtrans	EGFR (4I24A,B)
Dasatinib	1N1	BLAminus	ABL1 (2GQGA,B), PMYT1 (5VCVA), STK10 (5OWRA), STK24 (4QMSA)
		BLAplus	BTK (3K54A)
		BLBminus	EPHA2 (5I9YA)
		BLBplus	PTK6 (5H2UA,B,C,D)
		BABtrans	ABL1 (4XEYB), BMX (3SXRA,B), BTK (3OCTA)
		DFGout-noise	DDR1 (5BVWA)
Erlotinib	AQ4	BLAminus	EGFR (1M17A)
		BLBtrans	EGFR (4HJOA)
Gefitinib	IRE	BLAminus	EGFR (2ITOA, 2ITYA, 2ITZA, 3UG2A, 4WKQA)
		BLBtrans	EGFR (4I22A)
Ibrutinib	1E8	BLBplus	BTK (5P9IA)
Imatinib	STI	BLAminus	KSYK (1XBBA)
		BBAminus	ABL1 (2HYYA,B,C,D, 3PYYA,B), ABL2 (3GVUA), CSF1R (4R7IA), DDR1 (4BKJA,B), KIT (1T46A), LCK (2PL0A)
Lapatinib	FMM	BLBplus	EGFR (1XKKA), ERBB4 (3BBTB,D)
Lenvatinib	LEV	BLBplus	VGFR2 (3WZDA)
Neratinib	HKI	BLAminus	EGFR (3W2QA)
		BLBplus	EGFR (2JIVB)
Nilotinib	NIL	BBAminus	ABL1 (3CS9A,B,C,D), MK11 (3GP0A)
Nintedanib	XIN	BLAminus	AAK1 (5TE0A)
Osimertinib	YY3	BLAminus	EGFR (4ZAUA)
Palbociclib	LQQ	BLAminus	CDK6 (2EUFB)
		BLBplus	CDK6 (5L2IA)
Ponatinib	0LI	BBAminus	DDR1 (3ZOSA,B), FGFR4 (4TYJA, 4UXQA)
Ribociclib	6ZZ	BLBplus	CDK6 (5L2TA)
Sorafenib	BAX	BBAminus	BRAF (1UWHA,B, 1UWJA,B, 5HI2A), CDK8 (3RGFA), MK14 (3GCSA, 3HEGA), VGFR2 (3WZEA)
Sunitinib	B49	BLAminus	PAK6 (4KS8A), PHKG2 (2Y7JA,B,C,D), STK24 (4QMZA)
		BLAplus	ITK (3MIYB)
		BLBminus	CDK2 (3TI1A)
		BBAminus	KIT (3G0FA,B)
		DFGout-noise	KIT (3G0EA)
Tofacitinib	MI1	BLAplus	PKN1 (4OTIA)
		ABAminus	JAK1_2 (3EYGA), JAK2_2 (3FUPA), JAK3_2 (3LXKA), TYK2 (3LXNA)
Vemurafenib	032	BLAplus	BRAF (3OG7A, 4RZVA,B), MLTK (5HESA,B)

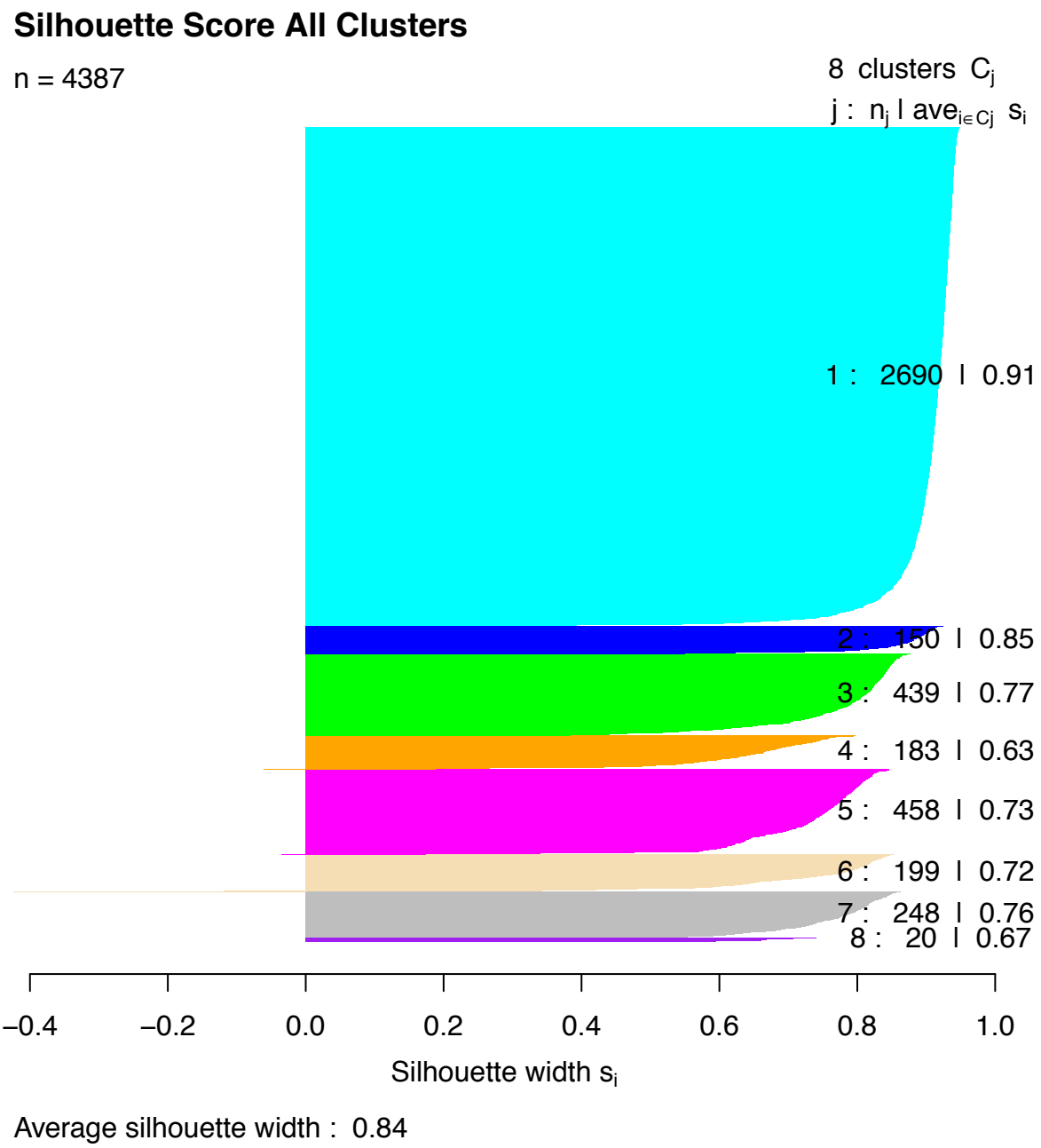

Fig. S1. Silhouette scores for all the clusters. Average value over all data points is 0.84, which indicates that the clusters are well separated and well defined.

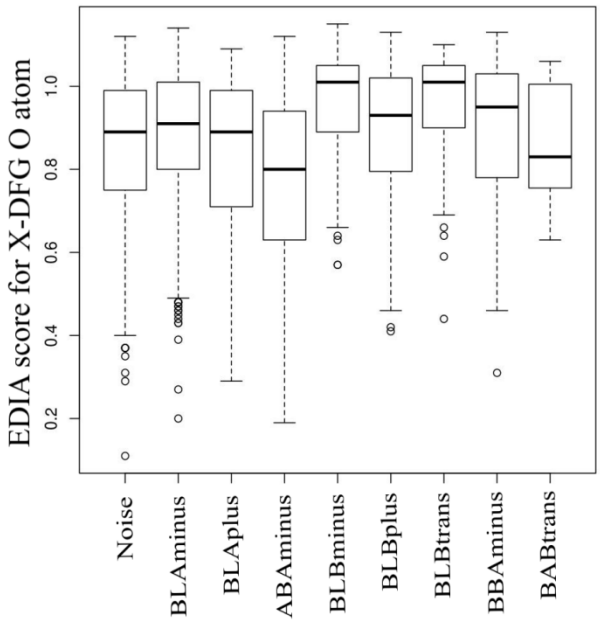

Fig. S2. EDIA score of X-DFG O atom across different clusters. An atom is modeled correctly if its score is 0.8 or more.

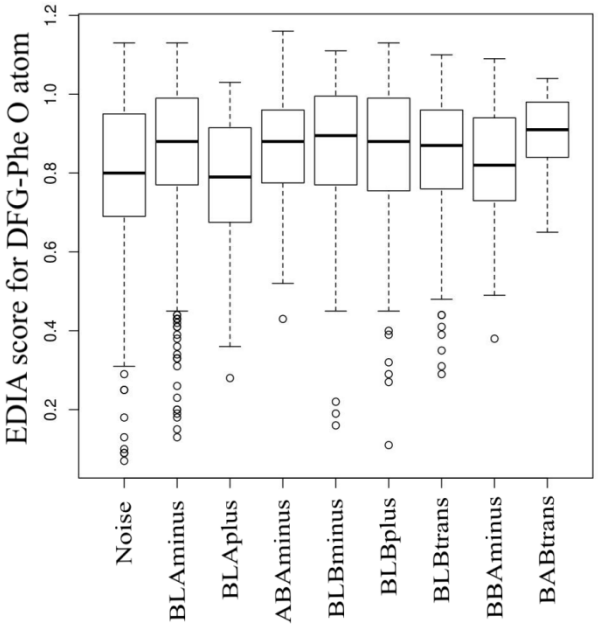

Fig. S3. EDIA score for DFG-Phe O atom.

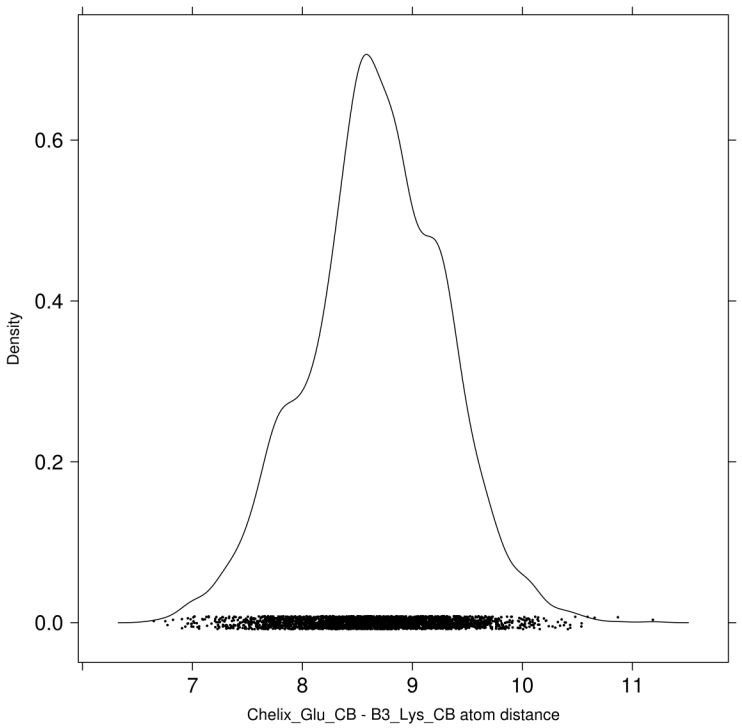

Fig. S4. Distribution of distances between C-helix-Glu-CB and β3-Lys-CB atoms in the structures which have an intact salt bridge between these residues (3327 chains, 200 kinases).

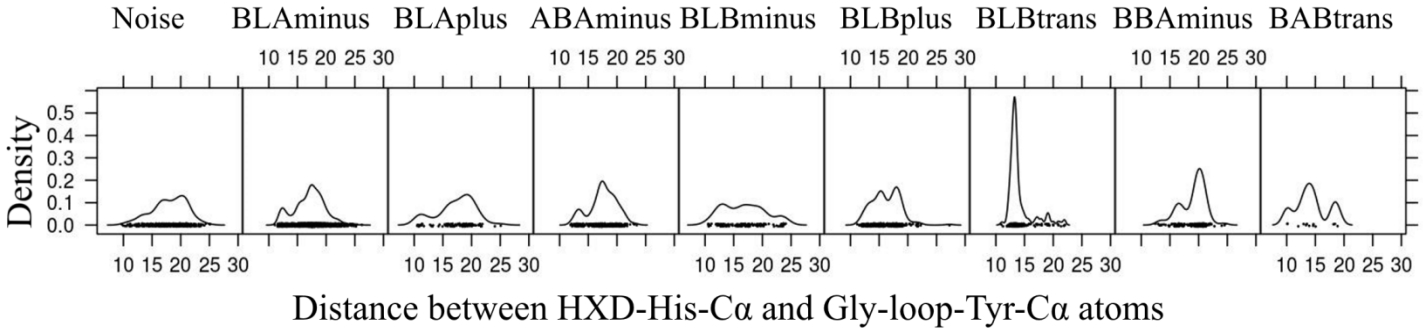

Fig. S5. Distribution of Gly-rich loop conformations - distance between HRD-His-Cα and Gly-rich loop-Phe/Tyr-Cα atoms (Phe144 in Aurora A kinase).

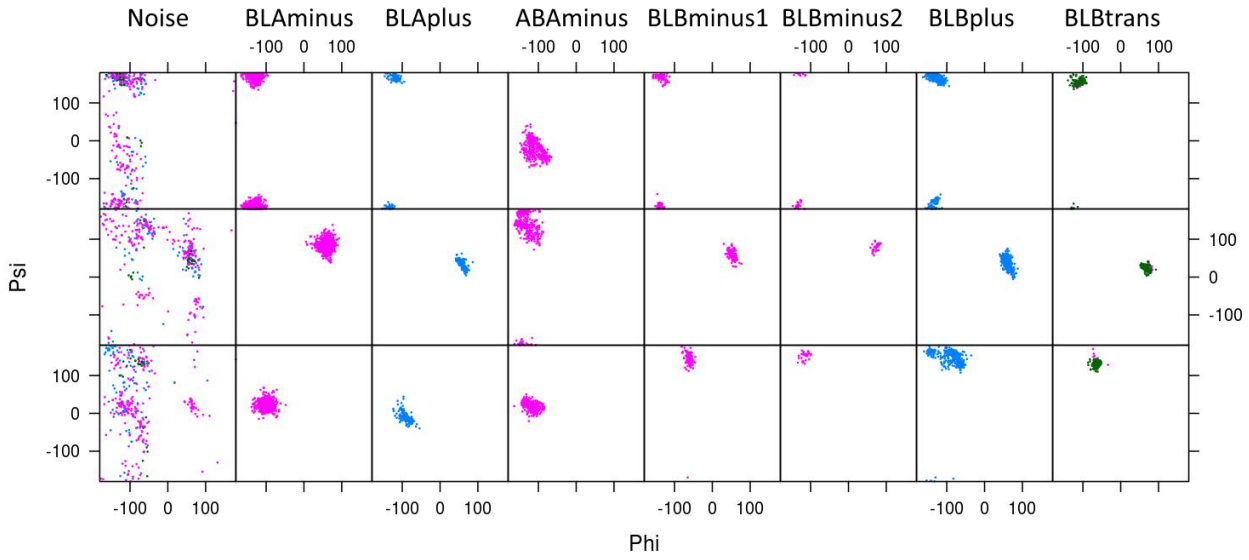

Fig. S6: DBSCAN clustering of XDF conformations of DFGin structures with ε=0.04 and *MinPts*=20 (clustering in Fig. 4 used ε=0.05 and *MinPts*=20). Backbone dihedral angles of X-DFG (top row), DFG-Asp (middle row) and DFG-Phe (bottom row) and the χ_1 of DFG-Phe (minus in magenta; plus in blue; trans in green) are plotted. The lower value of ε splits the "B" region of DFG-Phe residue in the BLBminus cluster into two subclusters labeled BLBminus1 and BLBminus2. These two clusters differ by about 40° in the ϕ dihedral of the DFG-Phe residue. We could not find any biologically meaningful difference between the structures in these two subclusters.

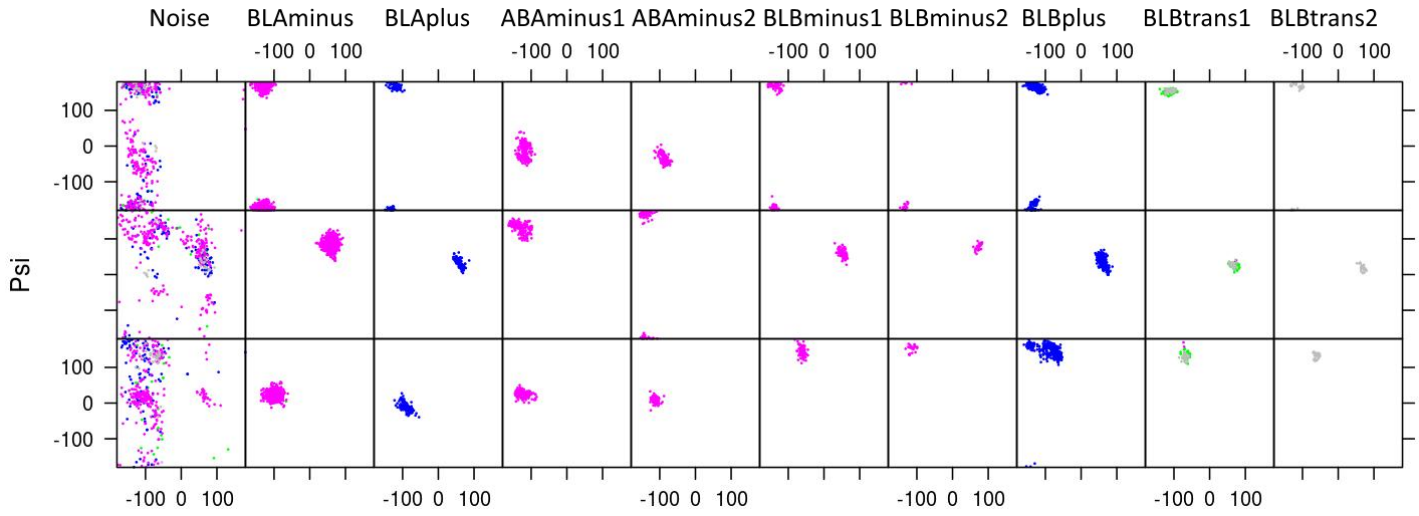

Fig. S7: DBSCAN clustering of XDF conformations of DFGin structures with ε=0.03 and *MinPts*=20 (clustering in Fig. 4 used ε=0.05 and *MinPts*=20). Backbone dihedral angles of X-DFG (top row), DFG-Asp (middle row) and DFG-Phe (bottom row) and the χ_1 of DFG-Phe (minus in magenta; plus in blue; trans in green) are plotted. The lower value of ε splits the ABAminus, BLBminus, and BLBtrans clusters in a way that is not productive (labeled ABAminus1,2, etc.).

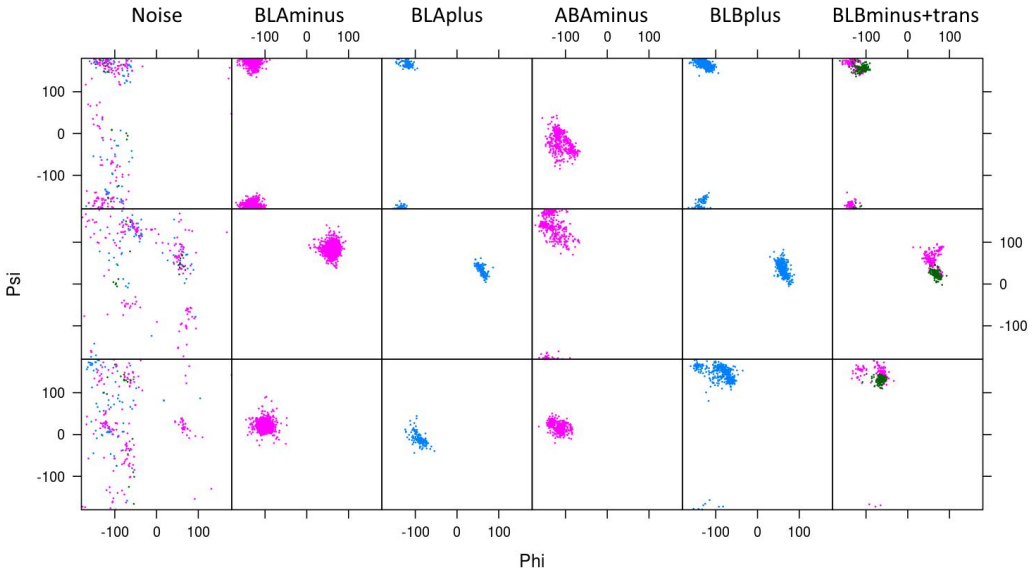

Fig. S8: DBSCAN clustering of XDF conformations of DFGin structures with ε=0.06 and *MinPts*=20 (clustering in Fig. 4 used ε=0.05 and *MinPts*=20). Backbone dihedral angles of X-DFG (top row), DFG-Asp (middle row) and DFG-Phe (bottom row) and the χ_1 of DFG-Phe (minus in magenta; plus in blue; trans in green) are plotted. With the higher value of ε, the BLBminus and BLBtrans clusters with different DFG-Phe rotamers merge into one cluster. By examining structures in these two clusters, we observe that the directionality of the activation loop is different in each structure as is the space available in the C-helix pocket for inhibitors.

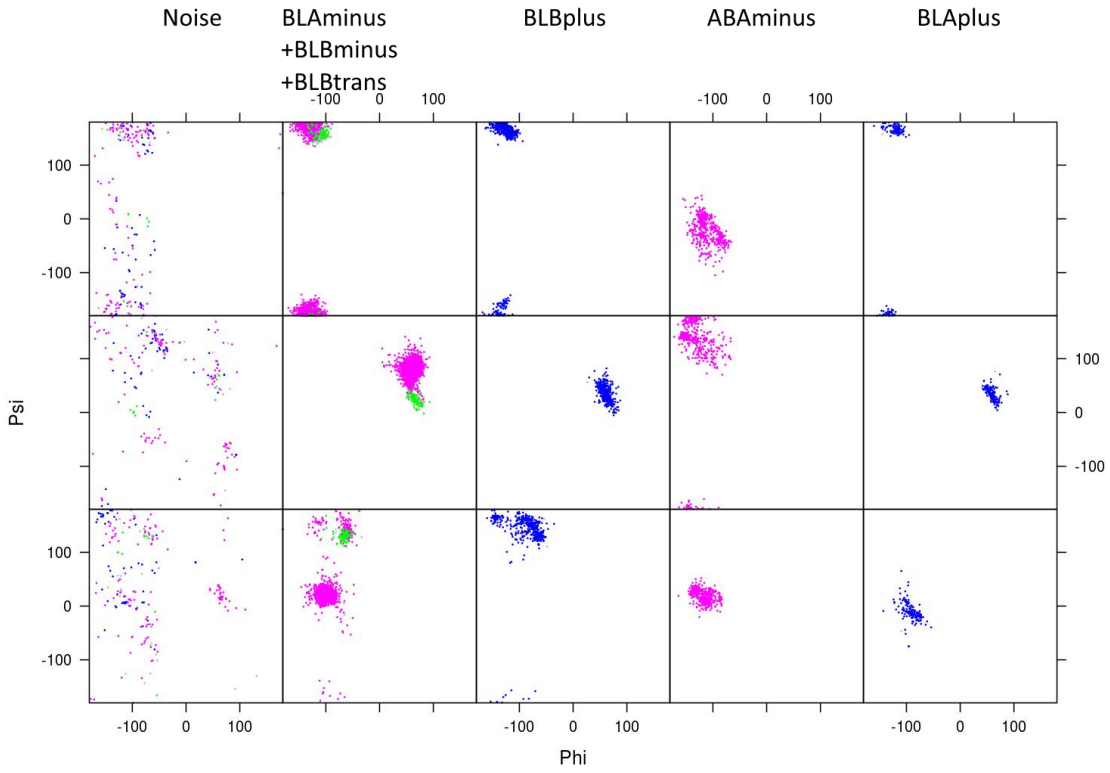

Fig. S9: DBSCAN clustering of XDF conformations of DFGin structures with ε=0.09 and *MinPts*=20 (clustering in Fig. 4 used ε=0.05 and *MinPts*=20). Backbone dihedral angles of X-DFG (top row), DFG-Asp (middle row) and DFG-Phe (bottom row) and the χ_1 of DFG-Phe (minus in magenta; plus in blue; trans in green) are plotted. With the higher value of ε, BLAminus, BLBminus, and BLBtrans structures are grouped together (2nd column), merging distinct regions of the Ramachandran map and different rotamers into the same cluster.