Coupled regulation by the juxtamembrane and sterile α motif (SAM) linker is a hallmark of Ephrin tyrosine kinase evolution

Annie Kwon, Mihir John, Zheng Ruan, Natarajan Kannan

Supporting information

- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6
- Figure S7
- Figure S8
- Figure S9
- Table S1
- Table S2
- SI Dataset 1
- SI Methods and Materials

Figure S1

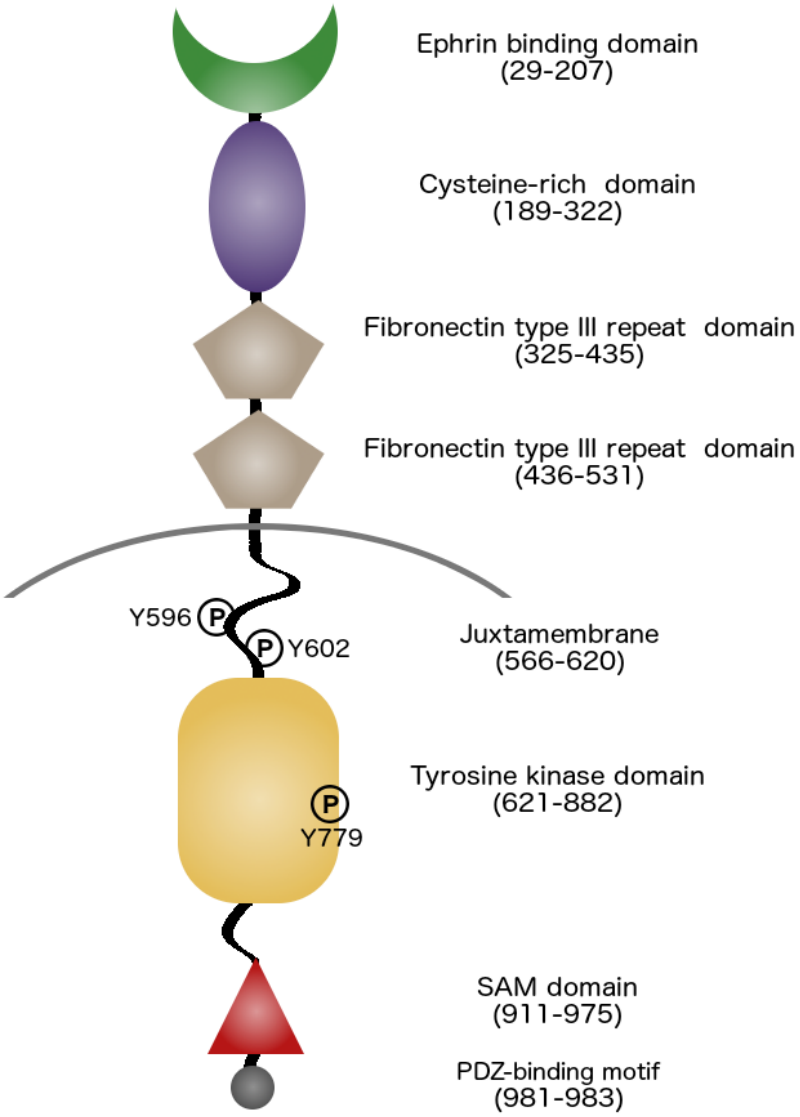

Figure S1. Schematic representation of the domain organization of human EphA3. The domain organization of human EphA3 is illustrated above, with residue numbers corresponding to each region indicated in parentheses. The three major autophosphorylation sites in the juxtamembrane and kinase domain of EphA3 are also indicated.

Figure S2

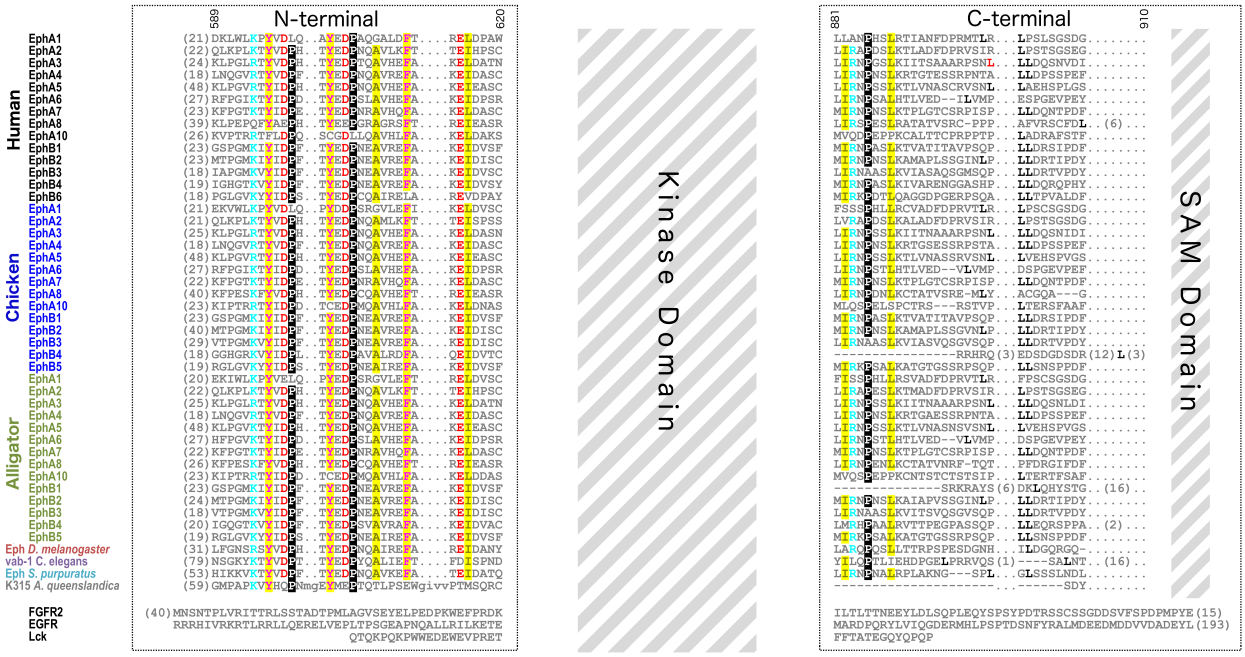

Figure S2. Alignment of N-terminal and C-terminal Eph sequences flanking the kinase domain. Well-aligned portions of the N-terminal juxtamembrane and C-terminal SAM domain linker are shown, with lengths of omitted segments denoted in parentheses. Sequence numbering above the alignment corresponds to the human EphA3 sequence. Columns with >60% amino acid similarity are highlighted, and divergent N- and C-terminal sequences of other tyrosine kinases are shown below the alignment. The leucine of the SAM linker tethering network (Leu901 in EphA3) and potentially equivalent leucines in other Eph orthologs are shown in red and black font, respectively.

Figure S3

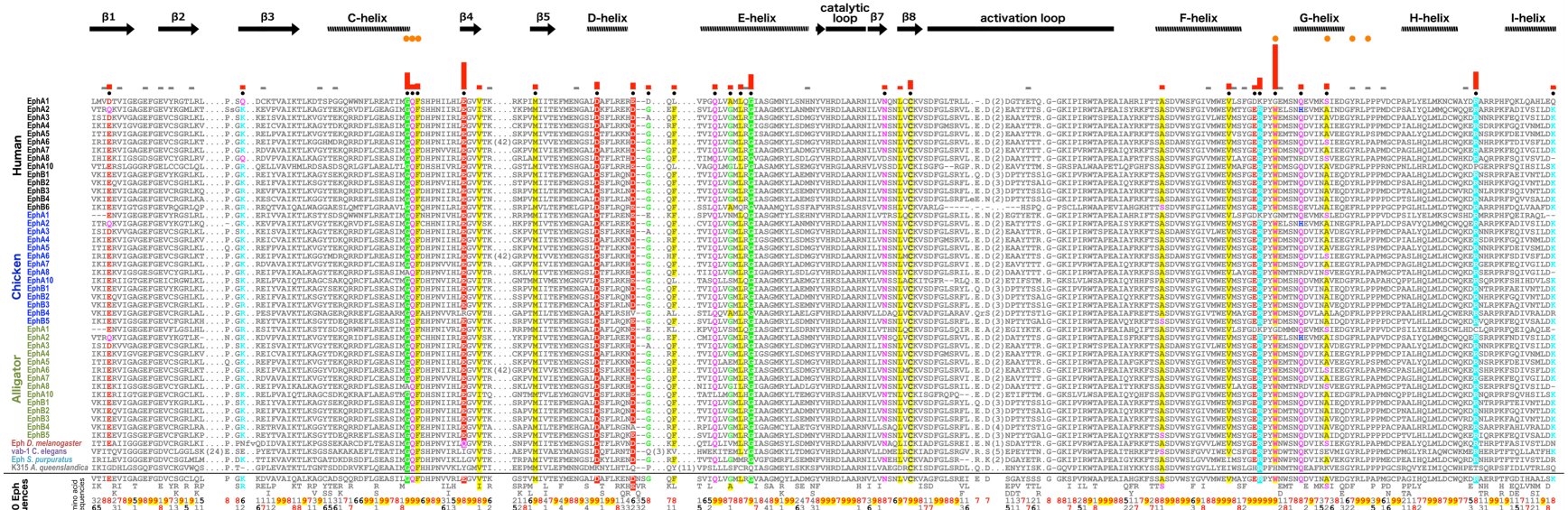

Figure S3. Unique sequence features distinguishing the Eph family from other tyrosine kinases identified across tyrosine kinase domain alignment. Residues involved in interactions with the juxtamembrane and SAM domain linker are indicated with orange dots.

Figure S4

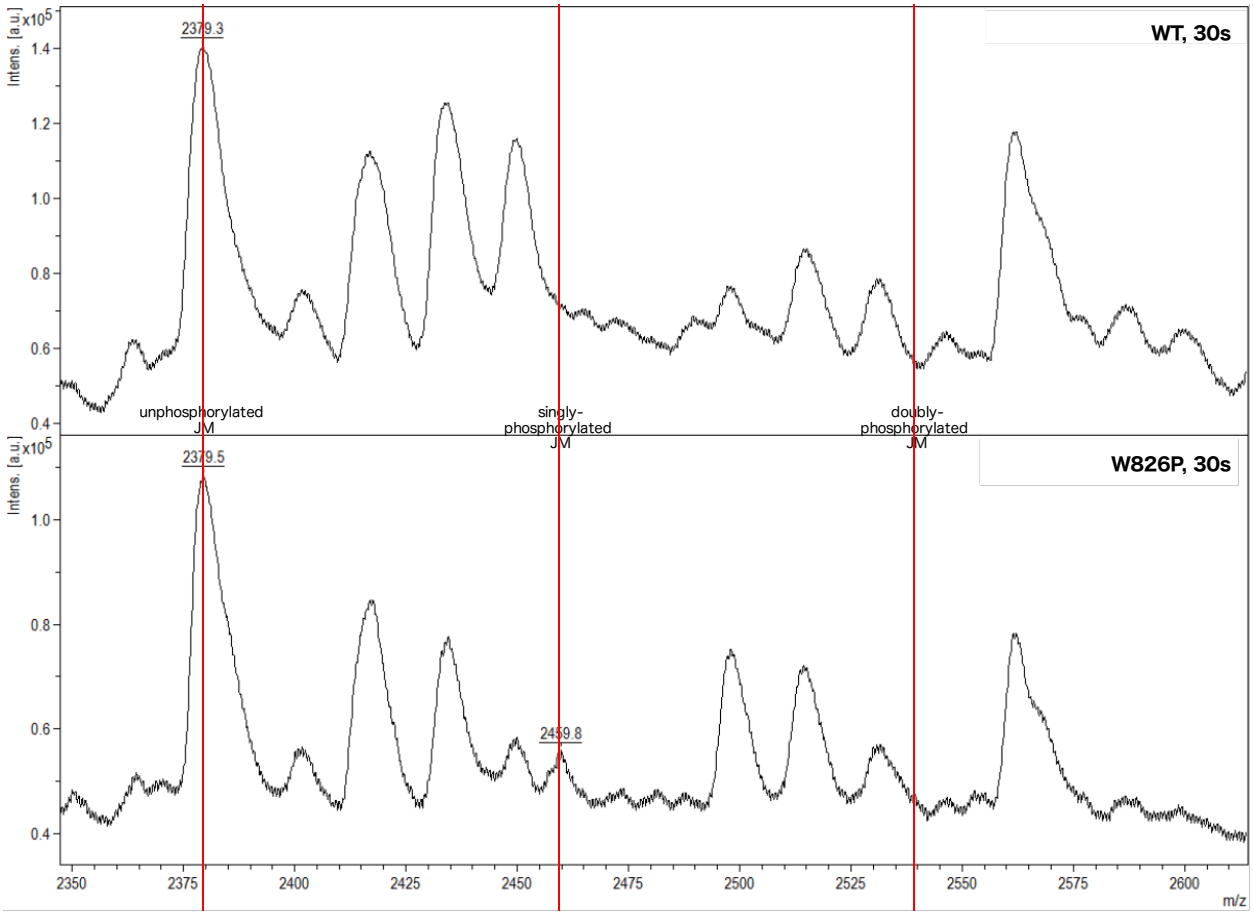

Figure S4. MALDI peptide mass fingerprinting of WT and W826P samples after 30s incubation with MgATP. Red lines indicate expected m/z values for the unphosphorylated, singly-phosphorylated, and doubly-phosphorylated juxtamembrane peptide.

Figure S5

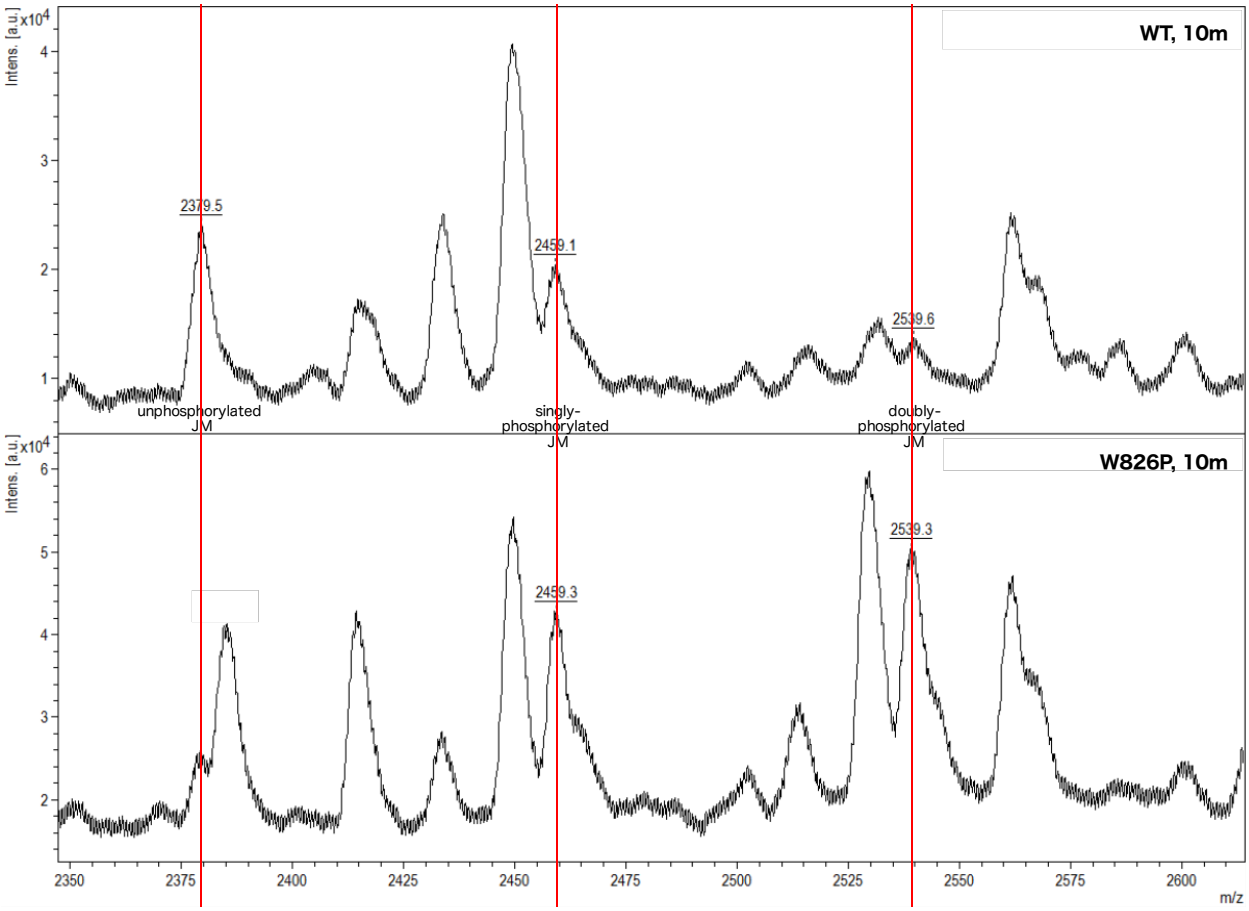

Figure S5. MALDI peptide mass fingerprinting of WT and W826P samples after 10 min incubation with MgATP. Red lines indicate expected m/z values for the unphosphorylated, singly-phosphorylated, and doubly-phosphorylated juxtamembrane peptide.

Figure S6

Sequence: TYVDPHTYEDPTQTVHEFAK, Y8-Phospho (79.96633 Da)
Charge: +4, Monoisotopic m/z: 615.26978 Da (-0.66 mmu/-1.08 ppm), MH+: 2458.05727 Da, RT: 39.76 min,
Identified with: Mascot (v1.30); IonScore:28, Exp Value:3.5E-002, Ions matched by search engine: 14/152

#1	c^{+}	c^{2+}	$c+1^{+}$	$c+1^{2+}$	Seq.	z^{+}	z^{2+}	$z+1^{+}$	$z+1^{2+}$	#2
1	119.08151	60.04439	120.08933	60.54830	T					20
2	282.14483	141.57605	283.15265	142.07996	Y	2340.99352	1171.00040	2342.00134	1171.50431	19
3	381.21325	191.11026	382.22107	191.61417	V	2177.93020	1089.46874	2178.93802	1089.97265	18
4	496.24020	248.62374	497.24802	249.12765	D	2078.86178	1039.93453	2079.86960	1040.43844	17
5	593.29297	297.15012	594.30079	297.65403	P	1963.83483	982.42105	1964.84265	982.92496	16
6	730.35188	365.67958	731.35970	366.18349	H	1866.78206	933.89467	1867.78988	934.39858	15
7	831.39956	416.20342	832.40738	416.70733	T	1729.72315	865.36521	1730.73097	865.86912	14
8	**1074.42921**	**537.71824**	**1075.43703**	**538.22215**	**Y-Phospho**	**1628.67547**	**814.84137**	**1629.68329**	**815.34528**	**13**
9	1203.47181	602.23954	1204.47963	602.74345	E	1385.64582	693.32655	1386.65364	693.83046	12
10	1318.49876	659.75302	1319.50658	660.25693	D	1256.60322	628.80525	1257.61104	629.30916	11
11	1415.55153	708.27940	1416.55935	708.78331	P	1141.57627	571.29177	1142.58409	571.79568	10
12	1516.59921	758.80324	1517.60703	759.30715	T	1044.52350	522.76539	1045.53132	523.26930	9
13	1644.65779	822.83253	1645.66561	823.33644	Q	943.47582	472.24155	944.48364	472.74546	8
14	1745.70547	873.35637	1746.71329	873.86028	T	815.41724	408.21226	816.42506	408.71617	7
15	1844.77389	922.89058	1845.78171	923.39449	V	714.36956	357.68842	715.37738	358.19233	6
16	1981.83280	991.42004	1982.84062	991.92395	H	615.30114	308.15421	616.30896	308.65812	5
17	2110.87540	1055.94134	2111.88322	1056.44525	E	478.24223	239.62475	479.25005	240.12866	4
18	2257.94382	1129.47555	2258.95164	1129.97946	F	349.19963	175.10345	350.20745	175.60736	3
19	2328.98094	1164.99411	2329.98876	1165.49802	A	202.13121	101.56924	203.13903	102.07315	2
20					K	131.09409	66.05068	132.10191	66.55459	1

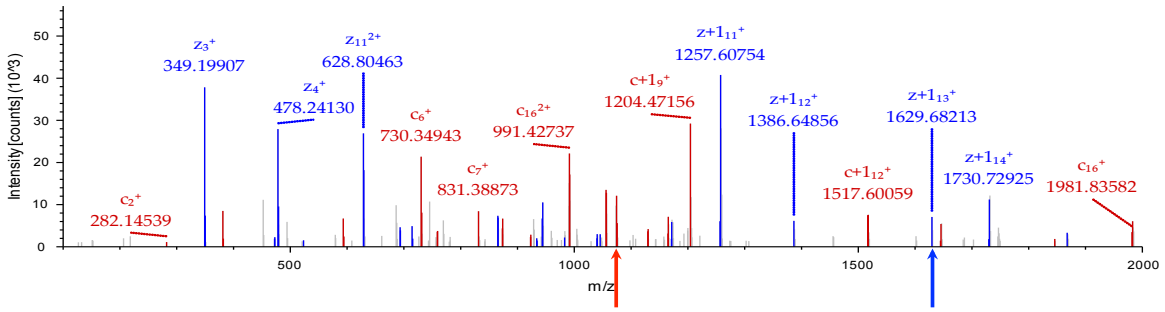

Figure S6. CID spectrum of the C-terminal phosphorylated juxtamembrane from LC-MS/MS analysis of the W826P sample after 30s incubation with MgATP. Red and blue values indicate matched peptide ions.

Figure S7

Sequence: TYVDPHTYEDPTQTVHEFAK, Y2-Phospho (79.96633 Da)
Charge: +3, Monoisotopic m/z: 820.02344 Da (-1.39 mmu/-1.69 ppm), MH+: 2458.05576 Da, RT: 45.25 min,
Identified with: Mascot (v1.30); IonScore:24, Exp Value:9.2E-002, Ions matched by search engine: 14/198
Fragment match tolerance used for search: 0.02 Da

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	102.05496	51.53112	T			20
2	**345.08461**	**173.04594**	**Y-Phospho**	**2357.01224**	**1179.00976**	**19**
3	444.15303	222.58015	V	2113.98259	1057.49493	18
4	559.17998	280.09363	D	2014.91417	1007.96072	17
5	656.23275	328.62001	P	1899.88722	950.44725	16
6	793.29166	397.14947	H	1802.83445	901.92086	15
7	894.33934	447.67331	T	1665.77554	833.39141	14
8	1057.40266	529.20497	Y	1564.72786	782.86757	13
9	1186.44526	593.72627	E	1401.66454	701.33591	12
10	1301.47221	651.23974	D	1272.62194	636.81461	11
11	1398.52498	699.76613	P	1157.59499	579.30113	10
12	1499.57266	750.28997	T	1060.54222	530.77475	9
13	1627.63124	814.31926	Q	959.49454	480.25091	8
14	1728.67892	864.84310	T	831.43596	416.22162	7
15	1827.74734	914.37731	V	730.38828	365.69778	6
16	1964.80625	982.90676	H	631.31986	316.16357	5
17	2093.84885	1047.42806	E	494.26095	247.63411	4
18	2240.91727	1120.96227	F	365.21835	183.11281	3
19	2311.95439	1156.48083	A	218.14993	109.57860	2
20			K	147.11281	74.06004	1

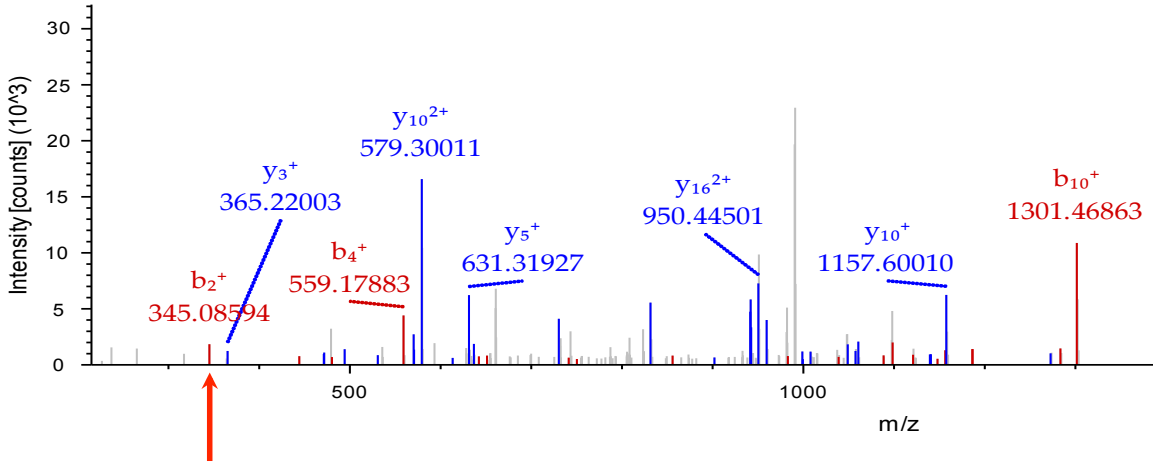

Figure S7. CID spectrum of the N-terminal phosphorylated juxtamembrane from LC-MS/MS analysis of the W826P sample after 30s incubation with MgATP. Red and blue values indicate matched peptide ions.

Figure S8

Sequence: TYVDPHTYEDPTQTVHEFAK, Y2-Phospho (79.96633 Da), Y8-Phospho (79.96633 Da)
Charge: +3, Monoisotopic m/z: 846.68201 Da (+1.74 mmu/+2.05 ppm), MH+: 2538.03147 Da, RT: 39.59 min,
Identified with: Mascot (v1.30); IonScore:16, Exp Value:4.1E-001, Ions matched by search engine: 8/152

#1	c^+	c^{2+}	$c+1^+$	$c+1^{2+}$	Seq.	z^+	z^{2+}	$z+1^+$	$z+1^{2+}$	#2
1	119.08151	60.04439	120.08933	60.54830	T					20
2	**362.11116**	**181.55922**	**363.11898**	**182.06313**	**Y-Phospho**	**2420.95985**	1210.98356	**2421.96767**	1211.48748	**19**
3	461.17958	231.09343	462.18740	231.59734	V	2177.93020	1089.46874	2178.93802	1089.97265	18
4	576.20653	288.60690	577.21435	289.11081	D	2078.86178	1039.93453	2079.86960	1040.43844	17
5	673.25930	337.13329	674.26712	337.63720	P	1963.83483	982.42105	1964.84265	982.92496	16
6	810.31821	405.66274	811.32603	406.16665	H	1866.78206	933.89467	1867.78988	934.39858	15
7	911.36589	456.18658	912.37371	456.69049	T	1729.72315	865.36521	1730.73097	865.86912	14
8	**1154.39554**	**577.70141**	**1155.40336**	**578.20532**	**Y-Phospho**	1628.67547	**814.84137**	1629.68329	**815.34528**	**13**
9	1283.43814	642.22271	1284.44596	642.72662	E	1385.64582	693.32655	1386.65364	693.83046	12
10	1398.46509	699.73618	1399.47291	700.24009	D	1256.60322	628.80525	1257.61104	629.30916	11
11	1495.51786	748.26257	1496.52568	748.76648	P	1141.57627	571.29177	1142.58409	571.79568	10
12	1596.56554	798.78641	1597.57336	799.29032	T	1044.52350	522.76539	1045.53132	523.26930	9
13	1724.62412	862.81570	1725.63194	863.31961	Q	943.47582	472.24155	944.48364	472.74546	8
14	1825.67180	913.33954	1826.67962	913.84345	T	815.41724	408.21226	816.42506	408.71617	7
15	1924.74022	962.87375	1925.74804	963.37766	V	714.36956	357.68842	715.37738	358.19233	6
16	2061.79913	1031.40320	2062.80695	1031.90711	H	615.30114	308.15421	616.30896	308.65812	5
17	2190.84173	1095.92450	2191.84955	1096.42841	E	478.24223	239.62475	479.25005	240.12866	4
18	2337.91015	1169.45871	2338.91797	1169.96262	F	349.19963	175.10345	350.20745	175.60736	3
19	2408.94727	1204.97727	2409.95509	1205.48118	A	202.13121	101.56924	203.13903	102.07315	2
20					K	131.09409	66.05068	132.10191	66.55459	1

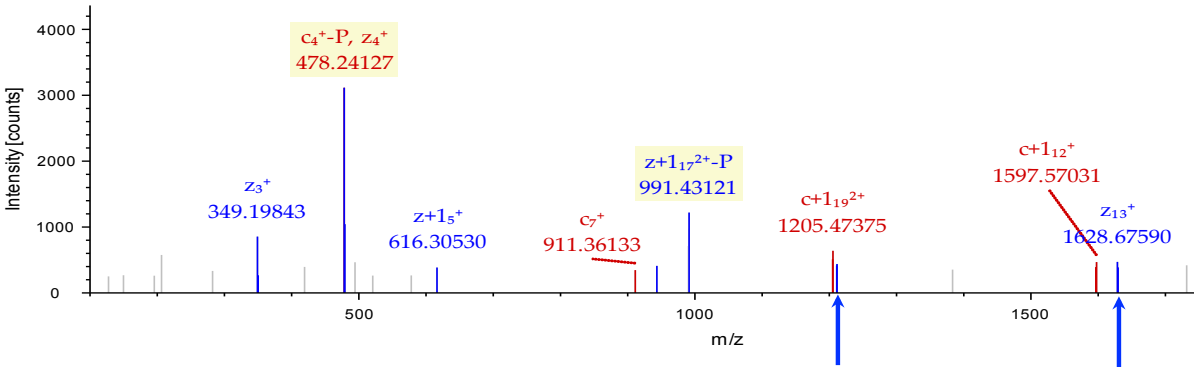

Figure S8. CID spectrum of the doubly-phosphorylated juxtamembrane from LC-MS/MS analysis of the W826P sample after 30s incubation with MgATP. Red and blue values indicate matched peptide ions.

Figure S9

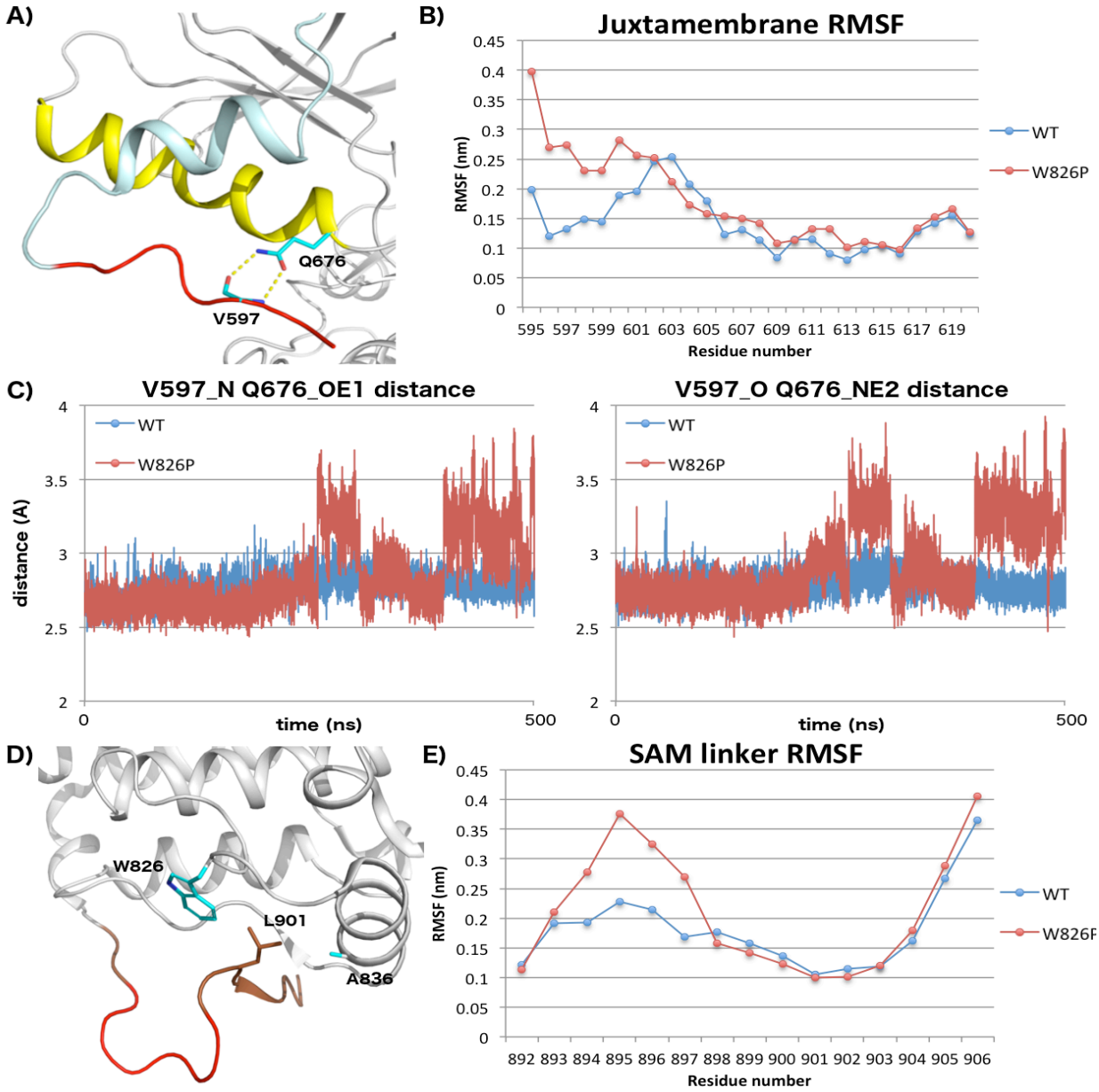

Figure S9. Molecular dynamics analysis of WT and W826P inactive models. A) Key polar interactions in the autoinhibitory conformation of EphA3 (PDB 2QO2). The red segment indicates the 7-residue region that was found to be more flexible in the W826P mutant relative to WT (B). B) RMSF values of juxtamembrane Cα atoms in WT and W826P molecular dynamics simulations. C) Distance between oxygen and nitrogen atoms involved in the two polar interactions in the autoinhibited form calculated from WT and W826P molecular dynamics simulations. D) Key interactions between the C-terminal lobe of the kinase domain and the SAM domain linker in EphA3 (PDB 2QOC). The red segment indicates the 5-residue region that was found to be more flexible in the W826P mutant relative to WT (E). E) RMSF values of SAM linker Cα atoms in WT and W826P simulations.

Table S1. Rates of peptide phosphorylation by inactive and active WT and W826P proteins using NADH-coupled assay.

	specific activity (nmol/min*μmol)	
	active	inactive
WT	3.70 ± 0.19	0.607 ± 0.08
W826P	2.93 ± 0.16	0.954 ± 0.18

Table S2. Tryptic juxtamembrane peptides identified by MALD peptide mass fingerprinting.

			WT 30s			WT 10m			W826P 30s			W826P 10m		
Sequence	Modifications	Theoretical m/z	Observed m/z	Intensity	Mass error (ppm)	Observed m/z	Intensity	Mass error (ppm)	Observed m/z	Intensity	Mass error (ppm)	Observed m/z	Intensity	Mass error (ppm)
TYVDPHTYEDPTQTVHEFAK		2379.561	2379.349	140280.000	-88.88196	2379.520	23529.000	-17.01995	2379.595	108491.000	14.498476	2379.298	25899.000	-110.31449
TYVDPHTYEDPTQTVHEFAK	phosphorylation (1)	2459.54				2459.129	19835.000	-167.22637	2459.802	54618.000	106.402	2459.333	41238.000	-84.284043
TYVDPHTYEDPTQTVHEFAK	phosphorylation (2)	2539.52										2539.290	50184.000	-90.647044

SI Methods and Materials

Expression and purification of EphA3 WT and mutant proteins
The WT human EphA3 plasmid was obtained from Addgene (plasmid id 25136). The EphA3 coding region corresponds to residues Asp577-Ser947 with an N-terminal 6X His tag. Mutagenesis was carried out using high fidelity Phusion DNA polymerase (New England Biolabs). Plasmids were expressed in BL21 DE3 Rosetta II cells. Cultures were grown to O.D. 600 of 0.8 and induced with 0.2 mM IPTG at 18°C for 16 hours. Cells were harvested by centrifugation and resuspended in 50 mM sodium phosphate buffer pH 8.0, 0.5M NaCl followed by sonication. Cell debris was discarded, and the supernatant was applied on a Ni-NTA agarose column. The protein was eluted in 25 mM HEPES, 0.1 mM NaCl, 300 mM imidazole pH 7.5. Mass spectrometry revealed that the protein had degraded to boundaries corresponding to Asp577-Thr912. The cleaved product was purified out using a MonoQ column, and the smaller fragment was used for subsequent experiments. Proteins were dialyzed into 25 mM HEPES, 0.5 mM NaCl at pH 7.5. Protein concentrations were quantified by UV absorbance at 280 nm.

Mass spectrometry analysis of autophosphorylated samples
All mass spectrometry experiments were performed at the Proteomics and Mass Spectrometry Facility, University of Georgia. Autophosphorylation samples were run on 10% acrylamide gel and stained with coomassie. An in-gel trypsin digest was performed by slicing the gel into small pieces and rinsing with 50% acetonitrile/20 mM ammonium bicarbonate (~pH7.5-8) twice. Gel pieces were dehydrated by adding 100% of acetonitrile and dried out by a SpeedVac. Trypsin solution (0.01µg/µL in 20 mM ammonium bicarbonate) was added until completely absorbed. The tubes were placed in an incubator at 37°C overnight. Tryptic peptides were extracted by incubating with 50% acetonitrile/0.1% formic acid twice. Extracts were dried down by a SpeedVac. For MALDI analysis, samples were brought up in 10 ul of 20% acetonitrile w/0.1% formic acid (FA): 80% water with 0.1% FA. 1 ul was spotted with 1 ul of matrix (15mg/ml of DHB in 50:50 acetonitrile w/0.1% formic acid (FA):water with 0.1% FA). Data was collected using a Bruker Daltonics Autoflex operated in reflectron mode and analyzed using Matrix Science online software.

LC-MS/MS analyses were performed on a Thermo-Fisher LTQ Orbitrap Elite Mass Spectrometer coupled with a Proxeon Easy NanoLC system. Peptides were loaded into a reversed-phase column (self-packed column/emitter with 200 Å 5 µM Bruker MagicAQ C18 resin), then directly eluted into the mass spectrometer. The two-buffer gradient elution (0.1% formic acid as buffer A and 99.9% acetonitrile with .0.1% formic acid as buffer B) starts with 5% B, holds at 5%B for 2 minutes, then increases to 25% B in 60 minutes, to 40% B in 10 minutes, and to 95% B in 10 minutes. The data-dependent acquisition (DDA) method was used to acquire MS data. A survey MS scan was acquired first, then the top 5 ions in the MS scan were selected for following CID and HCD MS/MS analysis. Both MS and MS/MS scans were acquired by Orbitrap at resolutions of 120,000 and 30,000, respectively. Data were acquired using Xcalibur software (2.2, Thermo Fisher Scientific). Proteins identification and modification characterization were performed using Thermo Proteome Discoverer (1.4) with Mascot (Matrix Science) and Uniprot database. The spectra of possible modified peptides were inspected further to verify the accuracy of the assignments. The semi-quantitative analyses were achieved using a label-free quantification workflow within Proteome Discoverer, which a Precursor Ion Areas

Detector node calculates the average of the top three highest peptide areas in the Extracted Ion Chromatograms (mass precision, 2 ppm).

NADH-coupled assay

We employed a NADH coupled enzyme assay in 96 well format. In the assay, reaction progress (conversion of ATP to ADP) was monitored by a coupled reaction that leads to oxidation of NADH. PK/LDH, PEP, and NADH were purchased from Sigma. The reactions were carried out at 35°C in a reaction volume of 50 µl, and absorbance was measured at 340nm every 2 minutes using the absorbance mode in the Synergy H4 microplate Reader. The reaction mixture comprised of 25 mM HEPES (pH 7.5), 500 mM NaCl, 10 mM $MgCl_2$, 1 mM ATP, 1 mM PEP, 500 µM NADH, 400 µM peptide substrate (NH2-KQWDNYEFIW-COOH, obtained from Biomatik corporation), and 100-200 nM of EphA3. A no-peptide control was included to account for any ATPase activity or auto-phosphorylation exhibited by the protein. The amount of protein used was optimized such that the kinase reaction was at least 25 times slower than the coupled PK/LDH reaction. Specific activity for the proteins was calculated by taking a ratio of the Vmax and the concentration of enzymes used in the assay.

Molecular dynamics simulations

All-atom MD simulations were performed using GROMACS (5.0.2) (2) and resources at the Georgia Advanced Computing Resource Center (GACRC), University of Georgia. The starting conformation of EphA3 in the autoinhibited state was modeled from crystal structure 2QO2. The two disordered regions (2 residues in the juxtamembrane and 20 residues in the activation loop) were modeled through a fragment-based loop modeling procedure using Rosetta (3.8)(3). The W826P mutation was introduced in the modeled WT autoinhibited structure using the loop refine module in Modeller.(9.19)(4). The protein was parameterized with amber99SB-ildn force field and solvated with TIP3P water model in a dodecahedron box that was 1 nm bigger than the protein on all sides (5). The system was then neutralized with 0.1M sodium chloride. To take advantage of the latest GPU acceleration, the Verlet cutoff scheme is used to maintain the neighbor list (6). The long-range electrostatics is calculated using particle mesh ewald (PME) algorithms. The resulting simulation system contains approximately 50000 atoms. Energy minimization was performed by coupling steepest descent and conjugate-gradient algorithm until the Fmax was less than 100 kcal/mol. Atom velocities were generated according to a Maxwell distribution at 310K. Canonical (NVT) ensemble simulations were carried out by heating from 0 to 310 K using velocity rescaling for 200 ps with a timestep of 2 fs (7). Isothermal-isobaric (NPT) ensemble (P = 1 bar, T = 310 K) were carried out using the berendsen barostat to maintain pressure and density of the system. Position restraint was applied to heavy atoms of the protein during the NVT and NPT equilibration steps. The unrestrained MD productions were run for 500 ns from the NPT ensemble. Analysis of MD trajectories were done using packages from the Gromacs suite, and the cross-correlation matrices of Cα-Cα motions were generated using Prody (8).

1. Yang, J., Zappacosta, F., Annan, R. S., Nurse, K., Tummino, P. J., Copeland, R. A., and Lai, Z. (2009) The catalytic role of INCENP in Aurora B activation and the kinetic mechanism of Aurora B/INCENP. *Biochem J* **417**, 355-360
2. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25
3. Rohl, C. A., Strauss, C. E. M., Misura, K. M. S., and Baker, D. (2004) Protein structure prediction using rosetta. *Numerical Computer Methods, Pt D* **383**, 66-+
4. Fiser, A., Do, R. K., and Sali, A. (2000) Modeling of loops in protein structures. *Protein Sci* **9**, 1753-1773
5. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950-1958
6. Páll, S., and Hess, B. (2013) A flexible algorithm for calculating pair interactions on SIMD architectures. *Computer Physics Communications* **184**, 2641-2650
7. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J Chem Phys* **126**, 014101
8. Bakan, A., Meireles, L. M., and Bahar, I. (2011) ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575-1577