# Supporting Information

# Insights into Aurora-A Kinase Activation Using Unnatural Amino Acids Incorporated by Chemical Modification

Fiona C. Rowan,[†,‡] Meirion Richards,[†] Rachel A. Bibby,[‡] Andrew Thompson,[§] Richard Bayliss,*[,‡,‖] and Julian Blagg*[,†]

[†] Cancer Research UK Cancer Therapeutics Unit, Division of Cancer Therapeutics, The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, U.K.

[‡] Division of Structural Biology and [§]Proteomics Core Facility, The Institute of Cancer Research, Chester Beatty Laboratories, Chelsea, London SW3 6JB, U.K.

[‖] Department of Biochemistry, University of Leicester, Henry Wellcome Building, Lancaster Road, Leicester LE1 9HN, U.K.

**E-mail: julian.blagg@icr.ac.uk; richard.bayliss@le.ac.uk*

## SUPPLEMENTARY FIGURES

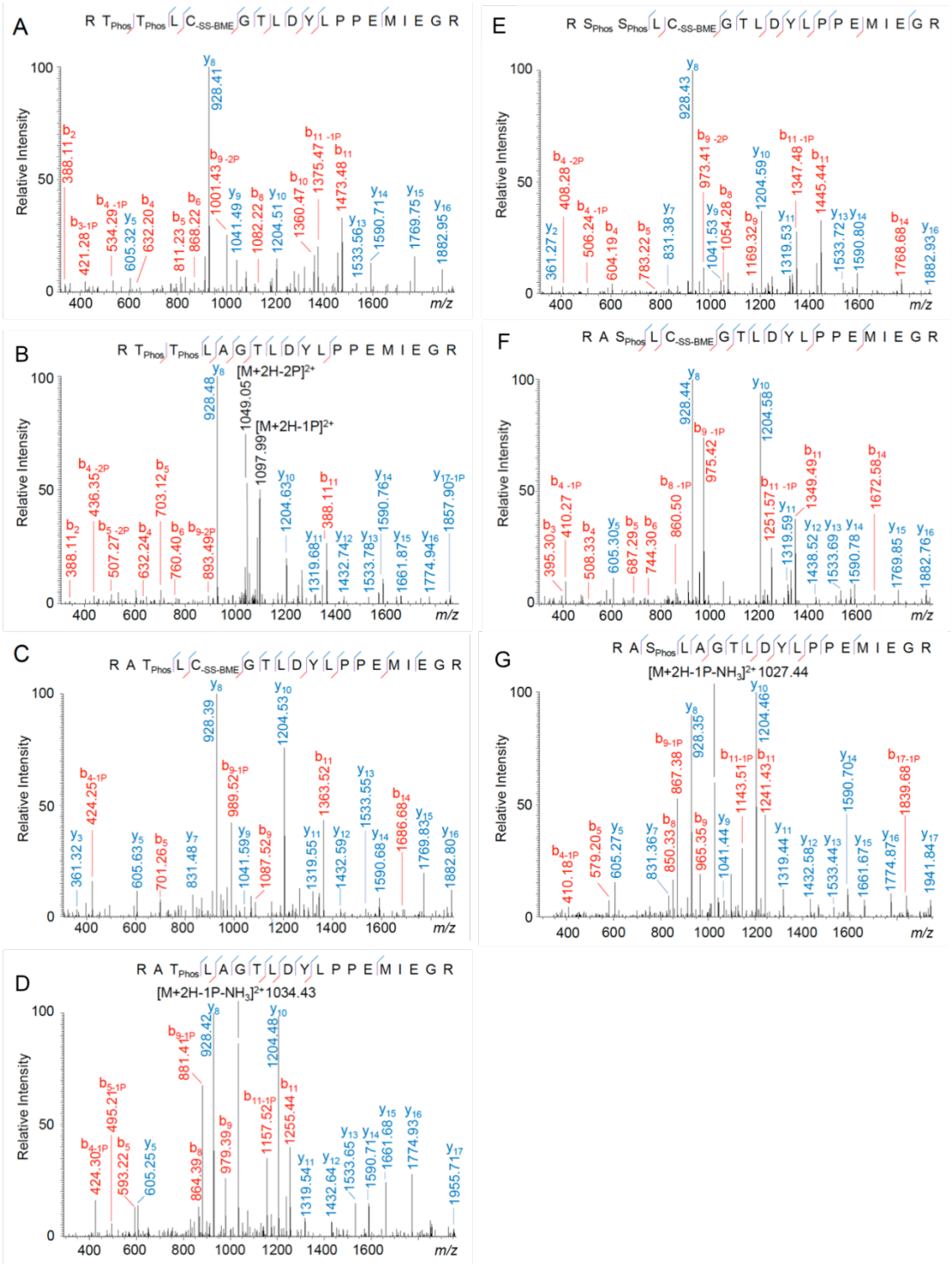



**Figure S1.** LC-MS/MS spectra indicating phosphorylation of residues (A) T287 and T288 in the wild-type sample (precursor ion 1,147.0198$^{2+}$, error -0.71 ppm, ion score 47.6), (B) T287 and T288 in the pseudo wild-type sample (precursor ion 1,163.0056$^{2+}$, error -0.89 ppm, ion score 56.9), (C) T288 in the AT sample (precursor ion 1,146.0164$^{2+}$, error -0.84 ppm, ion score 60.7), (D) T288 in the pseudo AT sample (precursor ion: 1,092.0322$^{2+}$, error 0.095 ppm, ion score 86.4), (E) S287 and S288 in the SS

Sample (precursor ion 1,186.9898$^{2+}$, error -0.28 ppm, ion score 41.9), (F) S288 in the AS sample (precursor ion 1,139.0090$^{2+}$, error -1.1 ppm, ion score 70.2), and (G) S288 in the pseudo AS sample (precursor ion 1,085.0233$^{2+}$, error -0.88 ppm, ion score 103.4). Neutral losses of phosphorylation (-98 Da) are marked as '-nP' and prominent b- and y- fragment ions are flagged for each peptide sequence. Cysteines were protected as disulfide BME modified residues (–SS-BME).

MHHHHHHSSGLVPRGSGMKETAAAKFEENLYFQGAMESKKRQWALEDFEIGRPLGKG
KFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGY
FHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRD
IKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLW
SLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQ
RPMLREVLEHPWITANSSKPSNCQNKESASKQS

**Figure S2**. Sequence of AurA kinase domain 122–403, with N-terminal His-tag (underlined). T287 and T288, sites subject to the instalment of phospho-mimics, are highlighted yellow. The surface cysteines mutated to alanine for pseudo wild-type templates are highlighted blue.

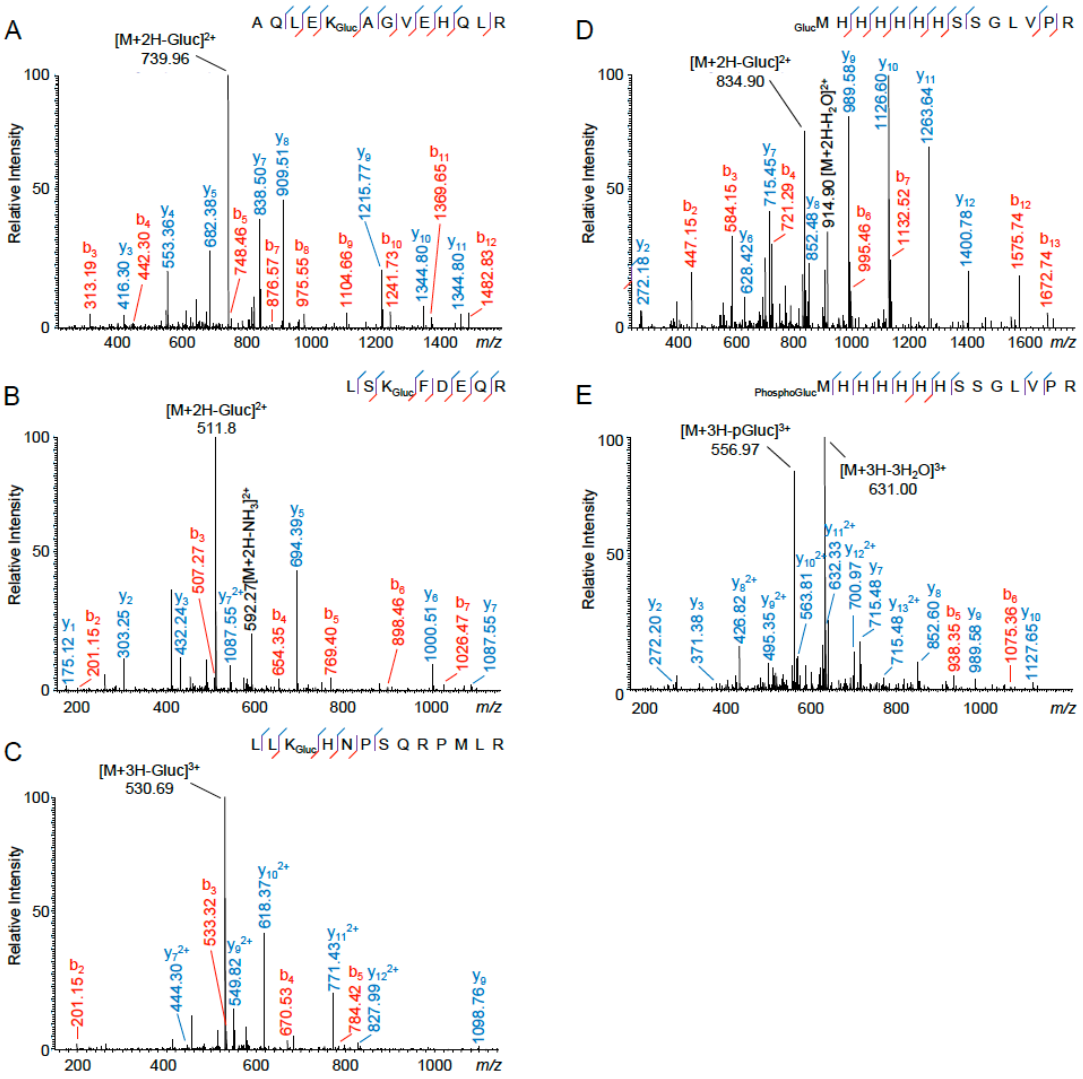



**Figure S3**. LC-MS/MS spectra indicating (A-C) lysine gluconoylation of residues K171 (precursor $828.9308^{2+}$, mass error -0.60 ppm, ion score 46.2), K227 (precursor $600.7906^{2+}$, mass error -0.41 ppm, ion score 30.9), and K365 (precursor $589.9894^{3+}$, mass error -0.0015 ppm, ion score 13.7), respectively, (D) N-terminal gluconoylation (precursor $923.9295^{2+}$, mass error 0.19 ppm, ion score 53.7, and (E) N-terminal phosphogluconoylation (precursor $642.9436^{3+}$, mass error -0.76 ppm, ion score 21.0). Neutral losses of gluconoylation (-178 Da) and phosphogluconoylation (-258 Da) are marked and prominent b- and y- fragment ions are flagged for each peptide sequence.

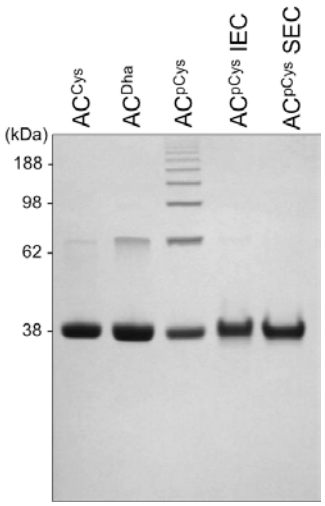



**Figure S4**. SDS-PAGE (4–12% acrylamide) of AurA AC variant at each step of the phospho-cysteine chemical reaction. IEC: sample after ion exchange chromatography. SEC: sample after size exclusion chromatography. Prior to loading on the gel, samples were boiled in SDS loading buffer without reducing agent.

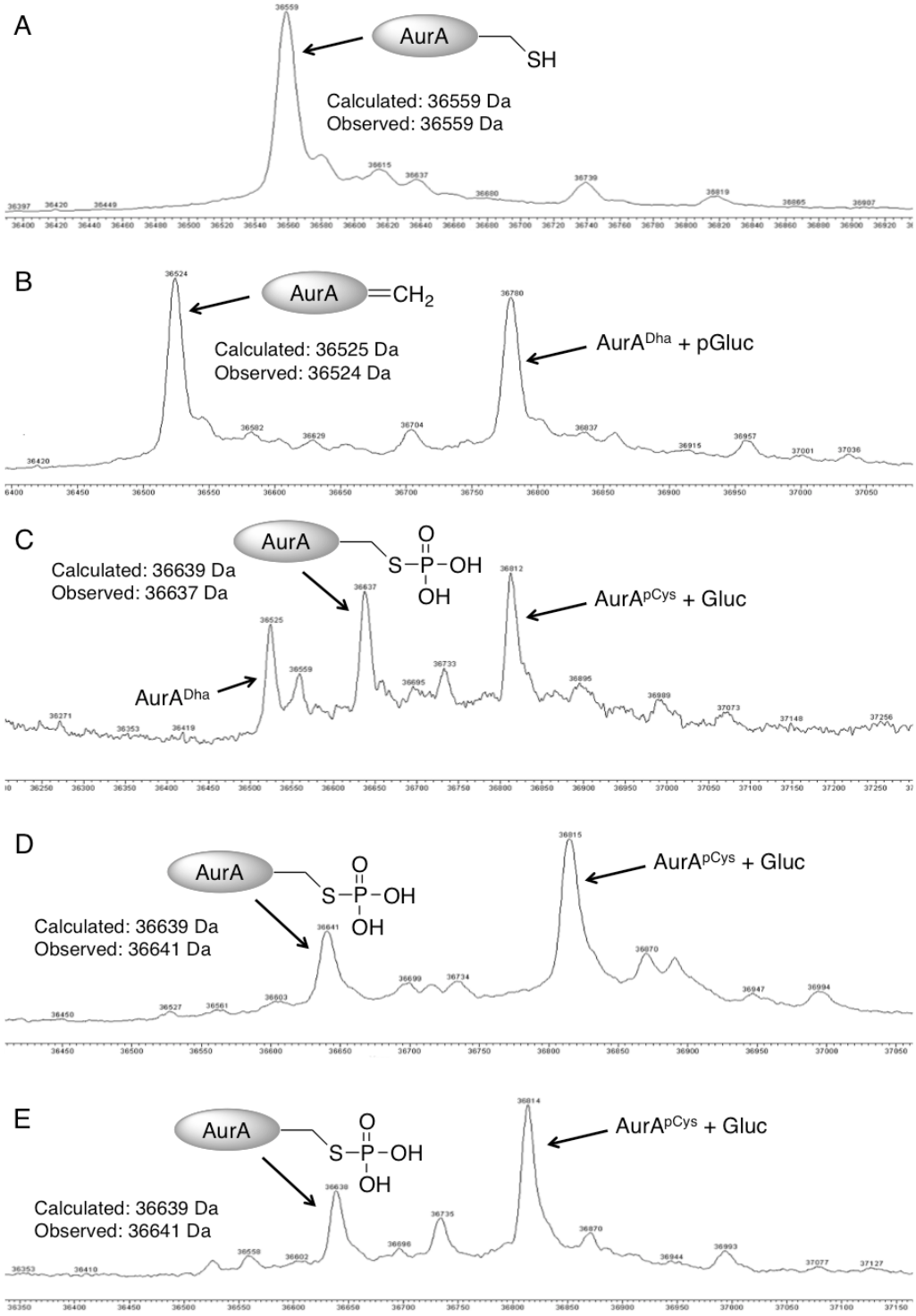



**Figure S5**. Intact protein mass spectra illustrating conversion of cysteine to phospho-cysteine on AurA kinase domain. Cysteine (A) reacts completely to dehydroalanine (B). This undergoes partial conversion to phospho-cysteine (C). The desired product (D) is isolated by high-resolution anion exchange chromatography. The reaction yield is calculated by measuring the total protein recovery of pure AurA$^{pCys}$ (i.e. after IEC and SEC) and comparing against the amount of starting material, as it is not feasible to accurately quantitate Dha to Cys conversion (C) by intact mass spectrometry. The protein remains unchanged after the kinase activity assay (E).

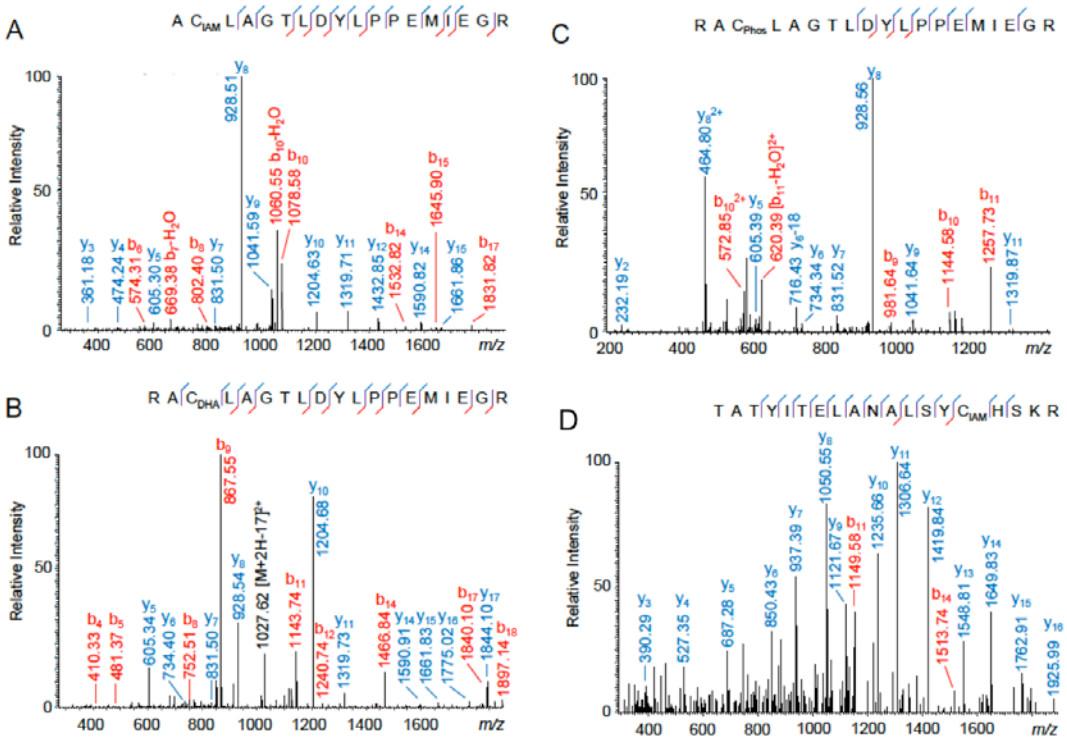
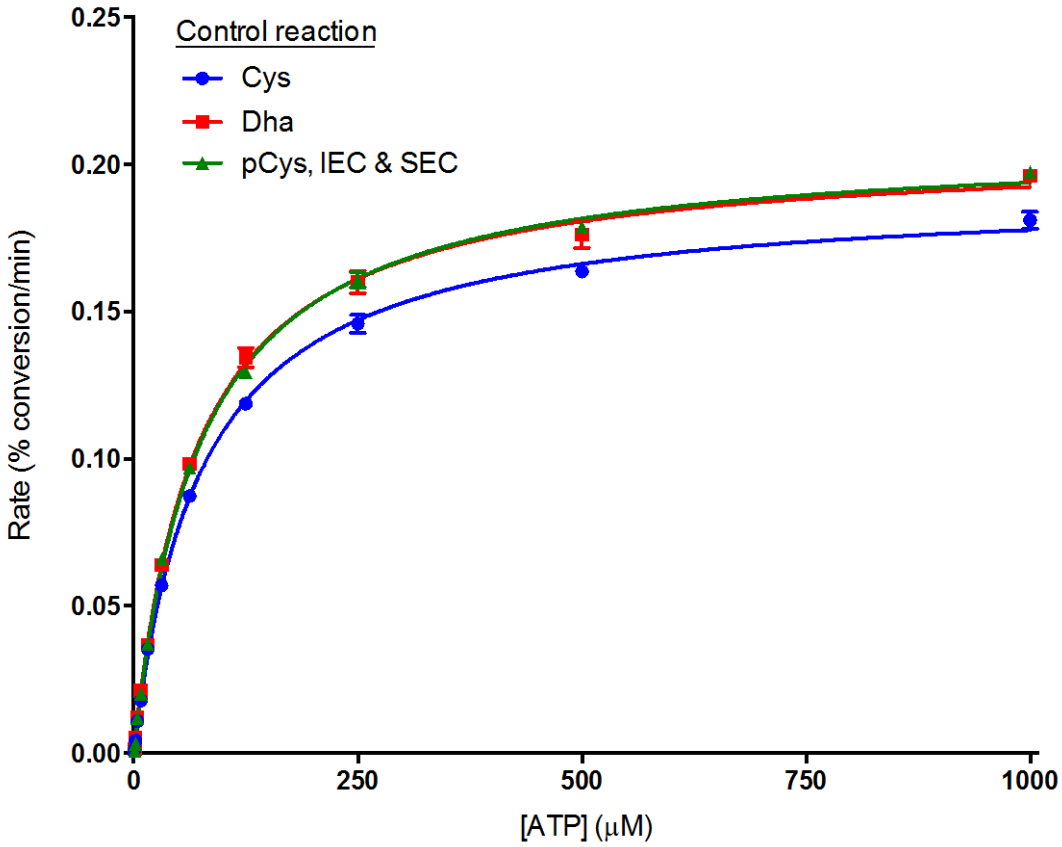



**Figure S6**. LC-MS/MS spectra indicating (A) mutation of T288C, detected as the iodo-acetamide alkylated cysteine (precursor $1003.4893^{2+}$, mass error -0.46 ppm, ion score 79.7), (B) conversion of T288C to dehydroalanine (precursor $1036.0349^{2+}$, mass error -0.95 ppm, ion score 43.6), (C) phospho-cysteine (precursor $729.0111^{3+}$, mass error 0.025 ppm, ion score 27.6), and (D) Confirmation of non-mutation and non-reaction of the buried cysteine C247 (precursor $1100.0459^{2+}$, -0.81 ppm error, ion score 86.5).




| Sample | ATP $K_m$ (μM) | $V_{max}$ (% conversion/min) |
|---|---|---|
| Cys control reaction | 74 ± 2 | 0.191 ± 0.002 |
| Dha control reaction | 69 ± 2 | 0.206 ± 0.002 |
| pCys control reaction, IEC & SEC | 72 ± 2 | 0.208 ± 0.002 |

**Figure S7**. AurA kinase activity following phospho-cysteine control reactions. AurA pseudo AT was subjected to the reaction and purification conditions used to produce and isolate AC and CA phospho-cysteine variants. The conditions do not affect Michaelis-Menten kinetic parameters. AurA pseudo AT control samples were all assayed at 4 nM.

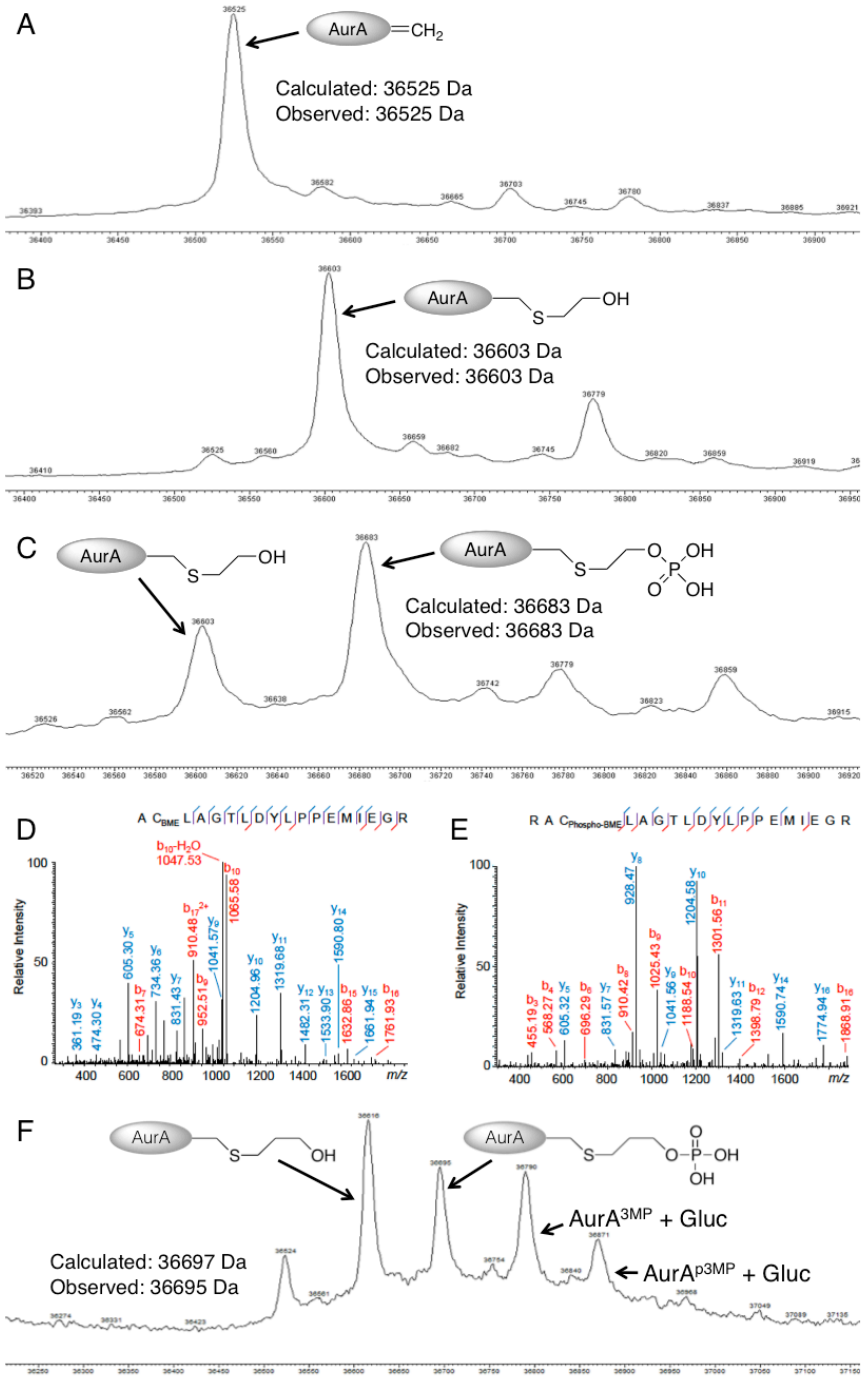



**Figure S8**. Intact and LC-MS/MS mass spectra indicating auto-phosphorylation of 2-hydroxyethylcysteine on AurA kinase domain AC construct. Dehydroalanine (A) reacts with BME to yield 2-hydroxyethylcysteine (B), which becomes phosphorylated in the presence of ATP (C). LC-MS/MS spectra indicate (D) 2-hydroxyethylcysteine (precursor $996.99^{2+}$, mass error -1.5 ppm, ion score 78.4), and (E) phospho-2-hydroxyethylcysteine (precursor $1{,}115.0250^{2+}$, mass error -0.86 ppm, ion score 40.6). (F) AurA AC modified to contain a longer serine mimic, 3-hydroxypropylcysteine, autophosphorylates to yield phospho-3-hydroxypropylcysteine.

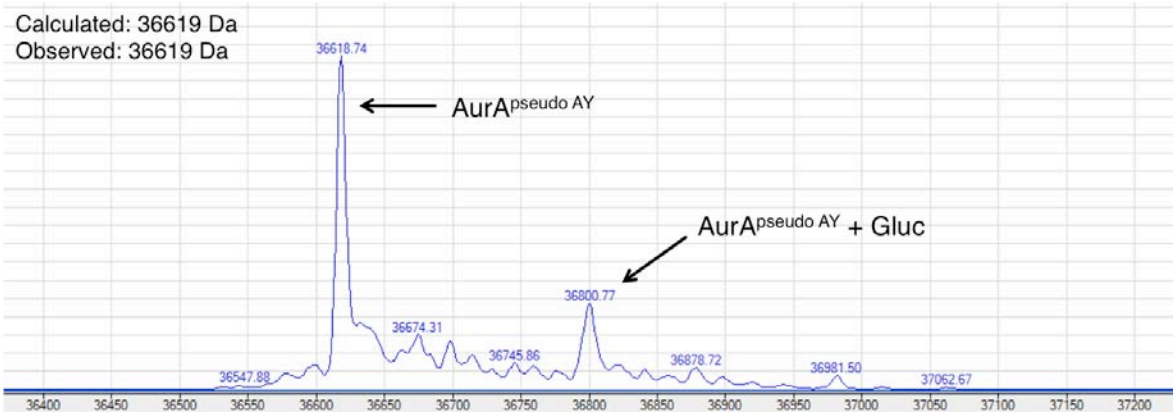



**Figure S9**. Intact protein mass spectrum of AurA pseudo AY, indicating complete absence of autophosphorylation after the kinase activity assay.

## MASS SPECTROMETRY METHODS

**MS/MS for intact proteins.** 1 µl injections of the sample were made onto a Security Guard C8 column cartridge (4 x 3 mm, AJO-4290, Phenomenex). Samples were refrigerated at 4°C in a G1367B auto-sampler (Agilent) with G1330B thermostat module prior to injection. Chromatographic separation at 60°C was carried out using an Agilent 1200 Series HPLC over a 1 minute gradient elution of 95:5 to 5:95 water and acetonitrile (both modified with 0.1% formic acid) at a flow rate of 0.5 ml/min. The post column eluent flow was infused into an Agilent 6520 Series qToF mass spectrometer fitted with a dual ESI ionisation source. An initial divert to waste was used to aid desalting. LC eluent and nebulising gas was introduced into the grounded nebuliser with spray direction orthogonal to the capillary axis. The aerosol was dried by heated gas (10 L/min of nitrogen at 350°C, 50 psi), producing ions by ESI. Ions entered the transfer capillary along which a potential difference of 4kV was applied. The fragmentor voltage was set at 190V and skimmer at 65V. Signal was optimised by AutoTune.m. Profile mass spectrometry data were acquired in positive ionisation mode over a scan range of m/z 650-2000 (scan rate 1.0) with reference mass correction at m/z 922.009798 hexakis(1H,1H,3H-perfluoropropoxy)phosphazene. Raw data were processed using Agilent MassHunter Qualitative Analysis B.06.00 and MagTran 1.02.[1]

**LC-MS/MS for Trypsin-digested proteins.** In initial experiments, AurA samples in Tris buffer (typically 50 mM Tris pH 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 10 mM BME, 10% glycerol) were reduced by incubation with 5 mM tris(2-carboxyethyl)phosphine for 20 mins at 37°C, alkylated with 55 mM iodoacetamide at ambient temperature in the dark, digested with modified porcine trypsin (Promega) for 4 hours at 37°C and the reaction quenched by addition of trifluoroacetic acid (2% final concentration). Reduction and alkylation was observed in some cases to interfere with elucidating the status of key cysteine modifications, and therefore subsequent samples were directly digested with trypsin, without prior reduction and alkylation. In all cases, digested samples were purified using ziptips (Millipore) essentially according to manufacturer instructions and the dried eluates reconstituted in 0.1% formic acid for LC-MS/MS analysis. Reversed phase chromatography was performed using an HP1200 platform (Agilent) coupled to an LTQ Velos Orbitrap mass spectrometer (Thermo Fisher Scientific) in one of two configurations; Configuration 1: Peptides were resolved on a 75 µm I.D. 10 cm C18 Pepmap column (3 µm particle size; LC Packings/Dionex) over 30 mins using a linear gradient of 96:4 to 50:50 buffer A:B (buffer A: 2% acetonitrile/0.1% formic acid; buffer B: 80% acetonitrile/0.1% formic acid) at 300 nl/min. Peptides were ionised by electrospray ionisation using 1.7 kV applied directly to the post-column LC eluent via a microtee built into the nanospray source, and sample was infused into the mass spectrometer using a non-coated SilicaTip emitter (20 µm I.D., 10 µm tapered tip; New Objectives); Configuration 2: Peptides were resolved on a 75 µm I.D. 15 cm C18 packed emitter column (3 µm particle size; Nikkyo Technos Co., Ltd.) using the same gradient conditions as above at 250 nl/min. Peptides were ionised by electrospray ionisation using 1.8 kV applied immediately pre-column to the packed emitter via a microtee built into the nanospray source. Sample was infused into the mass spectrometer directly from the end of the tapered tip silica column

(6-8 µm tapered tip). For both configurations, the ion transfer tube was heated to 200°C and the S-lens set to 60%. MS/MS were acquired using data dependent acquisition based on a full 30,000 resolution FT-MS scan to sequence the top 10 most intense ions using enhanced ion trap scans. Automatic gain control was set to 1,000,000 for FT-MS and 30,000 for IT-MS/MS, full FT-MS maximum inject time was 500 ms and normalised collision energy was set to 35% with an activation time of 10 ms. Multistage activation was used to target phospho-serine/threonine peptides by co-fragmenting precursor ions undergoing neutral loss of 32.70, 49.00, 65.40 and 98.00 m/z, corresponding to neutral loss of phosphate, if observed in the top 3 most intense fragment ions. MS/MS was acquired for selected precursor ions with a single repeat count followed by dynamic exclusion with a 10 ppm mass window for 15 s duration for the experiments using the C18 Pepmap column, or 10 s duration for the packed emitter column, based on a maximal exclusion list of 500 entries. Raw MS/MS data were compiled into peaklists using Proteome Discoverer v1.3 (Thermo Fisher Scientific) default parameters. Peaklists were interrogated against a SwissProt 2011_01 *Homo sapiens* subset database (20,282 sequences), customised to include AurA construct sequences, using Mascot v2.2 (www.matrixscience.com) and assuming tryptic enzyme specificity with up to two missed cleavages. A precursor ion tolerance of 5 ppm and fragment ion tolerance of 0.25 Da was applied and the following variable modifications were accounted for depending on the sample in question: acetylation of the protein N-terminus, carbamidomethylation of cysteine, oxidation of methionine, pyroglutamisation of peptide N-terminal glutamine, phosphorylation of serine, threonine and tyrosine, gluconoylation and phosphogluconoylation of the protein N-terminus and lysine residues, mutation of threonine to alanine, mutation of threonine to cysteine, conversion of threonine-cysteine mutations to dehydroalanine, phospho-cysteine, -SS-phospho-cysteine, BME-cysteine, phosphorylation of BME-cysteine, and reductive disulfide linkage of BME to cysteine. The Mascot peptide and protein identification results were grouped and validated using Scaffold v3.6 (Proteome Software Inc.). Protein identifications were automatically accepted if they contained at least 2 unique peptides assigned with at least 95% confidence by Peptide Prophet.[2] To confirm sequence modifications including phosphorylation, mutations and chemical derivitisations, the peptide confidence threshold was reduced to 50% and candidate spectra were visually assessed.

## STRUCTURE DETERMINATION OF AURA$^{V174M}$

The catalytic domain of AurA$^{V174M}$ (residues 122-403) was expressed in methionine auxotroph *E. coli* B834 cells co-transformed with the CodonPlus -RIL plasmid grown in Seleno-Met Medium (Molecular Dimensions). The protein was purified to homogeneity as previously described for wild-type AurA and concentrated to 10 mg/ml. The protein was supplemented with 2 mM ATPγS and 10 mM each of two peptides sequence (LRRWSLGA and SYSYDAPSDFINFS). Barrel-shaped crystals were grown by vapor diffusion using a well solution comprising 0.1 M HEPES pH 7.5, 10% PEG 6000, 5% MPD and 0.1 M calcium chloride. This condition was obtained from initial screening against a focused set of 71 conditions in which Aurora kinase crystals have previously been described, followed by use of the Additive Screen (Hampton Research) to improve crystal morphology. A crystal was washed in the same buffer, supplemented with 20% glycerol, and flash-frozen in liquid nitrogen. A native dataset was collected at DIAMOND beamline I03. The structure was solved by molecular replacement using the structure of unphosphorylated AurA (1OL6) as a model. There was no apparent electron density due to either peptide, and the ATPyS appears to have been hydrolyzed to ADP (as previously observed in PDB code 1OL5), which is complexed with two $Ca^{2+}$ ions in place of $Mg^{2+}$. Iterative rounds of model building (using COOT[3]) and refinement (using PHENIX[4]) were carried out to complete the structure. The geometry of the final model was checked using MOLPROBITY.[5]

**Table S1.** Summary of crystallographic analysis

| | |
|---|---|
| *Lattice parameters* | |
| Space group | $P6_122$ |
| $a$ (Å) | 81.62 |
| $b$ (Å) | 81.62 |
| $c$ (Å) | 175.39 |
| $\alpha$ (°) | 90 |
| $\beta$ (°) | 90 |
| $\gamma$ (°) | 120 |
| *Data collection at Diamond I03* | |
| Wavelength (Å) | 0.9721 |
| Resolution range | 70.71-2.50 |
| (highest-resolution shell) (Å) | (2.64-2.50) |
| Unique reflections | 12693 (1798) |
| Completeness (%) | 100.0 (100.0) |
| Multiplicity | 11.1 (10.3) |
| $R_{merge}$ (%) | 11.9 (62.7) |
| $R_{pim}$ (%) | 3.3 (19.2) |
| $I/\sigma(I)$ | 16.3 (3.3) |
| *Refinement* | |
| Resolution range (Å) | 65.56-2.50 |
| $R_{factor}$ (%) | 19.19 |
| $R_{free}$ (%) | 24.41 |
| Bond deviation (Å) | 0.009 |
| Angle deviation (°) | 1.34 |
| *MolProbity analysis* | |
| All-atom Clashscore | 25.55 |
| Rotamer outliers (%) | 3.2 |
| Ramachandran outliers (%) | 0.4 |
| Ramachandran favoured (%) | 95.5 |

## SUPPLEMENTARY REFERENCES

(1) Zhang, Z., and Marshall, A. G. (1998) A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra, *J. Am. Soc. Mass Spectrom. 9*, 225–233.

(2) Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search, *Anal. Chem. 74*, 5383–5392.

(3) Emsley, P. and Cowtan K (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126–2132.

(4) Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchetini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination, *Acta Crystallogr. D Biol. Crystallogr. 58*, 1948–1954.

(5) Chen, V.B., Arendall III, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr. 66*, 12–21.